Notice

Medicine is an ever-changing science. As new research and clinical experience broaden our knowledge, changes in treatment and drug therapy are required. The authors and the publisher of this work have checked with sources believed to be reliable in their efforts to provide information that is complete and generally in accord with the standards accepted at the time of publication. However, in view of the possibility of human error or changes in medical sciences, neither the authors nor the publisher nor any other party who has been involved in the preparation or publication of this work warrants that the information contained herein is in every respect accurate or complete, and they disclaim all responsibility for any errors or omissions or for the results obtained from use of the information contained in this work. Readers are encouraged to confirm the information contained herein with other sources. For example and in particular, readers are advised to check the product information sheet included in the package of each drug they plan to administer to be certain that the information contained in this work is accurate and that changes have not been made in the recommended dose or in the contraindications for administration. This recommendation is of particular importance in connection with new or infrequently used drugs.
Contents

Volume 1

Part I

Basic Considerations 1

1. Leadership in Surgery ........................................ 3
   Stephen Markowiak, Hollis Merrick, Shiela Beroukhim,
   Jeremy J. Laukka, Amy Lightner, Munier Nazzal,
   Lee Hammerling, James R. Macho, and F. Charles Brunicardi

2. Systemic Response to Injury and Metabolic Support ............. 27
   Siobhan A. Corbett

3. Fluid and Electrolyte Management of the Surgical Patient .......... 83
   Matthew D. Neal

4. Hemostasis, Surgical Bleeding, and Transfusion .................. 103
   Ronald Chang, John B. Holcomb, Evan Leibner,
   Matthew Pommery, and Rosemary A. Kozar

5. Shock ............................................................ 131
   Brian S. Zuckerbraun, Andrew B. Peitzman, and Timothy R. Billiar

6. Surgical Infections ............................................. 157
   Robert E. Bulander, David L. Dunn, and
   Greg J. Beilman

7. Trauma ........................................................... 183
   Clay Cothren Burlew and Ernest E. Moore

8. Burns .............................................................. 251
   Jeffrey H. Anderson, Samuel P. Mandell, and
   Nicole S. Gibran

9. Wound Healing .................................................. 271
   Munier Nazzal, Mohammad F. Osman,
   Heitham Albeshri, Darren B. Abbas, and Carol A. Angel

10. Oncology ......................................................... 305
    William E. Carson III, Funda Meric-Bernstam, and
    Raphael E. Pollock

11. Transplantation .................................................. 355
    David L. Dunn, Angelika C. Gruessner, and
    Rainer W.G. Gruessner

12. Quality, Patient Safety, Assessments of Care, and Complications .... 397
    Martin A. Makary, Peter B. Angood, and
    Mark L. Shapiro

Part II

Specific Considerations 511

13. Physiologic Monitoring of the Surgical Patient ..................... 433
    Anthony R. Cyr and Louis H. Alarcon

14. Minimally Invasive Surgery, Robotics, Natural Orifice Transluminal
    Endoscopic Surgery, and Single-Incision Laparoscopic Surgery ...... 453
    Donn H. Spight, Blair A. Jobe, and John G. Hunter

15. Molecular Biology, The Atomic Theory of Disease, and Precision Surgery ... 479
    Xin-Hua Feng, Xia Lin, Xinran Li, Juehua Yu,
    John Nemunaitis, and F. Charles Brunicardi

Specific Considerations

16. The Skin and Subcutaneous Tissue .................................. 513
    Patrick Harbour and David H. Song

17. The Breast .................................................................. 541
    Catherine C. Parker, Senthil Damodaran,
    Kirby J. Bland, and Kelly K. Hunt

18. Disorders of the Head and Neck ...................................... 613
    Antoine Eskander, Stephen Y. Kang,
    Michael S. Harris, Bradley A. Otto, Oliver Adunka,
    Randal S. Weber, and Theodoros N. Teknos

19. Chest Wall, Lung, Mediastinum, and Pleura ......................... 661
    Katie S. Nason, Rose B. Ganin, and James D. Luketich

20. Congenital Heart Disease ............................................ 751
    Ragav Murthy, Tbilitha G. Moe, Glen A. Van Arsdell,
    John J. Nigro, and Tara Karamlou

21. Acquired Heart Disease ............................................... 801
    Matthew R. Schill, Ali J. Khiabani, Puja Kachroo,
    and Ralph J. Damiano Jr

22. Thoracic Aneurysms and Aortic Dissection .......................... 853
    Scott A. Lefair, Ourania Preventza, and Joseph S. Coselli

23. Arterial Disease ....................................................... 897
    Peter H. Lin, Carlos F. Bechara, Changyi Chen,
    and Frank J. Veith

24. Venous and Lymphatic Disease ....................................... 981
    Atish Chopra, Timothy K. Lien,
    and Gregory L. Moneta

25. Esophagus and Diaphragmatic Hernia ................................ 1009
    Blair A. Jobe, John G. Hunter, and David I. Watson

26. Stomach ............................................................... 1099
    Robert E. Roses and Daniel T. Dempsey
Volume 2

27. The Surgical Management of Obesity ..........1167
   Anita P. Courcoulas and Philip R. Schauer

28. Small Intestine ........................................ 1219
   Ali Tavakkoli, Stanley W. Ashley, and Michael J. Zinner

29. Colon, Rectum, and Anus ............................ 1259
   Mary R. Kwaan, David B. Stewart Sr, and Kelli Bullard Dunn

30. The Appendix .......................................... 1331
   Fadi S. Dahdah, David Heidt, and Kiran K. Turaga

31. Liver .................................................... 1345
   David A. Geller, John A. Goss, Ronald W. Busuttil, and Allan Tsung

32. Gallbladder and the Extrahepatic Biliary System .................................................. 1393
   Kelly R. Haisley and John G. Hunter

33. Pancreas .................................................. 1429
   William E. Fisher, Dana K. Andersen, John A. Windsor, Vikas Dudeja, and F. Charles Brunicardi

34. The Spleen ................................................. 1517
   Adrian E. Park, Eduardo M. Targarona, Adam S. Welz, and Carlos Rodriguez-Otero Luppi

35. Abdominal Wall, Omentum, Mesentery, and Retropertioneum ...................................... 1549
   Scott Kizy and Sayeed Ikramuddin

36. Soft Tissue Sarcomas ................................. 1567
   Ricardo J. Gonzalez, Alessandro Gronchi, and Raphael E. Pollock

37. Inguinal Hernias ........................................ 1599
   Chandan Das, Tahir Jamil, Stephen Stanek, Ziya Baghmanli, James R. Macho, Joseph Sferra, and F. Charles Brunicardi

38. Thyroid, Parathyroid, and Adrenal .............. 1625
   Geeta Lal and Orlo H. Clark

39. Pediatric Surgery ....................................... 1705
   David J. Hackam, Jeffrey Upperman, Tracy Girkischelt, Kasper Wang, and Henri R. Ford

40. Urology .................................................... 1759
   Ahmad Shahsigh, Michael Sourial, Farah F. Bellow, Christopher McLung, Rama Jayanthi, Stephanie Kielb, Geoffrey N. Box, Bodo E. Knudsen, and Cheryl T. Lee

41. Gynecology ............................................... 1783
   Sarah M. Temkin, Thomas Gregory, Elise C. Kohn, and Linda Duska

42. Neurosurgery ............................................ 1827
   Ashwin G. Ramayya, Saurabh Sinha, and M. Sean Grady

43. Orthopedic Surgery .................................... 1879
   Nabil A. Ebraheim, Bert J. Thomas, Freddie H. Fu, Bart Muller, Dharmesh Vyas, Matt Niesen, Jonathan Pribaz, and Klaus Draenert

44. Surgery of the Hand and Wrist ..................... 1925
   Scott D. Lifchez and Brian H. Cho

45. Plastic and Reconstructive Surgery ............. 1967
   Rajiv Y. Chandawarkar, Michael J. Miller, Brian C. Kellogg, Steven A. Schulz, Ian L. Valerio, and Richard E. Kirschner

46. Anesthesia for Surgical Patients .................. 2027
   Junaid Nizamuddin and Michael O’Connor

47. Surgical Considerations in Older Adults ........ 2045
   Anne M. Suskind and Emily Finlayson

48. Ethics, Palliative Care, and Care at the End of Life .................................................. 2061
   Daniel E. Hall, Eliza W. Beal, Peter A. Angelos, Geoffrey P. Dunn, Daniel B. Hinshaw, and Timothy M. Pavlik

49. Global Surgery ........................................... 2077
   Katherine E. Smiley, Haile T. Debis, Catherine R. deVries, and Raymond R. Price

50. Optimizing Perioperative Care: Enhanced Recovery and Chinese Medicine ................... 2113
   Catherine R. deVries, and Raymond R. Price

51. Understanding, Evaluating, and Using Evidence for Surgical Practice ......................... 2137
   Andrew J. Benjamin, Andrew B. Schneider, Jeffrey B. Matthews, and Gary An

52. Ambulatory Surgery .................................... 2153
   Marcus Adair, Stephen Markowiak, Hollis Merrick, James R. Macho, Kara Richardson, Moriah Masculo, Munier Nazzal, and F. Charles Brunicardi

53. Skills and Simulation .................................. 2163
   Neal E. Seymour and Carla M. Pugh

54. Web-Based Education and Implications of Social Media ............................................ 2187
   Lillian S. Kao and Michael E. Zenilman

Index/2197
Contributors

Darren B. Abbas, MD
Clinical Instructor
Department of Surgery
University of Toledo College of Medicine and Life Sciences
Toledo, Ohio
Chapter 9, Wound Healing

Marcus Adair, MD
Clinical Instructor
Department of Surgery
University of Toledo College of Medicine and Life Sciences
Toledo, Ohio
Chapter 52, Ambulatory Surgery

Oliver Adunka, MD, FACS
Professor
Vice-Chair, Clinical Operations
Department of Otolaryngology-Head and Neck Surgery
Director, Division of Otology/Neurotology and Cranial Base Surgery
The Ohio State University-James Comprehensive Cancer Center
Columbus, Ohio
Chapter 18, Disorders of the Head and Neck

Louis H. Alarcon, MD, FACS, FCCM
Professor of Surgery and Critical Care Medicine
Medical Director, Division of Trauma Surgery
University of Pittsburgh
Pittsburgh, Pennsylvania
Chapter 13, Physiologic Monitoring of the Surgical Patient

Heitham Albeshri, MD
Clinical Instructor
Department of Surgery
University of Toledo College of Medicine and Life Sciences
Toledo, Ohio
Chapter 9, Wound Healing

Gary An, MD
Professor of Surgery
Department of Surgery
The University of Chicago Medicine
Chicago, Illinois
Chapter 51, Understanding, Evaluating, and Using Evidence for Surgical Practice

Dana K. Andersen, MD, FACS
Scientific Program Manager
Division of Digestive Diseases and Nutrition
National Institute of Diabetes and Digestive and Kidney Diseases
National Institutes of Health
Bethesda, Maryland
Chapter 33, Pancreas

Jeffrey H. Anderson, MD
Resident, Department of Surgery
Harborview Medical Center
Seattle, Washington
Chapter 8, Burns

Carol A. Angel, MD
Clinical Instructor
Department of Surgery
University of Toledo College of Medicine and Life Sciences
Toledo, Ohio
Chapter 9, Wound Healing

Peter A. Angelos, MD, PhD, FACS
Linda Kohler Anderson Professor of Surgery and Surgical Ethics
Chief, Endocrine Surgery
Associate Director
MacLean Center for Clinical Medical Ethics
The University of Chicago Medicine
Chicago, Illinois
Chapter 48, Ethics, Palliative Care, and Care at the End of Life

Peter B. Angood, MD, CPE, FRCS(C), FACS, MCCM
President and Chief Executive Officer
American Association for Physician Leadership
Tampa, Florida
Chapter 12, Quality, Patient Safety, Assessments of Care, and Complications

Glen S. Van Arsdell, MD
Chief, Pediatric Cardiac Surgery
Mattel Children’s Hospital
University of California
Los Angeles, California
Chapter 20, Congenital Heart Disease
Contributors

Stanley W. Ashley, MD
General & Gastrointestinal Surgery
Brigham and Women’s Hospital
Frank Sawyer Professor of Surgery
Harvard Medical School
Boston, Massachusetts
Chapter 28, Small Intestine

Ziya Baghmanli, MD
Clinical Instructor
Department of Surgery
University of Toledo Medical Center
Toledo, Ohio
Chapter 37, Inguinal Hernias

Eliza W. Beal
Clinical Instructor, Department of General Surgery
The Ohio State University
Columbus, Ohio
Chapter 48, Ethics, Palliative Care, and Care at the End of Life

Carlos F. Bechara, MD
Associate Professor of Surgery
Program Director, Vascular Surgery Fellowship
Department of Surgery
Loyola University Medical Center
Maywood, Illinois
Chapter 23, Arterial Disease

Greg J. Beilman, MD
Owen H and Sarah Davidson Wangensteen Chair of Experimental Surgery
University of Minnesota
Minneapolis, Minnesota
Chapter 6, Surgical Infections

Fara F. Bellows, MD
Clinical Assistant Professor
Department of Urology
The Ohio State University
Columbus, Ohio
Chapter 40, Urology

Andrew J. Benjamin, MD, MS
The University of Chicago Medicine
Chicago, Illinois
Chapter 51, Understanding, Evaluating, and Using Evidence for Surgical Practice

Shiela Beroukhim, MD
Clinical Instructor
Harbor-UCLA Medical Center
Torrance, California
Chapter 1, Leadership in Surgery

Timothy R. Billiar, MD, FACS
George Vance Foster Professor and Chair
Department of Surgery
University of Pittsburgh School of Medicine
Pittsburgh, Pennsylvania
Chapter 5, Shock

Kirby I. Bland, MD
Fay Fletcher Kerner Professor
The University of Alabama at Birmingham
Department of Surgery
Birmingham, Alabama
Chapter 17, The Breast

Geoffrey N. Box, MD
Assistant Professor
Department of Urology
The Ohio State University
Columbus, Ohio
Chapter 40, Urology

F. Charles Brunicardi, MD, FACS
John Howard Endowed Professor of Pancreatic Surgery
Chair, Department of Surgery
Director, Cancer Program
University of Toledo College of Medicine and Life Sciences
Academic Chief of Surgery
ProMedica Health System
Toledo, Ohio
Chapter 1, Leadership in Surgery
Chapter 15, Molecular Biology, The Atomic Theory of Disease, and Precision Surgery
Chapter 33, Pancreas
Chapter 37, Inguinal Hernias
Chapter 52, Ambulatory Surgery

Robert E. Bulander, Jr., MD, PhD
Assistant Professor of Surgery
University of Minnesota
Minneapolis, Minnesota
Chapter 6, Surgical Infections

Clay Cothren Burlew, MD, FACS
Professor of Surgery
Director, Surgical Intensive Care Unit
Program Director, Surgical Critical Care Fellowship
Program Director, Trauma & Acute Care Surgery Fellowship
The Ernest E. Moore Shock Trauma Center
Denver Health Medical Center
University of Colorado School of Medicine
Denver, Colorado
Chapter 7, Trauma

Ronald W. Busuttil, MD, PhD
William P. Longmire, Jr., Chair in Surgery
Professor and Executive Chair
Department of Surgery
David Geffen School of Medicine at University of California-Los Angeles
Los Angeles, California
Chapter 31, Liver
William E. Carson III, MD, FACS
Professor of Surgery and Vice Chair for Promotion and Tenure
The John B. and Jane T. McCoy Chair in Cancer Research
Interim Chief, Division of Surgical Oncology
Associate Director for Clinical Research
OSU Comprehensive Cancer Center
The Ohio State University
Columbus, Ohio
Chapter 10, Oncology

Rajiv Y. Chandawarkar, MD, MBA
Acting Chair
Department of Plastic and Reconstructive Surgery
The Ohio State University Wexner Medical Center
Columbus Ohio
Chapter 45, Plastic and Reconstructive Surgery

Ronald Chang, MD
Department of Surgery
University of Texas Houston
Houston, Texas
Chapter 4, Hemostasis, Surgical Bleeding, and Transfusion

Changyi Chen, MD, PhD
Professor of Surgery
Michael E. DeBakey Department of Surgery
Baylor College of Medicine
Houston, Texas
Chapter 23, Arterial Disease

Brian H. Cho, MD
Resident in Plastic Surgery
Johns Hopkins Department of Plastic Surgery
Baltimore, Maryland
Chapter 44, Surgery of the Hand and Wrist

Atish Chopra, MD
Fellow, Division of Vascular Surgery
Department of Surgery
Oregon Health & Science University
Portland, Oregon
Chapter 24, Venous and Lymphatic Disease

Orlo H. Clark, MD, FACS
Professor Emeritus, Department of Surgery
UCSF Mt Zion Medical Center
San Francisco, California
Chapter 38, Thyroid, Parathyroid, and Adrenal

Siobhan A. Corbett, MD
Associate Professor of Surgery
Department of Surgery
Rutgers-Robert Wood Johnson Medical School
Rutgers Biomedical and Health Sciences
New Brunswick, New Jersey
Chapter 2, Systemic Response to Injury and Metabolic Support

Joseph S. Coselli, MD
Vice-Chair, Michael E. DeBakey Department of Surgery
Professor, Cullen Foundation Endowed Chair
Chief, Division of Cardiothoracic Surgery
Baylor College of Medicine
Chief, Adult Cardiac Surgery Texas Heart Institute
Chief, Adult Cardiac Surgery Section
Associate Chief, Cardiovascular Service
Baylor St. Luke’s Medical Center
Houston, Texas
Chapter 22, Thoracic Aneurysms and Aortic Dissection

Anita P. Courcoulas, MD, MPH, FACS
Professor of Surgery
Section Head, MIS Bariatric & General Surgery
University of Pittsburgh Medical Center
Pittsburgh, Pennsylvania
Chapter 27, The Surgical Management of Obesity

Zhigang Cui, MD
Associate Professor of Surgery
Department of Surgery
Tianjin Medical University
Nankai Hospital
Institute of Acute Abdomen of Tianjin, Tianjin, China
Chapter 50, Optimizing Perioperative Care: Enhanced Recovery and Chinese Medicine

Anthony R. Cyr, MD, PhD
Clinical Instructor
Department of Surgery
University of Pittsburgh School of Medicine
Pittsburgh, Pennsylvania
Chapter 13, Physiologic Monitoring of the Surgical Patient

Fadi S. Dahdaleh, MD
Department of Surgery
University of Chicago
Chicago, Illinois
Chapter 30, The Appendix

Ralph J. Damiano, MD
Evarts A. Graham Professor of Surgery
Chief, Division of Cardiothoracic Surgery
Co-Chair, Heart & Vascular Center
Washington University School of Medicine
St Louis, Missouri
Chapter 21, Acquired Heart Disease

Senthil Damodaran, MD, PhD
Assistant Professor
Departments of Breast Medical Oncology and Investigational Cancer Therapeutics
The University of Texas MD Anderson Cancer Center
Houston, Texas
Chapter 17, The Breast
Contributors

Chandan Das, MD
Clinical Instructor
Department of Surgery
University of Toledo College of Medicine and Life Sciences
Toledo, Ohio
Chapter 37, Inguinal Hernias

Haile T. Debas, MD, FACS
Maurice Galante Distinguished Professor of Surgery, Emeritus
Founding Executive Director, Global Health Sciences
Director Emeritus, University of California Global Health Institute
Dean Emeritus, School of Medicine
Former Chancellor
University of California, San Francisco
San Francisco, California
Chapter 49, Global Surgery

Daniel T. Dempsey, MD, FACS
Professor of Surgery
Department of Surgery
Perelman School of Medicine
University of Pennsylvania
Philadelphia, Pennsylvania
Chapter 26, Stomach

Catherine R. deVries, MD
Professor of Surgery
Adjunct Associate Professor of Family Medicine/Public Health
University of Utah
Salt Lake City, Utah
Chapter 49, Global Surgery

Vikas Dudeja, MD
Assistant Professor of Surgery
Division of Surgical Oncology
Department of Surgery
University of Miami
Miami, Florida
Chapter 33, Pancreas

David L. Dunn, MD, PhD
Prospect, Kentucky
Chapter 6, Surgical Infections
Chapter 11, Transplantation

Geoffrey P. Dunn, MD
Medical Director
Department of Surgery
Hamot Medical Center
Erie, Pennsylvania
Chapter 48, Ethics, Palliative Care, and Care at the End of Life

Kelli Bullard Dunn, MD, FACS, FASCRS
Vice Dean, Community Engagement and Diversity
Professor of Surgery
University of Louisville School of Medicine
Louisville, Kentucky
Chapter 29, Colon, Rectum, and Anus

Nabil A. Ebraheim, MD
Chairman and Professor
Department of Orthopaedic Surgery
The University of Toledo Medical Center
Toledo, Ohio
Chapter 43, Orthopedic Surgery

Antoine Eskander, MD, ScM, FRCSC
Assistant Professor, University of Toronto
Department of Otolaryngology—Head & Neck Surgery
Sunnybrook Health Sciences Centre, Surgical Oncologist
Michael Garron Hospital, Endocrine Surgery
Institute for Clinical Evaluative Sciences (ICES), Adjunct Scientist
Toronto, Ontario, Canada
Chapter 18, Disorders of the Head and Neck

Stephen A. Esper, MD, MBA
Assistant Professor
Department of Anesthesiology and Perioperative Medicine
Director, UPMC Centers for Perioperative Care
University of Pittsburgh Medical Center
Pittsburgh, Pennsylvania
Chapter 50, Optimizing Perioperative Care: Enhanced Recovery and Chinese Medicine

Xin-Hua Feng, PhD
Distinguished Investigator
Life Sciences Institute
Zhejiang University
Hangzhou, Zhejiang, China
Professor of Molecular Cell Biology
Michael E. DeBakey Department of Surgery and Department of Molecular & Cellular Biology
Baylor College of Medicine
Houston, Texas
Chapter 15, Molecular Biology, The Atomic Theory of Disease, and Precision Surgery

Emily Finlayson, MD, MS, FACS
Professor in Residence
Department of Surgery, Division of General Surgery
Department of Medicine, Division of Geriatrics
Phillip R. Lee Institute for Health Policy Studies
Director, Center for Surgery in Older Adults
Director, Department of Surgery Faculty Mentoring Program
Department of Surgery
University of California San Francisco
San Francisco, California
Chapter 47, Surgical Considerations in Older Adults
William E. Fisher, MD, FACS
Professor of Surgery, Clinical Vice Chair and
Chief, Division of General Surgery
George L. Jordan, MD Chair of General Surgery
Michael E. DeBakey Department of Surgery
Director, Elkins Pancreas Center
Baylor College of Medicine
Houston, Texas
*Chapter 33, Pancreas*

Henri R. Ford, MD, MHA
Dean and Chief Academic Officer
University of Miami Miller School of Medicine
Don Soffer Clinical Research Center
Miami, Florida
*Chapter 39 Pediatric Surgery*

Freddie H. Fu, MD, DSc (Hon.), DPs (Hon.)
Distinguished Service Professor
University of Pittsburgh
David Silver Professor and Chairman
Department of Orthopaedic Surgery
University of Pittsburgh School of Medicine
Head Team Physician
University of Pittsburgh Department of Athletics
Pittsburgh, Pennsylvania
*Chapter 43, Orthopedic Surgery*

David A. Geller, MD
Richard L. Simmons Professor of Surgery
Chief, Division of Hepatobiliary and Pancreatic Surgery
University of Pittsburgh School of Medicine
Pittsburgh, Pennsylvania
*Chapter 31, Liver*

Nicole S. Gibran, MD, FACS
Professor, Department of Surgery
Associate Dean, Research and Graduate Education
Harborview Medical Center
University of Washington School of Medicine
Seattle, Washington
*Chapter 8, Burns*

Ricardo J. Gonzalez, MD, FACS
Chair, Sarcoma Department
Professor of Surgery
Chief of Surgery, Moffitt Cancer Center
Tampa, Florida
*Chapter 36, Soft Tissue Sarcoma*

John A. Goss, MD
Professor of Surgery
Michael E. DeBakey Department of Surgery
Division of Abdominal Transplantation and
Hepatobiliary Surgery
Baylor College of Medicine
Houston, Texas
*Chapter 31, Liver*

M. Sean Grady, MD
Charles Harrison Frazier Professor
Chairman
Department of Neurosurgery
Perelman School of Medicine at University of
Pennsylvania
Philadelphia, Pennsylvania
*Chapter 42, Neurosurgery*

Angelika Gruessner, PhD
Professor of Medicine
State University of New York (SUNY-Downstate)
New York, New York
*Chapter 11, Transplantation*

Rainer Gruessner, MD
Clarence & Mary Dennis Professor of Surgery
Chairman, Department of Surgery
State University of New York (SUNY-Downstate)
New York, New York
*Chapter 11, Transplantation*

David J. Hackam, MD, PhD, FACS
Garrett Professor and Chief of Pediatric Surgery
Professor of Surgery, Pediatrics and Cell Biology
Johns Hopkins University School of Medicine
Pediatric Surgeon-in-Chief and Co-Director
Johns Hopkins Children’s Center
The Charlotte R. Bloomberg Children’s Center
Baltimore, Maryland
*Chapter 39, Pediatric Surgery*

Kelly R. Haisley, MD
Clinical Instructor
Department of Surgery
Oregon Health and Science University
Portland, Oregon
*Chapter 32, Gallbladder and the Extrahepatic Biliary System*

Lee Hammerling, MD
Chief Academic Officer
President, New Ventures
ProMedica Health System
Toledo, Ohio
*Chapter 1, Leadership in Surgery*

Patrick Harbour, MD
Department of Plastic Surgery
MedStar Georgetown University Hospital
Washington, DC
*Chapter 16, The Skin and Subcutaneous Tissue*

Michael S. Harris, MD
Assistant Professor, Department of Otolaryngology
Assistant Professor, Medical College of Wisconsin
Department of Otolaryngology & Communication Sciences
Division of Neurotology & Skull Base Surgery
Milwaukee, Wisconsin
*Chapter 18, Disorders of the Head and Neck*
Contributors

David G. Heidt, MD, FACS
Clinical Associate Professor
University of Toledo College of Medicine
Staff Surgeon, St. Joseph Mercy Medical Center
Ann Arbor, Michigan
Chapter 30, The Appendix

Daniel B. Hinshaw, MD
Professor
Department of Surgery
University of Michigan
Ann Arbor, Michigan
Chapter 48, Ethics, Palliative Care, and Care at the End of Life

John B. Holcomb, MD
Professor of Surgery
Department of Surgery
University of Texas Houston
Houston, Texas
Chapter 4, Hemostasis, Surgical Bleeding, and Transfusion

Jennifer Holder-Murray, MD, FACS, FASCRS
Vice Chair of Quality Integration
Assistant Professor Surgery
Division of Colon and Rectal Surgery
University of Pittsburgh Medical Center
Pittsburgh, Pennsylvania
Chapter 50, Optimizing Perioperative Care: Enhanced Recovery and Chinese Medicine

Kelly K. Hunt, MD, FACS
Hamill Foundation
Distinguished Professor of Surgery in Honor of Dr. Richard G. Martin Sr.
Chair, Department of Breast Surgical Oncology
The University of Texas MD Anderson Cancer Center
Houston, Texas
Chapter 17, The Breast

John G. Hunter, MD, FACS, FRCS Edin(Hon.)
Executive Vice President and CEO, OHSU Health System
Mackenzie Professor of Surgery, School of Medicine
Oregon Health and Science University
Portland, Oregon
Chapter 14, Minimally Invasive Surgery, Robotics, Natural Orifice Transluminal Endoscopic Surgery, and Single-Incision Laparoscopic Surgery
Chapter 25, Esophagus and Diaphragmatic Hernia
Chapter 32, Gallbladder and the Extrahepatic Biliary System

Sayeed Ikramuddin, MD, MHA
Jay Phillips Professor and Chair
Department of Surgery
University of Minnesota Medical School
Minneapolis, Minnesota
Chapter 35, Abdominal Wall, Omentum, Mesentery, and Retroperitoneum

Tahir Jamil, MD
Clinical Assistant Professor of Surgery
University of Toledo College of Medicine and Life Sciences
Department of Surgery
Promedica Health System
Toledo, Ohio
Chapter 37, Inguinal Hernias

Rama Jayanthi, MD
Chief, Pediatric Urology
Nationwide Children’s Hospital
Dayton Children’s Hospital
Clinical Professor, Department of Urology
The Ohio State University
Columbus, Ohio
Chapter 40, Urology

Blair A. Jobe, MD, FACS
Chair of Surgery, Western Pennsylvania Hospital
Director, Institute for the Treatment of Esophageal and Thoracic Disease, Allegheny Health Network
Pittsburgh, Pennsylvania
Chapter 14, Minimally Invasive Surgery, Robotics, Natural Orifice Transluminal Endoscopic Surgery, and Single-Incision Laparoscopic Surgery
Chapter 25, Esophagus and Diaphragmatic Hernia

Puja Kachroo, MD
Clinical Instructor
Division of Cardiothoracic Surgery, Department of Surgery
Washington University School of Medicine
Barnes-Jewish Hospital
St. Louis, Missouri
Chapter 21, Acquired Heart Disease

Stephen Y. Kang, MD
Assistant Professor
Department of Otolaryngology—Head and Neck Surgery
Division of Head and Neck Oncology
The Ohio State University—James Comprehensive Cancer Center
Columbus, Ohio
Chapter 18, Disorders of the Head and Neck

Lillian S. Kao, MD, MS
Jack H. Mayfield, MD, Chair in Surgery
Vice-Chair of Research and Faculty Development
Vice-Chair for Quality of Care
Department of Surgery
McGovern Medical School at the University of Texas Health Science Center at Houston
Houston, Texas
Chapter 54, Web-Based Education and Implications of Social Media
Tara Karamlou, MD, MSc
Professor of Surgery, Division of Pediatric Cardiac Surgery
Cleveland Clinic
Cleveland, Ohio
Chapter 20, Congenital Heart Disease

Brian C. Kellogg, MD
Plastic and Reconstructive Surgery
Nationwide Children’s Hospital
700 Children’s Dr
Columbus, Ohio

Ali J. Khiabani, MD, MHA
Resident Physician
Department of Surgery
Washington University School of Medicine
St. Louis, Missouri
Chapter 21, Acquired Heart Disease

Stephanie Kielb, MD
Associate Professor
Departments of Urology, Medical Education, and Obstetrics and Gynecology
Northwestern University Feinberg School of Medicine
Chicago, Illinois
Chapter 40, Urology

Richard E. Kirschner, MD
Chief, Plastic and Reconstructive Surgery
Nationwide Children’s Hospital
Columbus, Ohio
Chapter 45, Plastic and Reconstructive Surgery

Scott Kizy, MD
Resident in Surgery
University of Minnesota
Minneapolis, Minnesota
Chapter 35, Abdominal Wall, Omentum, Mesentery, and Retroperitoneum

Bodo E. Knudsen, MD, FRCSC
Henry A. Wise II Endowed Chair in Urology
Associate Professor
Department of Urology
The Ohio State University
Columbus, Ohio
Chapter 40, Urology

Elise C. Kohn, MD
Head, Gynecologic Cancer and Neuroendocrine Cancer Therapeutics
Clinical Investigations Branch
Cancer Therapy Evaluation Program
National Cancer Institute
Bethesda, Maryland
Chapter 41, Gynecology

Rosemary A. Kozar, MD, PhD
Professor of Surgery
Director of Translational Research at Shock Trauma
Associate Director of Shock Trauma Anesthesia Research (STAR) Center
University of Maryland School of Medicine
Baltimore, Maryland
Chapter 4, Hemostasis, Surgical Bleeding, and Transfusion

Mary R. Kwaan, MD, MPH
Associate Professor of Surgery
University of California, Los Angeles
Los Angeles, California
Chapter 29, Colon, Rectum, and Anus

Geeta Lal, MD, MSc, FRCS(C), FACS
Associate Professor, Surgery
Associate Chief Quality Officer, Inpatient
University of Iowa
Iowa City, Iowa
Chapter 38, Thyroid, Parathyroid, and Adrenal

Jeremy J. Laukka
Associate Professor of Medical Education & Neurology
Department of Medical Education
Associate Dean for Foundational Science Curriculum
University of Toledo College of Medicine and Life Sciences
Toledo, Ohio
Chapter 1, Leadership in Surgery

Cheryl T. Lee, MD
Dorothy M. Davis Endowed Chair in Cancer Research
Professor and Chairman
Department of Urology
The Ohio State University
Columbus, Ohio
Chapter 40, Urology

Evan Leibner, MD, PhD
Assistant Professor
Institute of Critical Care Medicine & Department of Emergency Medicine
Mount Sinai Medical Center
Icahn School of Medicine at Mount Sinai
New York, New York
Chapter 4, Hemostasis, Surgical Bleeding, and Transfusion
Contributors

Scott A. LeMaire, MD
Jimmy and Roberta Howell Professor of Cardiovascular Surgery
Vice Chair for Research, Michael E. DeBakey Department of Surgery
Professor of Molecular Physiology and Biophysics
Director of Research, Division of Cardiothoracic Surgery
Baylor College of Medicine
Department of Cardiovascular Surgery, Texas Heart Institute
Baylor St. Luke’s Medical Center
Houston, Texas
Chapter 22, Thoracic Aneurysms and Aortic Dissection

Xinran Li, PhD
Postdoctoral Fellow
Life Sciences Institute
Zhejiang University
Hangzhou, Zhejiang, China
Chapter 15, Molecular Biology, The Atomic Theory of Disease, and Precision Surgery

Timothy K. Liem, MD, FACS
Professor of Surgery, Vice-Chair for Quality
Department of Surgery, Knight Cardiovascular Institute
Oregon Health & Science University
Portland, Oregon
Chapter 24, Venous and Lymphatic Disease

Scott D. Lifchez, MD, FACS
Associate Professor of Plastic Surgery and Orthopedic Surgery
Program Director, Johns Hopkins/University of Maryland Plastic Surgery Residency
Director of Hand Surgery
Johns Hopkins Bayview Medical Center
Baltimore, Maryland
Chapter 44, Surgery of the Hand and Wrist

Amy Lightner, MD
Assistant Professor of Surgery
Mayo Clinic College of Medicine
Medical Director of Translation Into Practice Platforms Center for Regenerative Medicine, Mayo Clinic
Senior Associate Consultant, Department of Surgery
Mayo Clinic College of Medicine
Division of Colon and Rectal Surgery
Rochester, Minnesota
Chapter 1, Leadership in Surgery

Peter H. Lin, MD
Professor Emeritus of Surgery
Michael E. DeBakey Department of Surgery
Baylor College of Medicine
Houston, Texas
Chapter 23, Arterial Disease

Xia Lin, PhD
Associate Professor of Surgery
Michael E. DeBakey Department of Surgery
Baylor College of Medicine
Houston, Texas
Chapter 15, Molecular Biology, The Atomic Theory of Disease, and Precision Surgery

James D. Luketich, MD
Henry T. Bahnson Professor of Cardiothoracic Surgery, Chief, The Heart, Lung, and Esophageal Surgery Institute, Department of Surgery
Division of Thoracic and Foregut Surgery
University of Pittsburgh
Pittsburgh, Pennsylvania
Chapter 19, Chest Wall, Lung, Mediastinum, and Pleura

Carlos Rodriguez-Otero Luppi, MD, PhD
Unit of Gastrointestinal and Hematological Surgery, Hospital Sant Pau
Professor of Surgery, Autonomous University of Barcelona
Barcelona, Spain
Chapter 34, The Spleen

James R. Macho, MD, FACS
Emeritus Professor of Surgery
University of California, San Francisco
San Francisco, California
Chapter 1, Leadership in Surgery
Chapter 37, Inguinal Hernias
Chapter 52, Ambulatory Surgery

Martin A. Makary, MD, MPH, FACS
Professor of Surgery
Johns Hopkins University
Baltimore, Maryland
Chapter 12, Quality, Patient Safety, Assessments of Care, and Complications

Samuel P. Mandell, MD, MPH, FACS
Assistant Professor
Department of Surgery
University of Washington School of Medicine
UW Medicine Regional Burn Center
Harborview Medical Center
Seattle, Washington
Chapter 8, Burns

Stephen Markowiak, MD
Clinical Instructor
Department of Surgery
University of Toledo College of Medicine and Life Sciences
Toledo, Ohio
Chapter 1, Leadership in Surgery
Chapter 52, Ambulatory Surgery
Jeffrey B. Matthews, MD, FACS
Dallas B. Phemister Professor
Chair, Department of Surgery
Surgeon-in-Chief
The University of Chicago Pritzker School of Medicine
Chicago, Illinois
Chapter 51, Understanding, Evaluating, and Using Evidence for Surgical Practice

Christopher McClung, MD
Adjunct Professor
The Ohio State University
Reconstructive Urologist
The Central Ohio Urology Group
Columbus, Ohio
Chapter 40, Urology

Funda Meric-Bernstam, MD
Chair, Department of Investigational Cancer Therapeutics; Medical Director, Institute for Personalized Cancer Therapy; Professor, Divisions of Cancer Medicine and Surgery
MD Anderson Cancer Center
Houston, Texas
Chapter 10, Oncology

Hollis Merrick, MD, FACS
Professor Emeritus
Department of Surgery
University of Toledo College of Medicine and Life Sciences
Toledo, Ohio
Chapter 1, Leadership in Surgery

Michael J. Miller, MD, FACS
Professor and Chair
Department of Plastic and Reconstructive Surgery
The Ohio State University
Wexner Medical Center
Columbus, Ohio
Chapter 45, Plastic and Reconstructive Surgery

Tabitha G. Moe, MD, FACC
Assistant Professor of Surgery
Adult Congenital Cardiology
Pulmonary Hypertension
Pregnancy and Cardiovascular Disease
Creighton University School of Medicine
Omaha, Nebraska
Chapter 20, Congenital Heart Disease

Gregory L. Moneta, MD, FACS
Professor, Division of Vascular Surgery
Department of Surgery and Knight Cardiovascular Institute
Oregon Health & Science University
Portland, Oregon
Chapter 24, Venous and Lymphatic Disease

Ernest E. Moore, MD
Ernest E. Moore Shock Trauma Center at Denver Health
Distinguished Professor of Surgery, University of Colorado Denver
Editor, Journal of Trauma
Denver, Colorado
Chapter 7, Trauma

Bart Muller, MD, PhD
Orthopaedic Surgeon
Academic Medical Center
Amsterdam, The Netherlands
Chapter 43, Orthopedic Surgery

Raghav Murthy, MD, DABS, FACS
Assistant Professor
Director, Pediatric Heart Transplantation
Division of Pediatric Cardiac Surgery
Kravis Children’s at Mount Sinai Hospital
Icahn School of Medicine,
New York, New York
Chapter 20, Congenital Heart Disease

Moriah Muscaro, MD
Clinical Instructor
University of Toledo College of Medicine and Life Sciences
Toledo, Ohio
Chapter 52, Ambulatory Surgery

Katie S. Nason, MD, MPH
Vice-Chair for Surgical Research, Department of Surgery
Fellow, Institute for Healthcare Delivery and Population Science
Thoracic Surgeon, Division of Thoracic Surgery
University of Massachusetts Medical School—Baystate Health
Springfield, Massachusetts
Chapter 19, Chest Wall, Lung, Mediastinum, and Pleura

Munier Nazzal, MD, MBA, CPE, FACS, FRCS
Professor and Vice Chair, Department of Surgery
Chief, Division of Education
Chief, Division of Vascular, Endovascular and Wound Surgery.
Medical Director of the Wound Center, University of Toledo Medical Center
University of Toledo
Toledo, Ohio
Chapter 1, Leadership in Surgery
Chapter 9, Wound Healing
Chapter 52, Ambulatory Surgery
Matthew D. Neal, MD, FACS
Roberta G. Simmons Assistant Professor of Surgery
Assistant Professor of Clinical and Translational Science
and Critical Care Medicine
Departments of Surgery
Critical Care Medicine, and the Clinical and
Translational Science Institute (CTSI)
University of Pittsburgh
Attending Surgeon, Division of Trauma and Acute Care Surgery
University of Pittsburgh Medical Center
Pittsburgh, Pennsylvania
Chapter 3, Fluid and Electrolyte Management of the Surgical Patient

John Nemunaitis, MD
Professor of Medicine
Chief, Division of Hematology and Oncology
University of Toledo College of Medicine and Life Sciences
Director Cancer Research
ProMedica Health System
Toledo, Ohio
Chapter 15, Molecular Biology, The Atomic Theory of Disease, and Precision Surgery

Bradley Otto, MD
Assistant Professor
Department of Otolaryngology—Head and Neck Surgery
Division of Skull Base Surgery
Director, Division of General Otolaryngology
Director, Division of Rhinology
The Ohio State University—James Comprehensive Cancer Center
Columbus, Ohio
Chapter 18, Disorders of the Head and Neck

Mohamed F. Osman, MD
Assistant Professor of Surgery
Division of Vascular and Endovascular Surgery
Department of Surgery
Director of Surgical Intensive Care
The University of Toledo College of Medicine and Life Sciences
Toledo, Ohio
Chapter 9, Wound Healing

John J. Nigro, MD
Chief, Pediatric Cardiac Surgery
Director, Children’s Heart Center
Rady Children’s Hospital
San Diego, California
Chapter 20, Congenital Heart Disease

Catherine C. Parker, MD, FACS
Assistant Professor
Department of Surgery
The University of Alabama at Birmingham
Birmingham, Alabama
Chapter 17, The Breast

Timothy M. Pawlik, MD, MPH, MTS, PhD, FACS, RACS (Hon.)
Professor and Chair
Department of Surgery
The Urban Meyer III and Shelley Meyer Chair for Cancer Research
Surgeon-in-Chief
The Ohio State University Wexner Medical Center
Columbus, Ohio
Chapter 48, Ethics, Palliative Care, and Care at the End of Life

Michael F. O’Connor, MD, FCCM
Professor
Department of Anesthesia & Critical Care
The University of Chicago
Chicago, Illinois
Chapter 46, Anesthesia for Surgical Patients

Andrew B. Peitzman, MD
Mark M. Ravitch Professor and Vice Chairman
Department of Surgery, University of Pittsburgh School of Medicine
Pittsburgh, Pennsylvania
Chapter 5, Shock
Raphael E. Pollock, MD, PhD, FACS
Director, The Ohio State University Comprehensive Cancer Center
Professor of Surgery; Kathleen Klotz Chair in Cancer Research
The Ohio State University Wexner Medical Center
Columbus, Ohio
Chapter 10, Oncology
Chapter 36, Soft Tissue Sarcomas

Matthew Pommerening, MD, MS
Department of Cardiothoracic Surgery
University of Pittsburgh Medical Center (UPMC)
Pittsburgh, Pennsylvania
Chapter 4, Hemostasis, Surgical Bleeding, and Transfusion

Ourania Preventza, MD
Associate Professor, Michael E. DeBakey Department of Surgery
Division of Cardiothoracic Surgery
Baylor College of Medicine
Department of Cardiovascular Surgery, Texas Heart Institute
Baylor St. Luke’s Medical Center, Houston, Texas
Chapter 22, Thoracic Aneurysms and Aortic Dissection

Jonathan Pribaz, MD
Resident in Orthopaedic Surgery
UCLA Department of Orthopaedic Surgery
Santa Monica, California
Chapter 43, Orthopedic Surgery

Raymond R. Price, MD, FACS
Director, Graduate Surgical Education, Intermountain Medical Center, Intermountain Healthcare
Director, Center for Global Surgery, Professor
Department of Surgery, University of Utah
Salt Lake City, Utah
Chapter 49, Global Surgery

Carla M. Pugh, MD, PhD
Professor of Surgery
Director, Technology Enabled Clinical Improvement Center
Stanford University
Palo Alto, California
Chapter 53, Skills and Simulation

Ashwin G. Ramayya, MD, PhD
Department of Neurosurgery
Perelman School of Medicine at University of Pennsylvania
Philadelphia, Pennsylvania
Chapter 42, Neurosurgery

Kara Richardson, MD
Clinical Instructor
Department of Surgery
University of Toledo College of Medicine and Life Sciences
Toledo, Ohio
Chapter 52, Ambulatory Surgery

Robert E. Roses, MD, FACS
Assistant Professor of Surgery
Perelman School of Medicine
University of Pennsylvania
Philadelphia, Pennsylvania
Chapter 26, Stomach

Philip R. Schauer, MD
Professor of Surgery, Cleveland Clinic Lerner College of Medicine
Director, Bariatric and Metabolic Institute
Cleveland Clinic
Cleveland, Ohio
Chapter 27, The Surgical Management of Obesity

Matthew R. Schill, MD
Clinical Instructor
Division of Cardiothoracic Surgery Department of Surgery
Washington University School of Medicine
Barnes-Jewish Hospital
St Louis, Missouri
Chapter 21, Acquired Heart Disease

Andrew B. Schneider, MD, MS
The University of Chicago Medicine
Chicago, Illinois
Chapter 51, Understanding, Evaluating, and Using Evidence for Surgical Practice

Steven A. Schulz, MD
Assistant Professor, Department of Plastic and Reconstructive Surgery
The Ohio State University Wexner Medical Center
Columbus, Ohio
Chapter 45, Plastic and Reconstructive Surgery

Neal E. Seymour, MD, FACS
Vice Chair for Clinical Affairs, Department of Surgery
Division Chief, General Surgery, Baystate Health
Professor and Surgery Residency Program Director
University of Massachusetts Medical School—Baystate
Director, Baystate Simulation Center—Goldberg Surgical Skills Lab
Springfield, Massachusetts
Chapter 53, Skills and Simulation
Contributors

Joseph Sferra, MD, FACS
Clinical Associate Professor of Surgery
University of Toledo College of Medicine and Life Sciences
Vice President Department of Surgery
Promedica Health System
Toledo, Ohio
Chapter 37, Inguinal Hernias

Ahmad Shabsigh, MD, FACS
Assistant Professor
Department of Urology
The Ohio State University
Columbus, Ohio
Chapter 40, Urology

Mark L. Shapiro, MD, FACS
Trauma Medical Director and Chief of Trauma
Grady Memorial Hospital
Emory University Department of Surgery
Atlanta, Georgia
Chapter 12, Quality, Patient Safety, Assessments of Care, and Complications

Saurabh Sinha, MD
Department of Neurosurgery
Perelman School of Medicine at University of Pennsylvania
Philadelphia, Pennsylvania
Chapter 42, Neurosurgery

Katherine E. Smiley, MD
University of Utah Health
Division of General Surgery
Salt Lake City, Utah
Chapter 49, Global Surgery

David H. Song, MD, MBA, FACS
Physician Executive Director
MedStar Plastic & Reconstructive Surgery
Professor and Chairman
Department of Plastic Surgery
Georgetown University School of Medicine
Washington, DC
Chapter 16, The Skin and Subcutaneous Tissue

Michael W. Sourial, MD, FRCSC
Fellow, Endourology and Minimally Invasive Surgery
Department of Urology
The Ohio State University
Columbus, Ohio
Chapter 40, Urology

Stephen Stanek, MD
Clinical Assistant Professor of Surgery
University of Toledo College of Medicine and Life Sciences
Department of Surgery
Promedica Health System
Toledo, Ohio
Chapter 37, Inguinal Hernias

David B. Stewart, MD, FACS, FASCRS
Section Chief, Colorectal Surgery
University of Arizona – Banner University Medical Center
Tucson, Arizona
Chapter 29, Colon, Rectum, and Anus

Anne M. Suskind, MD, FACS
Associate Professor of Urology
Department of Urology
University of California San Francisco
San Francisco, California
Chapter 47, Surgical Considerations in Older Adults

Eduardo M. Targarona, MD, PhD, FACS
Chief of the Unit of Gastrointestinal and Hematological Surgery, Hospital Sant Pau
Professor of Surgery,
Autonomous University of Barcelona
Barcelona, Spain
Chapter 34, The Spleen

Ali Tavakkoli, MD
Division of General and GI Surgery
Brigham and Women’s Hospital
Associate Professor of Surgery
Harvard Medical School
Boston, Massachusetts
Chapter 28, Small Intestine

Theodoros N. Teknos, MD, FACS
Professor of Otolaryngology-Head and Neck Surgery
President and Scientific Officer of UH Seidman Cancer Center
UH Cleveland Medical Center
Cleveland, Ohio
Chapter 18, Disorders of the Head and Neck

Sarah M. Temkin, MD
Division of Gynecologic Oncology
Virginia Commonwealth University
Richmond, Virginia
Chapter 41, Gynecology

Bert J. Thomas, MD
Professor and Chief, Joint Replacement Service
Department of Orthopedic Surgery
David Geffen School of Medicine at UCLA
Los Angeles, California
Chapter 43, Orthopedic Surgery
Allan Tsung, MD
UPMC Professor of Surgery
Division of Hepatobiliary and Pancreatic Surgery
University of Pittsburgh
Pittsburgh, Pennsylvania
Chapter 31, Liver

Kiran K. Turaga, MD, MPH
Vice Chief, Section of General Surgery and Surgical Oncology
Associate Professor, Department of Surgery
Director, Surgical GI Cancer Program
Director, Regional Therapeutics Program
The University of Chicago Medical Center
Chicago, Illinois
Chapter 30, The Appendix

Jeffrey Upperman, MD
Associate Professor of Surgery
Director of Trauma, Pediatric Surgery
Children’s Hospital Los Angeles
Keck School of Medicine
University of Southern California
Los Angeles, California
Chapter 39, Pediatric Surgery

Ian L. Valerio, MD
Professor, Department of Plastic and Reconstructive Surgery
The Ohio State University Wexner Medical Center
Columbus Ohio
Chapter 45, Plastic and Reconstructive Surgery

Frank J. Veith, MD
Professor of Surgery
Department of Surgery
New York University Langone Health
New York, New York
Chapter 23, Arterial Disease

Dharmesh Vyas, MD, PhD
University of Pittsburgh Medical Center Orthopaedic Surgery
UPMC Lemieux Sports Complex
Medical Director and Head Team Physician
Pittsburgh Penguins
Pittsburgh, Pennsylvania
Chapter 43, Orthopedic Surgery

Kasper Wang, MD
Associate Professor of Surgery
Keck School of Medicine
University of Southern California
Los Angeles, California
Chapter 39, Pediatric Surgery

Ximo Wang, MD, PhD
Professor of Surgery
Department of Surgery
Tianjin Medical University, Nankai Hospital
Chief, Institute of Acute Abdomen of Tianjin
Tianjin, China
Chapter 50, Optimizing Perioperative Care: Enhanced Recovery and Chinese Medicine

Zhiliang Wang, MD
Professor of Surgery
Institute of Hepatobiliary Diseases
Zhongnan Hospital of Wuhan University
Hubei Providence, China
Institute of Acute Abdomen of Tianjin
Tianjin, China
Chapter 50, Optimizing Perioperative Care: Enhanced Recovery and Chinese Medicine

David I. Watson, MBBS, MD, PhD, FRACS, FRCSEd (Hon.), FAHMS
Professor & Head, Discipline of Surgery
Flinders University of South Australia
Adelaide, South Australia
Australia
Chapter 25, Esophagus and Diaphragmatic Hernia

Randal S. Weber, MD, FACS
Professor and Chairman, Director of Surgical Services
Department of Head and Neck Surgery
University of Texas MD Anderson Cancer Center
Houston, Texas
Chapter 18, Disorders of the Head and Neck

Adam S. Weltz, MD
Department of Surgery
Anne Arundel Medical Center
Annapolis, Maryland
Chapter 34, The Spleen

John A. Windsor, BSc, MD, FRACS, FACS, FRSNZ
HBP/Upper GI Surgeon and Professor of Surgery
Director, Surgical and Translational Research Centre
Department of Surgery, University of Auckland,
Auckland, New Zealand
Chapter 33, Pancreas

Juehua Yu, PhD
Associate Professor
Cancer Research Institute, Hangzhou Cancer Hospital,
Hangzhou
Department of Child Health Care, Xinhua Hospital
Shanghai JiaoTong University School of Medicine
Shanghai, China
Chapter 15, Molecular Biology, The Atomic Theory of Disease, and Precision Surgery
Contributors

Michael E. Zenilman, MD, MS
Professor of Surgery
Weill Cornell Medicine
Chair, Department of Surgery
New York Presbyterian Brooklyn Methodist Hospital
Brooklyn, New York
Chapter 54, Web-Based Education and Implications of Social Media

Michael J. Zinner, MD, FACS
Moseley Professor of Surgery, Emeritus
Harvard Medical School
Boston, Massachusetts
CEO and Executive Medical Director
Miami Cancer Institute
Miami, Florida
Chapter 28, Small Intestine

Brian S. Zuckerbraun, MD, FACS
Henry T. Bahnson Professor of Surgery
University of Pittsburgh
Chief, Trauma and General Surgery
University of Pittsburgh Medical Center and VA
Pittsburgh Healthcare System
Pittsburgh, Pennsylvania
Chapter 5, Shock
It was a singular privilege to serve as editor-in-chief of the first and subsequent six editions of *Principles of Surgery*. The invitation from the current editor-in-chief, Dr. F. Charles Brunicardi, who has discharged that responsibility for the ensuing four editions, to participate in the textbook’s 50th anniversary, is gratifying. The readers of the first seven editions often commented on the distinctive yellow cover. On this particular celebration of longevity, the color yellow connotes “gold.”

The past 50 years has witnessed an unimaginable growth in scientific knowledge available to students of surgery. The “science of surgery” has gained dominance over the “art of surgery.” Diverse technologies have been incorporated to expedite diagnosis and improve surgical excision or repair. The establishment of more precise criteria for categorization and analyzing data, coupled with advances in informatics, has allowed for the practice of “evidence-based medicine and surgery.” It is, as if, today’s surgeons have adopted a new language, new rules, new protocols—and anticipate new outcomes. The passage of time has been associated with transformative change, which has been beautifully captured in the 11th edition.

Among the “Basic Considerations” that transcend individual organ systems, change has occurred at an ever-accelerating pace, in multiple arenas, with variable consequences, since the first edition made its debut. Not all changes have been favorable. Increased effectiveness of antibiotics has improved the outcomes of the treatment of sepsis, but has been associated with the appearance of *clostridium difficile* colitis and lethal MRSA hospital outbreaks. HIV, AIDS, HPV (human papilloma virus), and hepatitis B and C had not entered the surgical lexicon prior to publication of *Principles of Surgery*.

Over the course of years, trauma has become an ever-increasing problem. Since publication of the first edition, improved diagnostic techniques have altered the approach to individuals who sustained major trauma. The concept of immediate “damage control to be followed by delayed definitive treatment,” the availability of angioembolization to control bleeding, and inert material to maintain protect the unclosed abdominal abdomen for protected state for a critical, at times prolonged, period of time, during which caloric requirements are satisfied parenterally.

In oncology, a more precise tumor classification based on size, nodal involvement, metastases, chemical and biologic characteristics has been accepted. This, in turn, has allowed for more meaningful assessment of a variety of therapeutic regimens. Chemotherapy has been joined by immunotherapy, and targeted, precision genomic therapy has recently been introduced.

At the time of publication of the first edition of *Principles of Surgery*, only the kidney was deemed clinically acceptable for homotransplantation and satisfactory immunosuppression had not been developed. Advances in immunosuppression have added the liver, pancreas, small bowel, heart, and lungs to the list of organs transplanted with anticipated success.

Among the 1805 pages of text in the first edition, “facts” and “declarations by experts” have failed to stand the test of time for a variety of reasons. Little effort is required to uncover statements that now would be judged “False!” For example: (1) Cancer of the hypopharynx is three to four times as common as cancer of the larynx (the reverse is true). (2) Effective treatment of a single ventricle in a neonate is not feasible. (3) The distal 1 to 2 cm of the esophageal lumen is normally lined by columnar rather than squamous epithelium (the description of a Barrett’s esophagus). (4) There is but one treatment for acute appendicitis…the only question to be resolved is the timing of surgical intervention. (5) The adenomatous (colon) polyp is a lesion of negligible malignant potential. (6) The only acceptable treatment for a splenic injury accompanied by any evidence of intraperitoneal bleeding in an adult is splenectomy. (7) Hundred percent of patients with primary hyponatremia have hypokalemia (most have no hypokalemia). More dramatic is the evidence that many of the prevalent surgical procedures that merited detailed illustration, consuming multiple pages in the first edition, are now, rarely if ever, performed.
It must be emphasized that a textbook chronicles a science during the contemporaneous time. The first edition, as is true for each of the 11 editions of *Principles of Surgery*, is a compendium that pertains, solely, up to the time of publication. Print does not imply permanence. Print often outlasts the fact it promulgates. I congratulate Dr. Brunicardi and the coeditors on a modern and beautifully written 11th edition that carries forward the tradition of the *Principles of Surgery* into the next 50 years.

Seymour I. Schwartz, MD, FACS
Distinguished Alumni Professor of Surgery
University of Rochester School of Medicine & Dentistry
It is both an honor and a privilege to be asked to become an associate editor for the 11th edition of *Schwartz’s Principles of Surgery*. Much has changed since the first edition was published in 1969, particularly in terms of how adult learners obtain knowledge. Today, approximately nine out of ten American adults use the internet and internet use by college graduates is nearly universal. Journal articles on any and all topics are available with a few keystrokes, with over 1,000 new articles being added daily to archives such as PubMed Central. Additionally, there are a multitude of online textbooks, videos of procedures, interactive surgical simulator applications, and other web-based resources that are widely available to medical students and professionals. *So, one might ask, do we still need surgical textbooks?*

The debate about whether textbooks are obsolete is not a new one. Opponents of textbooks suggest that they are expensive and inconvenient to access. Their content can be argued to become quickly outdated and to be unengaging to the modern learner who prefers interactive, multimedia content. On the other hand, proponents of textbooks note that evidence is lacking that comprehension is improved with digital technology. Furthermore, textbooks allow teachers to provide content within a clear framework, to ensure uniform delivery of content, and to have ease in re-referencing information.

*What is the right answer?* Modern and future learners should have textbooks available to them in multiple media formats. One media type does not fit all learners. Like surgery, optimal learning must be personalized based on an individual’s preferences. The editors and publishing company behind *Schwartz’s Principles of Surgery* have embraced this idea—the hardcover continues to be the best-selling general surgery textbook worldwide and there are no plans to eliminate the printed version. At the same time, the content is widely available on an interactive digital platform—Access Surgery—that includes access to multiple textbooks, quick references, a video atlas, and test review questions.

Regardless of the format, knowledge must come from a reliable source of information. For example, each chapter in the 11th edition of *Schwartz’s Principles of Surgery* is written by at least one, and often two or more, authors who are experts in the subject matter. These authors have frequently built on work by those who have preceded them. Furthermore, each chapter is supported by the evidence and vetted by one or more senior surgeons serving as editors. This new edition continues to provide up-to-date information on age-old topics in surgery such as the physiologic basis of disease as well as on the clinical diagnosis and management of surgical diseases.

The 11th edition deftly balances core knowledge that has stood the test of time with contemporary advances in science and technology. Examples include updated chapters on “Molecular Biology, The Atomic Theory of Disease, and Precision Surgery” and “Minimally Invasive Surgery, Robotics, and Natural Orifice Transluminal Endoscopic Surgery.” Additionally, there are multiple chapters focused on non-technical skills, which are often more important than technical skills, such as the first chapter of the textbook on “Leadership in Surgery.” This 11th edition also boasts five new chapters: “Enhanced Recovery after Surgery,” “Understanding and Evaluating Evidence for Surgical Practice,” “Ambulatory/Outpatient Surgery,” “Skills and Simulation,” and “Web-Based Education and Implications of Social Media.”

The fact that the 11th edition of *Schwartz’s Principles of Surgery* marks the textbook’s 50th anniversary is a testament to its continued relevance and contributions to surgical education. Moreover, its longevity is also a reflection of far-sighted editors-in-chief, first Dr. Seymour Schwartz followed by Dr. F. Charles Brunicardi, who have been able to not only keep up with but also to anticipate changes in the surgical landscape. Not only is surgery a continuously changing discipline, but also the world in which surgeons practice is constantly evolving, as reflected by the digital era. Nonetheless, textbooks and the knowledge they carry will continue to play an important role, regardless of their format and packaging.

Lillian S. Kao, MD
Jack H. Mayfield, MD, Chair in Surgery
Professor and Chief, Division of Acute Care Surgery
Vice Chair of Research and Faculty Development
Vice Chair of Quality of Care
Co-Director, Center for Surgical Trials and Evidence-based Practice (C-STEP)
Department of Surgery
McGovern Medical School at the University of Texas Health Science Center at Houston
For more free medical ebooks go to extern.ir
With the publication of its 11th edition, Schwartz’s Principles of Surgery celebrates its 50th anniversary. It is remarkable to consider the number of students, residents, fellows, surgeons, and patients who have benefitted from the collective knowledge compiled in this text over the last half-century. It is an honor for the current editorial board to carry forward the tradition of excellence in education established by Dr. Seymour Schwartz and previous editors. We recognize that surgeons have entered into an era of surgery in which the outcomes of operations and patient satisfaction scores are carefully monitored, demanding excellence through enhanced evidenced-based knowledge, patient-family–centered care, and the highest levels of professionalism.

The first chapter on leadership has taken on special meaning in light of the new demands placed on surgeons for both technical and nontechnical skills, underscoring the importance of instituting a formal leadership-training program for surgery students of all ages with an emphasis on mentoring. We have also entered into the dawn of a new era of surgery with advances in minimally invasive surgery using robots, molecular contrast, and full computerization, thus enhancing the safety of surgery and allowing surgeons a more comfortable environment in which to work. We recognize that the use of “omic” information is ushering in an era of precision surgery and the importance of surgeons, who have access to the tissues of the human body on a daily basis for “omic” profiling that will guide targeted therapies to enhance the outcomes of surgery.

Taking these constructs into consideration, the editors and authors of this 11th edition have done their very best to revise each chapter and convey the current state-of-the-art in surgery. Continuing in this effort, five new chapters have been added: Understanding and Evaluating Evidence for Surgical Practice, Enhanced Recovery after Surgery, Ambulatory/Outpatient Surgery, Skills and Simulation, and Web-based Education and Implications of Social Media. This edition contains the latest in leadership training, surgical science, surgical techniques, and therapy for students, residents, fellows, and surgeons. Another important component of this new edition is the artwork. We acknowledge the outstanding artistic team of Jason M. McAlexander & Associates who directed the full color art program, which provides clear and consistent learning aids throughout the text and visually reflects the comprehensive and updated nature of this book.

The editors are thankful that this text is a trusted source for training and crafting surgeons worldwide. Such success is due in large part to the extraordinary efforts of our contributors—leaders in their fields—who not only train up-and-coming surgeons but also impart their knowledge and expertise to benefit patients across the globe. The inclusion of many international authors to the chapters within is ultimately a testament to mentorship, albeit on a broader scale, and we thank these authors and mentors, both near and far. To our fellow editorial board members who have tirelessly devoted their time and knowledge to the integrity and excellence of their craft and this textbook, we extend our gratitude. We are thankful to Andrew Moyer, Christie Naglieri, and all at McGraw-Hill who continued to believe in and support this work, and we wish to thank Katie Elsbury for her dedication to the organization and editing of this edition. Lastly, we would like to thank our families for their support and love.

F. Charles Brunicardi, MD, FACS
Dedication

We, the editors of this leading textbook of surgery, *Schwartz’s Principles of Surgery* are pleased to dedicate the 11th edition to Dr. Frank Gordon Moody. While most academic surgeons recognize Dr. Moody, as a top echelon surgical leader of the last half century, we choose to dedicate this edition to him because of the profound influence he had on the careers of many of the editors of this textbook. To some of us, Dr Moody was our surgical chair and academic inspiration. To others he was a research collaborator. For those of us who are not direct descendants, academically speaking, Frank Moody had the ability to recognize and provide the gift of mentorship to talented academic surgeons, irrespective of their academic pedigree.

Dr. Moody was born in Franklin, New Hampshire, attended Dartmouth College and Dartmouth Medical School (when it was a two-year school) then received his MD from Cornell. He stayed at Cornell throughout his surgical training, enticed into upper GI surgery by Dr. Frank Glenn. His academic career started at the University of California, San Francisco, under the legendary leadership of Dr. Bert Dunphy. He was subsequently recruited to the University of Alabama, Birmingham, where he rose to the rank of professor. In 1971, he became the Chair of Surgery at the University of Utah, coupling his love for skiing and hiking with an intense desire to bring scientific inquiry to the Wasatch Front. There, his passion for mentorship was uncovered. Eight of his trainees became department chairs, and many more visited Utah where the academic ‘bug’ was inoculated. In 1982, Dr. Moody took his talents to the University of Texas, Houston, where he served as the Denton Cooley Chair of Surgery. While he stepped down as Chair 12 years later, Dr. Moody remained in Texas for the rest of his career. Dr. Moody’s influence was truly global; he was active in the International Surgical Society and was a founder of the International Surgical Group. It was often said that there was never a meeting that Dr. Moody missed— and at every meeting he truly “showed up”, contributing to the program, asking challenging questions, and spurring new lines of investigation for the many GI surgeons lucky enough to have Dr. Moody engage with their line of discovery. Nearly continuously funded by the National Institutes of Health (NIH) from 1967 to 2008, Dr. Moody was a force for surgical science, encouraging active participation by surgeons in the NIH study sections.

To many of the editors, the connection to Dr. Moody was even more personal. Attracted to training in Utah by the combination of skiing, science, and great surgical training, I first met Dr. Moody in the pages of the 3rd edition of this textbook, in which he authored the chapter on gallbladder disease. After many years of learning in the operating room and the laboratory, it is an honor to follow in his footsteps as the author of this chapter in this and the prior three editions of this classic surgical book. Dr. Moody, we will miss you, and hope to carry your many gifts forward, the greatest of which were your support and mentorship of the many who have been lucky enough to follow in your footsteps.

John G. Hunter MD and the editors of *Schwartz’s Principles of Surgery, 11th edition*
This page intentionally left blank
INTRODUCTION

The field of surgery has evolved greatly from its roots, and surgical practice now requires the mastery of modern leadership principles and skills as much as the acquisition of medical knowledge and surgical technique. Historically, surgeons took sole responsibility for their patients and directed proceedings in the operating room with absolute authority, using a command-and-control style of leadership. Modern surgical practice has now evolved from single provider–based care toward a team-based approach, which requires collaborative leadership skills. Surgical care benefits from the collaboration of surgeons, anesthesiologists, internists, radiologists, pathologists, radiation oncologists, nurses, pharmacists, social workers, therapists, hospital staff, and administrators. Occupying a central role on the healthcare team, surgeons have the potential to improve patient outcomes, reduce medical errors, and improve patient satisfaction through their leadership of the multidisciplinary team. Thus, in the landscape of modern healthcare systems, it is imperative that surgical training programs include formal instruction on leadership principles and skills to cultivate their trainees’ leadership capabilities.

Many medical and surgical communities, including residency training programs, acknowledge the need for improved physician leadership. Specifically, surveyed surgical residents felt a lack of confidence in multiple areas of leadership, particularly in conflict resolution. Surgical trainees identify leadership skills as important, but they report themselves as “not competent” or “minimally competent” in this regard. While a small number of surgical training programs have implemented formal curricula focused on teaching leadership principles, it is now imperative that all surgical training programs teach these important skills to their trainees. Interviews of academic chairpersons identified several critical leadership success factors, including mastery of visioning, communication, change management, emotional intelligence, team building, business skills, personnel management, and systems thinking. These chairpersons stated that the ability of emotional intelligence was “fundamental to their success and its absence the cause of their failures,” regardless of medical knowledge. Thus, residency programs need to include leadership training to prepare future surgeons for success in modern healthcare delivery.

In the United States, the Accreditation Council for Graduate Medical Education (ACGME) has established six core competencies—patient care, medical knowledge, practice-based learning and improvement, interpersonal and communication skills, professionalism, and systems-based practice (Table 1-1)—that each contain principles of leadership. The ACGME has mandated the teaching of these core competencies but has not established a formal guide on how to teach the leadership skills described within the core competencies. Therefore, this chapter offers a review of fundamental principles of leadership and an introduction of the concept of a leadership training program for future surgeons.

DEFINITIONS OF LEADERSHIP

Many different definitions of leadership have been described. Former First Lady Rosalynn Carter once observed that “A leader takes people where they want to go. A great leader takes people
Key Points

1. Effective surgical leadership improves patient care, safety, and clinical outcomes.
2. A fundamental principle of leadership is to provide a vision that people can live up to, thereby providing direction and purpose to the constituency.
3. Surgical leaders have the willingness to lead through an active and passionate commitment to the vision.
4. Surgical leaders have the willingness to commit to lifelong learning.
5. Surgical leaders have the willingness to communicate effectively and resolve conflict.
6. Surgical leaders must practice effective time management.
7. Different leadership styles are tools to use based on the team dynamic.
8. Surgical trainees can be taught leadership principles in formal leadership training programs to enhance their ability to lead.
9. Mentorship provides wisdom, guidance, and insight essential for the successful development of a surgical leader.

Levels of Leadership

When working toward organizational success, strong leadership is a critical component. The best study of the relationship between leadership skill and organizational success is in the field of business. In business, the processes of customer satisfaction, product development, and organization efficiency are the equivalent of patient satisfaction, medical advancement, and efficient delivery of care. Jim Collins, author of *Good to Great*, studied the success and leadership styles of Fortune 500 companies over a 30-year period. He found that leadership is strongly correlated with corporate success, and most importantly for our study, that leadership strength can be broken down by level and characteristic (See figure 1-1).

Of 11 particularly outstanding organizations identified, great leadership was the single major defining characteristic that distinguished them from their peers. These organizations were led by what Collins called the “Level 5 Leader,” one whose personal humility and professional will drove team success. Under this system of leadership study, surgeon-leaders begin at the bottom level and, through study, hard work, and professional development, advance to the ultimate level of leadership.

### Table 1-1

<table>
<thead>
<tr>
<th>CORE COMPETENCY</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient care</td>
<td>To be able to provide compassionate and effective healthcare in the modern-day healthcare environment</td>
</tr>
<tr>
<td>Medical knowledge</td>
<td>To effectively apply current medical knowledge in patient care and to be able to use medical tools (i.e., PubMed) to stay current in medical education</td>
</tr>
<tr>
<td>Practice-based learning and improvement</td>
<td>To critically assimilate and evaluate information in a systematic manner to improve patient care practices</td>
</tr>
<tr>
<td>Interpersonal and communication skills</td>
<td>To demonstrate sufficient communication skills that allow for efficient information exchange in physician-patient interactions and as a member of a healthcare team</td>
</tr>
<tr>
<td>Professionalism</td>
<td>To demonstrate the principles of ethical behavior (i.e., informed consent, patient confidentiality) and integrity that promote the highest level of medical care</td>
</tr>
<tr>
<td>Systems-based practice</td>
<td>To acknowledge and understand that each individual practice is part of a larger healthcare delivery system and to be able to use the system to support patient care</td>
</tr>
</tbody>
</table>

**FUNDAMENTAL PRINCIPLES OF LEADERSHIP**

Leadership is a complex concept. Surgeons should strive to adopt leadership qualities that provide the best outcomes for their patients, based on the following fundamental principles: vision, willingness, time management, conflict resolution,
recruitment, and culture (See Table 1-2). Surgeon-leaders will develop a team of faculty, residents, and other healthcare personnel who are aligned on mission, vision, and values. The team and leader must be willing to address complex problems with honest communication and well-developed conflict resolution skills. A culture must be established where faculty and staff will work towards the advancement of the medical arts and the greater good of society.9

**Vision**

The first and most fundamental principle of leadership is to establish a vision that people can live up to, thus providing direction and purpose to the constituency. Creating a vision is a declaration of the near future that inspires and conjures motivation.10 A classic example of a powerful vision that held effective impact is President Kennedy’s declaration in 1961 that “... this nation should commit itself to achieving the goal, before this decade is out, of landing a man on the moon and returning him safely to the earth.” Following his declaration of this vision with a timeline to achieve it, the United Sates mounted a remarkable unified effort, and by the end of the decade, Neil Armstrong took his famous walk and the vision had been accomplished (Fig. 1-2).

On a daily basis, surgeons are driven by a powerful vision: the vision that our surgical care will improve patients’ lives. The great surgical pioneers, such as Hunter, Lister (Fig. 1-3), Halsted, von Langenbeck, Billroth, Kocher (Fig. 1-4), Carrel, Gibbon, Blalock, Wangensteen, Moore, Rhoads, Huggins, Murray, Kountz, Longmire, Starzl, and DeBakey (Fig. 1-5), each possessed a vision that revolutionized the field of surgery. In the 19th century, Joseph Lister changed the practice of surgery with his application of Pasteur’s germ theory. He set a young boy’s open compound leg fracture, a condition with a 90% mortality rate at that time, using carbolic acid dressings and aseptic surgical technique. The boy recovered, and Lister gathered nine more patients. His famous publication on the use of aseptic technique introduced the modern era of sterile technique. Emil Theodor Kocher was the first to master the thyroidectomy, thought to be an impossible operation at the time, and went on to perform thousands of thyroidectomies with a mortality of less than 1%. He was awarded the Nobel Prize in Physiology or Medicine in 1909 for describing the thyroid’s physiologic role in metabolism. Michael E. DeBakey’s powerful vision led to the development of numerous groundbreaking procedures that helped pioneer the field of cardiovascular surgery. For example, envisioning an artificial artery for arterial bypass operations, Dr. DeBakey invented the Dacron graft, which has helped millions of patients suffering from vascular disease and enabled the development of endovascular surgery. Dr. Frederick Banting, the youngest recipient of the Nobel Prize in Physiology or Medicine, had a vision to discover the biochemical link between diabetes and glucose homeostasis. His vision and perseverance led to the discovery of insulin.11 In retrospect, the power and clarity of their visions were remarkable, and their willingness and dedication were inspiring. By studying their careers and accomplishments, surgical trainees can be inspired by the potential impact of a well-developed vision.

**Table 1-2**

<table>
<thead>
<tr>
<th>LEADERSHIP SKILL</th>
<th>DESCRIPTION AND APPLICATION IN THE FIELD OF MEDICINE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vision</td>
<td>The act of establishing tangible goals of care for patients on both a daily basis as well as for long-term purposes.</td>
</tr>
<tr>
<td>Effective communication</td>
<td>Establishing an open, respectful, and nonjudgmental forum for communication among different members of the healthcare team and with the patient.</td>
</tr>
<tr>
<td>Willingness to lead</td>
<td>Taking on full responsibility for the care of patients and remaining ethical, professional, and committed despite the especially challenging rigors of joining the field of surgery.</td>
</tr>
<tr>
<td>Willingness to learn</td>
<td>A commitment to lifelong learning of the latest scientific, medical, and surgical updates to deliver optimized patient care.</td>
</tr>
<tr>
<td>Conflict resolution</td>
<td>The art of resolving conflicts in a peaceful and ethical manner in team settings.</td>
</tr>
</tbody>
</table>

**Figure 1-2.** Apollo 11 Lunar Module moon walk. Astronaut Edwin “Buzz” Aldrin walks by the footpad of the Apollo 11 Lunar Module, July 1969. (Reproduced with permission from AP Photo/ NASA. © 2018 The Associated Press.)

**Figure 1-3.** Joseph Lister directing use of carbolic acid spray in one of his earliest antiseptic surgical operations, circa 1865. (Used with permission from Getty Images.)
Generating Belief in Your Vision

Surgical leaders with great visions will inevitably require help from colleagues, other healthcare professionals, scientists, administrators, patients, and nonmedical personnel. To get this help, surgical leaders must inspire their team and understand motivation. For the surgeon-leader, it is critical to know that people do not follow leaders because of what they do; people follow leaders because of why they do what they do. The people who help the leader execute the vision are motivated by the leader’s beliefs and attitudes more than the leader’s policy or agenda. This concept, based on Simon Sinek’s *Start With Why*, is rooted in understanding of the anatomy and function of the human brain. See figure 1-6.

For example, take a surgeon-leader who wants to implement a new perioperative checklist to reduce surgical errors. The “what” is very simple: a checklist to reduce errors. The operating room team may make a rational decision to adopt the checklist; however, it is also possible that the checklist may be perceived as “another piece of paperwork” and rejected, or that the checklist may have its implementation fought, undermined, delayed, or ignored. A surgeon-leader who does not understand how people are motivated might argue rationally, telling the team that the checklist was created with great care, that all of the best evidence was incorporated in its creation, and that the checklist is short and efficient. This is the “how,” and once again it appeals to the rational and analytical side of the team. With these arguments, the surgeon-leader’s vision remains susceptible to rejection for many of the same reasons. A leader who understands how to motivate a team towards a vision will start with “why.” Before ever discussing the checklist in detail with the team, the leader will speak of their shared mission to offer the best patient care possible, ask the team to imagine how they might want a family member treated, and emphasize that a careless error could lead to patient harm and embarrassment for the team. With these arguments, which constitute an emotional appeal to the team’s belief system, the leader can expect this vision for better patient care via a new surgical checklist to be adapted by the team. The team will be receptive to implementing a new checklist, not because they believe in the checklist as a tool, but because they believe in the surgeon-leader’s vision for optimizing patient care.

There is a biological reason why this concept works. “Why,” “how,” and “what” are correlated to the functions

---

**Figure 1-5.** Michael E. DeBakey. (Reproduced with permission from AP Photo/David J. Phillip. © 2018 The Associated Press.)

---

**Figure 1-4.** Emil Theodor Kocher. (Reproduced with permission from the National Library of Medicine.)

---

**Figure 1-6.** When leaders seek to generate belief in their vision, it is best to appeal to the team with “why” statements. (Data from Sinek S. *Start with why*: how great leaders inspire everyone to take action. London: Portfolio/Penguin; 2013.)
of separate anatomical levels in the human brain. The neocortex is, evolutionarily, the newest area of our brains, and it is responsible for the analytical and rational thoughts and decisions that we make. It corresponds to the “what” and the “how.” When the surgeon-leader in the previous example started with the checklist and its rational arguments, the leader was appealing to their team’s neocortex, and the vision was rejected. However, when the surgeon started with the “why,” the vision for better patient care was emotionally accepted by the team, who became receptive to the checklist as a tool for achieving the vision.

Surgery is a field that requires extraordinary dedication and great personal sacrifice. The very nature of vision—steps forward into a better future—implies that change and difficult work will be required of the leader. See figure 1-7. For this reason, surgeon-leaders should establish visions about which they are deeply passionate and committed so that when obstacles are encountered the leader has the strength of will to progress forward. Leaders should be selective about which options they pursue. Each opportunity and idea requires great effort to execute; ultimately only a few can be brought to completion. Therefore, leaders should understand what drives their organization’s economic engine: the ideas and opportunities that will bring patients better care, bring the organization more patients, and create new treatments, etc. Thousands of hospitals, companies, innovators, and physicians are addressing many of the same problems in healthcare, such as growing burdens of chronic disease, an aging population, and rising health costs. The best opportunities lie where talent and ability align, so leaders and organizations should be cognizant of choosing projects for which they have the potential to be the “best in the world” at doing. Once the vision is set and the project is chosen, it is up to the leader to generate momentum.

Momentum is either a cumulative effect of continuous steps towards improvement or, alternatively, in the negative sense, movements towards failure or stagnation. The “flywheel effect,” depicted in Fig. 1-8, demonstrates the building of momentum with (a) initial steps forward, (b) an accumulation of visible results, (c) realignment of the team in the new direction (accounting for new information and data), and then (d) an accumulation of momentum followed by more steps forward. Careful attention to the aforementioned principles is essential in building a successful surgical career, department, or division.

**Willingness**

The Willingness Principle represents the active commitment of the leader toward his or her vision. To do so, a surgical leader must be willing to lead, commit to lifelong learning, communicate effectively, and resolve conflict.

**To Lead.** A key characteristic of all great leaders is the willingness to serve as the leader. Dr. Martin Luther King Jr, who championed the civil rights movement with a powerful vision of equality for all based on a commitment to nonviolent methods, did so at a time when his vocalization of this vision ensured harassment, imprisonment, and threats of violence against himself, his colleagues, and his family and friends (Fig. 1-9). King, a young, highly educated pastor, had the security of employment and family, yet was willing to accept enormous responsibility and personal risk and did so in order to lead a nation toward his vision of civil rights, for which he was awarded the Nobel Peace Prize in 1964.

Willingness to lead is a necessity in any individual who desires to become a surgeon. By entering into the surgical theater, a surgeon accepts the responsibility to care for and operate on patients, despite the risks and burdens involved. They do so, believing fully in the improved quality of life that can be achieved. Surgeons must embrace the responsibility of leading surgical teams that care for their patients, as well as leading surgical trainees to become future surgeons. A tremendous sacrifice is required for the opportunity to learn patient care. Surgical trainees accept the hardships of residency with its accompanying steep learning curve, anxiety, long work hours, and time spent away from family and friends. The active, passionate commitment to excellent patient care reflects a natural willingness to lead based on altruism and a sense of duty toward those receiving care. Thus, to ensure delivery of the utmost level of care, surgical trainees should commit to developing and refining leadership skills. These skills include a commitment to lifelong learning, effective communication, and conflict resolution.

**To Learn.** Surgeons and surgical trainees, as leaders, must possess willingness to commit to continuous learning. Modern surgery is an ever-changing field with dynamic and evolving healthcare systems and constant scientific discovery and innovation. Basic and translational science relating to surgical care is growing at an exponential rate. The sequencing of the human
Basic Considerations

Part I

Genome and the enormous advances in molecular biology and signaling pathways are leading to the transformation of pre-cinematic medicine and personalized surgery in the 21st century (see Chapter 15). Performing prophylactic mastectomies with immediate reconstruction for BRCA1 mutations and thyroidectomies with thyroid hormone replacement for RET proto-oncogene mutations are two of many examples of genomic information guiding surgical care. Technologic advances in minimally invasive surgery and robotic surgery as well as electronic records and other information technologies are revolutionizing the craft of surgery. The expansion of minimally invasive and endovascular surgery over the past three decades required surgeons to retrain in new techniques using new skills and equipment. In this short time span, laparoscopy and endovascular operations are now recognized as the standard of care for many surgical diseases, resulting in shorter hospital stay, quicker recovery, and a kinder and gentler manner of practicing surgery. Remarkably, during the last century, the field of surgery has progressed at an exponential pace and will continue to do so with the advent of using genomic analyses to engineer cancer cells with molecular imaging agents that will guide personalized surgery, which will transform the field of surgery during this century. Therefore, surgical leadership training should emphasize and facilitate the continual pursuit of knowledge.

Willingness to learn encompasses the surgeon’s commitment to lifelong learning. This has been exemplified by the surgeons of the past several decades who have dedicated their peak practicing years to perfecting minimally invasive surgical techniques, including the use of robotic surgery. The field continues to advance, offering many advantages to patients including faster recovery, sometimes decreased pain depending on procedure type, and shorter hospital stays.16-18

Fortunately, surgical organizations and societies provide surgeons and surgical trainees a means to acquire new knowledge on a continuous basis. There are numerous local, regional, national, and international meetings of surgical organizations that provide ongoing continuing medical education credits, also required for the renewal of most medical licenses. The American Board of Surgery requires all surgeons to complete meaningful continuing medical education to maintain certification.19 These societies and regulatory bodies enable surgeons and surgical trainees to commit to continual learning and ensure their competence in a dynamic and rapidly growing field.

Surgeons and trainees now benefit from the rapid expansion of web-based education as well as mobile handheld technology. These are powerful tools to minimize nonproductive time in the hospital and make learning and reinforcement of medical knowledge accessible. Currently web-based resources provide quick access to a vast collection of surgical texts, literature, and surgical videos. Surgeons and trainees dedicated to continual learning should be well versed in the utilization of these information technologies to maximize their education. The next evolution of electronic surgical educational materials will likely include simulation training similar to laparoscopic and Da Vinci device training modules. The ACGME, acknowledging the importance of lifelong learning skills and modernization of information delivery and access methods, has included them as program requirements for residency accreditation.

To Communicate Effectively. The complexity of modern healthcare delivery systems requires a higher level and collaborative style of communication. Effective communication directly impacts patient care. In 2000, the U.S. Institute of Medicine published To Err Is Human: Building a Safer Health System, which raised awareness concerning the magnitude of medical errors. This work showcased medical errors as the eighth leading cause of death in the United States with an estimated 100,000 deaths annually.20 Subsequent studies examining medical errors have identified communication errors as one of the most common causes of medical error.21-23 In fact, the Joint Commission identifies miscommunication as the leading cause of sentinel events. Information transfer and communication errors cause delays in patient care, waste surgeon and staff time, and cause serious adverse patient events.23 Effective communication among surgeons, nurses, ancillary staff, and patients is not only a crucial element to improved patient outcomes, but it also leads to less medical litigation.24-26 A strong correlation exists between communication and patient outcomes.

Establishing a collaborative atmosphere is important since communication errors leading to medical mishaps are not simply failures to transmit information. Communication errors “are far more complex and relate to hierarchical differences, concerns with upward influence, conflicting roles and role ambiguity, and interpersonal power and conflict.”22,27-29 Errors frequently originate from perceived limited channels of communication and hostile, critical environments. To overcome these barriers, surgeons and surgical trainees should learn to communicate in an open, universally understood manner and remain receptive to any team member’s concerns. A survey of physicians, nurses, and ancillary staff identified effective communication as a key element of a successful leader.30 As
leaders, surgeons, and surgical trainees who facilitate an open, effective, and collaborative style of communication can reduce errors and enhance patient care. A prime example is that successful communication of daily goals of patient care from the team leader improves patient outcomes. In one recent study, the modest act of explicitly stating daily goals in a standardized fashion significantly reduced patient length of intensive care unit stay and increased resident and nurse understanding of goals of care.31 Implementing standardized daily team briefings in the wards and preoperative units led to improvements in staff turnover rates, employee satisfaction, and prevention of wrong-site surgery.27 In cardiac surgery, improving communication in the operating room and transition to the postanesthesia care unit was an area identified to decrease risk for adverse outcomes.32 Behaviors associated with ineffective communication, including absence from the operating room when needed, playing loud music, making inappropriate comments, and talking to others in a raised voice or a condescending tone, were identified as patient hazards; conversely, behaviors associated with effective collaborative communication, such as leading the time-out process and closed-loop communication technique, resulted in improved patient outcomes.

One model to ensure open communication is through standardization of established protocols. A commonly accepted protocol is the “time out” that is now required in the modern operating room. During the time-out protocol, all team members introduce themselves and state a body of critical information needed to safely complete the intended operation. This same standardization can be taught outside the operating room. Within the Kaiser system, certain phrases have been given a universal meaning: “I need you now” by members of the team is an understood level of urgency and generates a prompt physician response 100% of the time.27 As mentioned earlier, standardized forms can be useful tools in ensuring universally understood communication during sign-out. The beneficial effect of standardized team communication further demonstrates how effective communication can improve patient care and is considered a vital leadership skill.

Effective communication with patients in the modern era, necessitates understanding that many patients access health information via the internet and that patients are often ill equipped to evaluate the individual source.33,34 Discrepancies exist between surgeon’s self-perceived ability to communicate and patient’s actual satisfaction. A patient’s perceived interaction with their physician has an enormous impact on patient health outcomes, malpractice, and financial reimbursement;35,36 specifically, the association between poor doctor–patient communication and a patient’s perception that their doctor does not care about them. Good bedside manner has been shown to decrease litigation even in situations of error or undesirable outcome.37,38 Physicians who demonstrate concern, actively know their patients, and share responsibility for decision-making are more likely to be trusted by their patients.26,41,42 Strong doctor–patient relationships and effective communication skills have been incentivized by the Agency for Healthcare Research and Quality and the Centers for Medicare & Medicaid Services through their Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) and Clinical and Group Consumer Assessment of Healthcare Providers and Systems (CGCAHPS) programs, which measure patient satisfaction.43

To Resolve Conflict. Great leaders are able to achieve their vision through their ability to resolve conflict. Delivery of modern surgical care is complex; numerous conflicts arise on a daily basis when surgeons and surgical trainees provide high-quality care. Therefore, the techniques for conflict resolution are essential for surgical leaders.

To properly use conflict resolution techniques, it is important for the surgeon and surgical trainee to always remain objective and seek personal flexibility and self-awareness. The gulf between self-perception and the perception of others can be profound; in a study of cooperation and collaboration among operating room staff, the quality of their own collaboration was rated at 80% by surgeons, yet was rated at only 48% by operating room nurses.44 Systematic inclusion of modern conflict resolution methods that incorporate the views of all members of a multidisciplinary team help maintain objectivity. Reflection is often overlooked in surgical residency training, but it is a critical component of learning conflict resolution skills. Introspection allows the surgeon to understand the impact of his or her actions and biases. Objectivity is the basis of effective conflict resolution, which can improve satisfaction among team members and help deliver optimal patient care.

Modern conflict resolution techniques are based on objectivity, willingness to listen, and pursuit of principle-based solutions.45 For example, an effective style of conflict resolution is the utilization of the “abundance mentality” model, which attempts to achieve a solution that benefits all involved and is based on core values of the organization, as opposed to the utilization of the traditional fault-finding model, which identifies sides as right or wrong.46 Application of the abundance mentality in surgery elevates the conflict above the affected parties and focuses on the higher unifying goal of improved patient care. “Quality Improvement” (previously or alternatively “Morbidity and Mortality”) conferences are managed in this style and have the purpose of practice improvement and improving overall quality of care within the system, as opposed to placing guilt or blame on the surgeon or surgical trainees for the complication being reviewed. The traditional style of command-and-control technique based on fear and intimidation is no longer welcome in any healthcare system and can lead to sanctions, lawsuits, and removal of hospital privileges or position of leadership.

Another intuitive method that can help surgical trainees learn to resolve conflict is the “history and physical” model of conflict resolution. This model is based on the seven steps of caring for a surgical patient that are well known to the surgical trainee: (a) the “history” is the equivalent of gathering subjective information from involved parties with appropriate empathy and listening; (b) the “laboratory/studies” are the equivalent of collecting objective data to validate the subjective information; (c) a “differential diagnosis” is formed out of possible root causes of the conflict; (d) the “assessment/plan” is developed in the best interest of all involved parties; the plan, including risks and benefits, is openly discussed in a compassionate style of communication; (e) “preoperative preparation” includes the acquisition of appropriate consultations for clearances, consideration of equipment and supplies needed for implementation, and the “informed consent” from the involved parties; (f) the “operation” is the actual implementation of the agreed-upon plan, including a time-out; (g) and “postoperative care” involves communicating the operative outcome, regular postoperative follow-up, and the correction of any complications that arise. This seven-step method is an example of an objective, respectful method of conflict resolution.47 Practicing different styles of conflict resolution and effective communication in front of the entire group of
surgical trainees attending the leadership training program is an effective means of teaching conflict resolution techniques.

**Time Management**

It is important for leaders to practice effective time management. Time is the most precious resource, as it cannot be bought, saved, or stored. Thus, management of time is essential for a productive and balanced life for those in the organization. The effective use of one’s time is best done through a formal time management program to improve one’s ability to lead by setting priorities and making choices to achieve goals. The efficient use of one’s time helps to improve both productivity and quality of life.48-50

It is important for surgeons and surgical trainees to learn and use a formal time-management program. There are ever-increasing demands placed on surgeons and surgical trainees to deliver the highest quality care in highly regulated environments. Furthermore, strict regulations on limitation of work hours demand surgical trainees learn patient care in a limited amount of time.48-50 All told, these demands are enormously stressful and can lead to burnout, drug and/or alcohol abuse, and poor performance.48-50 A time-motion study of general surgery trainees analyzed residents’ self-reported time logs to determine resident time expenditure on educational/service-related activities (Fig. 1-10).30 Surprisingly, senior residents were noted to spend 13.5% of their time on non-service/mainly educational value activities (Fig. 1-10).50

Time is the most precious resource, as it cannot be bought, saved, or stored. Thus, management of time is essential for a productive and balanced life for those in the organization. The effective use of one’s time is best done through a formal time management program to improve one’s ability to lead by setting priorities and making choices to achieve goals. The efficient use of one’s time helps to improve both productivity and quality of life.48-50

It is important for surgeons and surgical trainees to learn and use a formal time-management program. There are ever-increasing demands placed on surgeons and surgical trainees to deliver the highest quality care in highly regulated environments. Furthermore, strict regulations on limitation of work hours demand surgical trainees learn patient care in a limited amount of time.48-50 All told, these demands are enormously stressful and can lead to burnout, drug and/or alcohol abuse, and poor performance.48-50

A frequently used basic technique is the “prioritized list,” also known as the ABC technique. Individuals list and assign relative values to their tasks. The use of the lists and categories serves solely as a reminder, thus falling short of aiding the user in allocating time wisely. Another technique is the “time management matrix technique.”49 This technique plots activities on two axes: importance and urgency, yielding four quadrants (Fig. 1-11). Congruous with the Pareto’s 80/20 principle and Parkinson’s law, the time management matrix technique channels efforts into quadrant II (important but nonurgent) activities. The activities in this quadrant are high yield and include planning, creative activity, building relationships, and maintaining productivity. Too often, surgeons spend a majority of their time attending to quadrant I (important and urgent) tasks. Quadrant I tasks include emergencies and unplanned or disorganized situations that require intensive and often inefficient effort. While most surgeons and surgical trainees have to deal with emergencies, they often develop the habit of inappropriately assigning activities into quadrant I; excess time spent on quadrant I tasks leads to stress or burnout for the surgeon and distracts from long-term goals. Efficient time management allows surgeons and surgical trainees to be proactive about shifting energy from quadrant I tasks to quadrant II, emphasizing preplanning and creativity over always attending to the most salient issue at hand, depending on the importance and not the urgency.

Finally, “the six areas of interest” is an alternative effective time management model that can help surgeons and surgical trainees achieve their goals, live a better-balanced lifestyle, and improve the quality of their lives.49 The process begins by performing a time-motion study in which the activities of 6-hour increments of time over a routine week are chronicled. At the end of the week, the list of activities is analyzed to determine how the 168 hours in 1 week have been spent. The surgical trainee then selects six broad categories of areas of interest.
LEADERSHIP IN SURGERY
CHAPTER 1
(i.e., family, clinical care, education, health, community service, hobbies) and sets a single activity goal in each category every day and monitors whether those goals are achieved. This technique is straightforward and improves one’s quality of life by setting and achieving a balanced set of goals of personal interest, while eliminating time-wasting activities.

A formal time management program is essential for modern leadership. The practice and use of time management strategies can help surgeons and surgical trainees achieve and maintain their goals of excellent clinical care for their patients, while maintaining a more balanced lifestyle.

Self-Care and Wellness
The challenges of practicing medicine place unique stresses on surgeons. A departmental program for improving wellness and teaching self-care can help alleviate these stresses. Acknowledging these stresses is an important step for any leader to help peers at risk. Quality of life surveys have identified individual protective factors that can be implemented prophylactically. These factors for improving self-care and wellness include regular exercise programs, maintenance of routine medical care, and health screening. The following may not apply to all physicians; however, religious practices, reflective writing, and maximizing work-life balance have also been demonstrated to be protective.

Surgeons and physicians overall experience increased rates of suicide, depression, substance abuse, marital and family problems, and other stress-related health effects as compared to the general population. Suicide rates in physicians are higher among those who are divorced, widowed, or never married. Depression is a common challenge, with rates as high as 30% among trainees, and higher when lifetime risk is considered. Drug and alcohol abuse among physicians mirrors the general population; however, physicians have higher rates of prescription drug abuse. The ability to self-medicate likely contributes to prescription drug abuse by physicians. Divorce and marriage unhappiness among physicians has been attributed to the “psychology of postponement,” compulsive personality traits that are reinforced and selected for during medical training, and lack of work-life balance. Residents, due to their inexperience, may be at higher risk than practicing physicians. For physicians who do not seek professional help, fear of losing their medical license is the most commonly provided reason. Departmental wellness programs may provide an alternative source of support for these surgeons.

The past 10 years have seen a significant increase in attention to the issue of physician wellness. Physician wellness has become an issue transcending specialties and resulting in significant research. The creation of wellness and self-care programs within departments represents an opportunity for surgeons to demonstrate leadership qualities.

Recruitment
The challenges of modern medicine and ever-larger medical centers have created a reality where no single surgeon-leader can exercise complete control—it takes a team of leaders with shared vision, mission, and goals. To this end, the previously discussed “level 5 leader” who embodies personal humility and professional will is essential. Previous generations whose leaders and departments were composed of self-proclaimed giants dominated and suppressed alternative points of view, communication, and innovation. In recent years, there has been a change to building teams with authentic leaders who have high ethical standards and well-developed nontechnical skills, who lead by example, and who never compromise excellence. The surgeon-leader must build a team where talented individuals are placed in the right job for their skills. The essence of a leader is one who enables others to succeed. Team work is imperative to change, and trust is the make-or-break component. Simply put, teams that trust each other work well, and teams that do not trust each other do not work well.

Creating a Culture of Empathy, Patient-Family-Centered Care, and Personalized Surgery
Creating the right culture is the most challenging of all the surgeon-leader’s tasks. Modern surgical departments should focus on creating a culture of empathy, patient–family-centered care, and personalized surgery. Instilling a positive culture requires both discipline and consistency because it may take considerable time to change how people think, feel, and behave. Organizational culture is built around the leader’s vision and values. Coming up with strong values requires genuine commitment. A leader should realize that staying true to his or her values can be challenging when conflicts arise.

WHY WE LEAD
Choosing to Become a Leader
There are many benefits to becoming a leader. Humankind has pondered the question of whether leaders are born or made for millennia. The best evidence to date indicates that leaders are both born and made. Leadership potential is a skill that all persons are born with, to some degree, and that can be formally trained, learned through observation, and honed with practice. The positive effects of a leader on others are innumerable, including a leader’s positive influence on innovation, diversity, culture, and quality. For modern surgeons, leadership skills are essential for the delivery of quality patient care; therefore, it is the duty of the surgeon to study leadership.

For the surgeon studying to be a better leader, effective leadership also has many individual benefits, including recognition from one’s peers, promotion, and autonomy. Modern leaders are increasingly required to be humble about their accomplishments in order to be successful and effective. Beyond recognition, promotion, and autonomy there are more selfless reasons for surgeons to desire leadership. Leadership is a tool to help make a difference. Leadership is a good path towards a career as an educator, which offers the leader a sense of accomplishment and satisfaction in seeing others succeed. Some choose to become leaders out of a sense of selfless service, taking on leadership for the benefit of others, or out of a desire to solve problems. Leadership may come with material rewards, including wealth and power, which motivate some.

Whatever the motivation, surgeons, in their role as leaders of patient care teams, have a duty to develop some skill in leadership. It would be best for individuals, departments, and patients if all surgeons sought to develop leadership skills and experience in some area of administration, patient care, education, or research. The benefits to the individual are numerous.

Leadership’s Effect on Healthcare Cost and Clinical Outcomes
Much attention has rightly been paid to historical leaders’ impact on humanity. Surgical leaders of the past have made great contributions on which we may build. All surgeons have a responsibility to be leaders, whether at the team level or in
an administrative or organizational capacity. To that end, it is worth noting the benefits of formal leadership education.

Large observational studies using trained observers assessed the effects of different surgical leadership styles on operative cases. Team cohesion and collective efficiency were reduced when leaders utilized abusive supervision or over-controlling methods. Abusive supervision alone was associated with decreased “psychological safety.” Surgeons perceived having positive leadership characteristics by their staff have lower 30-day all-cause mortality. This is likely due to creating a culture of safety where the staff can speak up if they notice an error and feel they have the latitude to do what is best for the patient quickly and autonomously.

With increased recognition and attention on human error, nontechnical skills, including leadership, play a role in patient safety. The landmark study, “To Err Is Human,” estimated that almost 100,000 people die each year due to medical errors. In the surgical setting, 40% to 50% of errors may be attributed to communication breakdown. The Multifactor Leadership Questionnaire scores subjects on their demonstration of transformational leadership behaviors. Transformational leaders exhibit the qualities of charisma, inspired motivation, intellectual stimulation, and individualized consideration. In video analysis of complex surgical operations, surgeons scoring even a single point higher on the transformational leadership score exhibited 3 times more information sharing behaviors, 5 times more positive voice behaviors, and 10 times more supportive behaviors, all while displaying poor behaviors 12.5 times less frequently than their peers. Exhibiting the characteristics of transformational leadership clearly has much to offer the surgeon-leader in preventing serious errors.

The field of trauma contains the largest body of formal study demonstrating the positive effects of leadership on clinical results. Strong leadership skills improve both the speed of resuscitation and completion of the initial trauma evaluation. There is no one optimal style of leadership covering all situations; some call for a more empowering leadership style while others call for a more directive style. The optimal style of leadership varies based on team composition, with less experienced teams better responding to the directive style, while more experienced teams work faster with trust and an empowering style. The formally educated surgical leader should be able to switch easily between styles based on the situation at hand.

Leadership styles affect responses to patient safety concerns and protect the organization as a whole. The surgical leader adopts a supervisory capacity while creating a culture of safety. In detail, frontline staff must be encouraged to participate in safety improvement. Staff ownership of safety must be established and upheld. In order to assure this outcome, whistle-blowers must be protected. A culture of psychological safety, organizational fairness, and continuous learning is required. Subordinates require appropriate authority, autonomy, and latitude to do their jobs and care for patients.

Formal leadership training has been well studied within the Veteran’s Health Administration system using the Surgical Care Improvement Program. The Medical Team Training Program, for instance, has been shown to result in a 18% decrease in 30-day mortality and a 17% decrease in 30-day morbidity.

Also at the organizational level, leaders using an empowering style may improve process of care protocols and increase efficiency. Operating room turnover times specifically have been shown to be reducible. Value-based purchasing benchmarks, such as hospital-acquired infections, which affect reimbursement, can be reduced or eliminated depending on the measure. Medical errors may be reduced, and significant medical errors may have their effects mitigated. Patient satisfaction may be improved. The overall financial performance of the institution can be affected in a positive manner.

There are positive correlations between mutual respect, clinical leadership, and surgical safety. Traditional command and control style leadership negatively impacts psychological safety resulting in the development of more modern leadership styles. The best clinical processes have the potential to break down when there is a toxic work environment and lack of psychological safety within the team.

The Importance of Diversity and Leadership

The past quarter century has seen a steady increase in diversity within the field of surgery. Women, as of 2015, represent 38% of surgical trainees and 10% of academic professors currently, but have doubled their representation in the past 20 years. Some fields, such as head and neck surgery and plastic surgery, have studied their own subspecialty groups with similar findings. African Americans comprise both 6% of medical school graduates, 6% surgical trainees, and 2% to 4% of professors of surgery nationwide. Hispanics represent 5% of graduating medical students, 9% of general surgery trainees, and 4% to 5% of persons at all levels of academic surgery. Physician diversity is crucial and may help to address disparities in social determinants of health.

Studies indicate that the bottleneck in diversity occurs at the level of the medical school application pool, which in turn is caused by educational deficiencies at the primary, secondary, and collegiate level. As an attempted solution, the University of Michigan developed a “pipeline” program that pairs grade-school and high-school students with physicians for experiential learning and the development of mentoring, presentation skills, and networking. It is important for departments of surgery to develop a diversity program for recruitment of residents and faculty. Multi-institutional blinded studies indicate that the implementation of formal leadership and diversity training improves diversity leadership and strategic human resource management.

Leadership Styles

The principles of leadership can be practiced in a variety of styles. Just as there are many definitions of leadership, many classifications of styles exist as well. A landmark study by Daniel Goleman in Harvard Business Review identified six distinct leadership styles, based on different components of emotional intelligence. Emotional intelligence is the ability to recognize, understand, and control the emotions in others and ourselves. By learning different styles, surgeons and trainees can recognize their own leadership style and the effect on the team dynamic. Furthermore, it teaches when the situation may demand change in style for the best outcome. The six leadership styles identified are coercive, authoritative, affiliative, democratic, pacesetting, and coaching.

The coercive leader demands immediate compliance. This style reflects the command and control style that has historically dominated surgery. Excessive coercive leadership erodes team members’ sense of responsibility, motivation, sense of participation in a shared vision, and ultimately, performance. The phrase, “Do what I tell you!” brings to mind the coercive leader. However, it is effective in times of crisis to deliver clear,
concise instruction. This style should be used sparingly and is best suited for emergencies.

The authoritative leader embodies the phrase “Come with me,” focusing on mobilizing the team toward a common, grand vision. This type of leader allows the team freedom to innovate, experiment, and devise its own means. Goleman’s research indicates this style is often the most effective. These leaders display self-confidence, empathy, and proficiency in initiating new ideas and leading people in a new direction. This is best used when a shift in paradigm is needed.

The affiliative leader creates harmony and builds emotional bonds. This requires employment of empathy, building relationships, and emphasis on communication. An affiliative leader frequently gives positive feedback. This style can allow poor performance to go uncorrected if too little constructive/critical advice is given. Affiliative leadership is most useful when motivating people during stressful circumstances or healing rifts in a team.

The coaching style of leadership focuses on developing people for the future. Coaching is leadership through mentorship. The coach gives team members challenging tasks, counsels, encourages, and delegates. Unlike the affiliative leader who focuses on positive feedback, the coach helps people identify their weaknesses and improve their performance, and ties their work into their long-term career aspirations. This leadership style builds team capabilities by helping motivated learners improve. However, this style does not work well when team members are defiant and unwilling to change or learn, or if the leader lacks proficiency.

The democratic leader forges consensus through participation. This leadership style listens to and values each member’s input. It is not the best choice in an emergency situation, when time is limited, or when teammates cannot contribute informed guidance to the leader. It can also be exasperating if a clear vision does not arise from the collaborative process. This style is most appropriate when it is important to obtain team consensus, quell conflict, or create harmony.

The pacesetter leader sets high standards for performance and exemplifies them. These leaders identify poor performers and demand more from them. However, unlike the coach, the pacesetter does not build the skills of those who are not keeping up. Rather, a pacesetter will either take over the task himself or delegate the task to another team member. This style can be summed up best by the phrase, “Do as I do, now.” This leadership style works well when it is important to obtain high-quality results and there is a motivated, capable team. However, pacesetters can easily become micromanagers who have difficulty delegating tasks to team members, which leads to burn out on the part of the leader. Additionally, team members can feel overwhelmed and demoralized by the demands for excellence without an empathic counter balance.

Each of the above styles of leadership has strengths and weakness. Importantly, leaders who are the most successful do not rely only on one leadership style alone. They use several of them seamlessly depending on the situation and the team members at hand. Therefore, the more styles a leader has mastered, the better, with particular emphasis on the authoritative, affiliative, democratic, and coaching styles. Each leadership style is a tool that is ultimately employed to guide a team to realizing a vision or goal. Thus, leadership training programs should teach the proper use of all leadership styles while adhering to the principles of leadership.

---

**Table 1-3**

<table>
<thead>
<tr>
<th>Eighteen leadership training modules</th>
<th>IMPORTANCE MEAN SCORE</th>
<th>COMPETENCE MEAN SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic program development</td>
<td>3.2</td>
<td>2.4*</td>
</tr>
<tr>
<td>Leadership training</td>
<td>3.8</td>
<td>2.3*</td>
</tr>
<tr>
<td>Leadership theory</td>
<td>3.2</td>
<td>2.1*</td>
</tr>
<tr>
<td>Effective communication</td>
<td>3.7</td>
<td>2.7*</td>
</tr>
<tr>
<td>Conflict resolution</td>
<td>3.8</td>
<td>3*</td>
</tr>
<tr>
<td>Management principles</td>
<td>3.7</td>
<td>2.7*</td>
</tr>
<tr>
<td>Negotiation</td>
<td>3.7</td>
<td>2.8*</td>
</tr>
<tr>
<td>Time management</td>
<td>4</td>
<td>2.8*</td>
</tr>
<tr>
<td>Private or academic practice, managed care</td>
<td>3.6</td>
<td>2*</td>
</tr>
<tr>
<td>Investment principles</td>
<td>3.5</td>
<td>2.2*</td>
</tr>
<tr>
<td>Ethics</td>
<td>3.6</td>
<td>3.2</td>
</tr>
<tr>
<td>Billing, coding, and compliance</td>
<td>3.5</td>
<td>1.7*</td>
</tr>
<tr>
<td>Program improvement</td>
<td>3</td>
<td>2*</td>
</tr>
<tr>
<td>Writing proposals</td>
<td>3.3</td>
<td>2.2*</td>
</tr>
<tr>
<td>Writing reports</td>
<td>3.4</td>
<td>2.4*</td>
</tr>
<tr>
<td>Public speaking</td>
<td>3.7</td>
<td>2.7*</td>
</tr>
<tr>
<td>Effective presentations</td>
<td>3.7</td>
<td>2.7*</td>
</tr>
<tr>
<td>Risk management</td>
<td>3.5</td>
<td>2.1*</td>
</tr>
<tr>
<td>Total</td>
<td>3.6</td>
<td>2.5*</td>
</tr>
</tbody>
</table>

*P <0.001 by Student t-test between mean importance and mean competence scores.

for physicians with an MBA or management background but have not been incorporated into the core residency training program.80 Also, there are many institutions that have published experiences with leadership retreats or seminars for residents or young physicians.31-34 The ACGME hosts multiple leadership skills workshops for chief residents, mostly targeted toward pediatricians, family practitioners, and psychiatrists.85 Similarly, the American College of Surgeons leads an annual 3-day leadership conference focusing on leadership attributes, consensus development, team building, conflict resolution, and translation of leadership principles into clinical practice.86-87 These programs were all received well by participants and represent a call for a formal leadership program for all surgical trainees.

An innovative leadership curriculum first implemented in 2000, prior to work-hour restrictions, taught general surgery trainees’ collaborative leadership skills at a time when the traditional command-and-control leadership style predominated.2,39,90 Surgical residents participated in 18-hour-long modules based on the leadership principles and skills listed in Table 1-2, taught by the surgical faculty. A number of leadership techniques, including time management techniques and applied conflict resolution techniques described earlier, were designed and implemented as part of this leadership training program. Within 6 months of implementation, residents’ self-perceived total commitment to the highest personal and professional standards, communication skills, visualization of clear missions of patient care, and leadership of others toward that mission increased significantly.2,89,90 Remarkably, the positive impact of this leadership curriculum was significant when measured using tools, such as the Multifactor Leadership Questionnaire (MLQ), social skills inventory, personality inventory, and internal strength scorecard.2,89,90 The MLQ is a well-validated instrument that objectively quantifies leadership beliefs and self-perceived outcomes across medical and nonmedical disciplines. Based on the MLQ, surgical residents more often use a passive-avoidance style of leadership that emphasizes taking corrective action only after a problem is “significant and obvious.” This tool can also be used to track progress toward more effective, collaborative styles of leadership. These studies demonstrated the ability to measure leadership behavior of surgical trainees in a standardized, quantifiable format.2,89,90 Taken together, these studies support the concept that leadership skills can and should be taught to surgical trainees, and there are validated tools to measure outcomes.

**Designing the Program**

Success in designing a formal leadership development program can be achieved through the following method. First, select the right participants at the right time in their career. Junior surgeons new to practice are ideal; however, they should be given a chance to get their clinical and research interests off the ground before they are asked to lead others. Candidates, should be open to taking on leadership roles and have the right combination of introspection and humility that lends to professional development. High-quality speakers from the business, legal, creative, and medical worlds should be brought as guest speakers. Topics could include leadership overall, strategy, finance, management skills, feedback, and coaching. Constructive criticism is essential because prospective leaders will need guidance and mentoring. Surgeons who have been through a formal leadership training program will become proficient at team-building skills and management and will become self-empowered individuals.91

Formal leadership training is not restricted to faculty alone. Leadership training should begin early and continue throughout residency. Surgical residents’ leadership styles have been studied in environments where they are given an assistant to supervise, as if they were an attending. Most residents were able to adapt to difficult operative challenges, in this setting, by providing a more directed style of leadership to their assistants. When faced with a less challenging task, or when the surgery resident’s confidence was particularly high, their leadership score was also high. For the surgical resident preparing to move on to the attending level, such skills are necessary to develop.92

Nontechnical surgical skills, such as leadership, demonstrate a number of desired effects for the operative team. Patient safety, including all cause 30-day mortality, is improved by stronger nontechnical skills.59 Development of clear and effective communication, situational awareness, team skills, and decision-making all are correlated with reduced surgical errors. Interruptions, such as needing to answer a page during an operation, are the only nontechnical factors in surgical error that are not directly attributable to leadership style.93

Surgical leaders have a responsibility to make ethical decisions. At this time, there is no standard curriculum to formally train surgical residents in ethics, despite interest from a majority of residency program directors.94-97 Several solutions have been proposed. A case-based approach to ethics training appears to have some merit, where monthly hour long ethical dilemmas are discussed in an informal, nonhierarchical setting.98 In another study, an ICU-based simulation model demonstrated promise for teaching compassion and end-of-life ethics to surgical residents. In this model, surgery residents have their first end-of-life conversations with standardized patients simulating the surgical ICU environment.99,100

**Practicing Leadership Skills and Assessing Leadership Formally With Objective Structured Clinical Examination (OSCE) and Simulation**

The past decade has seen a demonstrable increase in our knowledge of how to develop leadership skills, particularly through simulation, as well as leadership evaluation through OSCE and other tools. Multiple groups have assessed multidisciplinary teams, typically composed of nurses, anesthesia groups, and surgeons for the leadership associated nontechnical skills of communication, teamwork, and situational awareness. Through increasingly validated instruments and assessment tools, these nontechnical skills have been found to be trainable.101 The OSCE has been established as the gold standard102 for the training and assessment of a wide range of clinical and nontechnical skills with high reliability and validity.103-106

The OSCE was developed by Harden, at the Ninewells Hospital in Dundee, Scotland, and first published in 1975.107 He subsequently coined the term “OSCE” in his 1979 publication “Assessment of Clinical Competence Using an Objective Structured Clinical Examination (OSCE).”108 The purpose of the OSCE was to address the lack of a reliable method to evaluate the clinical abilities of physicians and featured a comprehensive assessment of history-taking and physical examination skills. Early versions also assessed nontechnical skills, patient interaction, and professionalism. Since its inception, the OSCE has matured, been subjected to rigorous tests of reliability and validity, and has seen widespread adoption.109-111

OSCEs remain a critical portion of resident evaluation. They have been well validated for teaching leadership skills in
trauma and interacting with simulated patients in difficult scenarios. OSCEs can be tailored to a variety of circumstances, including practicing breaking bad news or discussing end of life care, dealing with angry or aggressive patients, and simulating disagreements with other providers or family members.109-112 The potential for OSCEs to train, test, and perfect nontechnical skills, such as leadership, is extraordinary.

A pilot project for the Medical Council of Canada was conducted by the University of Toronto and published in 1988 describing the use of an OSCE for evaluating the clinical skills of international medical graduates applying to Canadian residency.113 Effective communication and language proficiency have been key components since the beginning. A comprehensive review of this program 2 years later confirmed the reliability and validity of using an OSCE for this purpose.114 The Medical Council of Canada has subsequently mandated a requirement for an OSCE evaluation of all international graduates applying for positions in Canada. In place for the past two decades, the program has ensured a baseline proficiency of skill, attitude, knowledge, and other nontechnical skills.115

OSCEs quickly gained acceptance as an established tool to assess learners in a comprehensive manner and became the inspiration for the creation of the USMLE Step 2 Clinical Skills (CS) examination, required for all U.S. medical students prior to licensure.116 Indeed, medical students whose schools use OSCE as practice do better on USMLE Step 2.117 The USMLE Step 2 CS examination meets the criteria, discussed in the following section, for a thorough and well-designed OSCE examination, due to its 12-station design which takes 8 hours to complete. It has been found to be a valid and comprehensive evaluation of a student’s clinical abilities, admittedly at massive expense to medical students.118 In the United States, osteopathic medical students take the OSCE-style Level 2 Performance Evaluation.

Although station number and total duration are not completely agreed upon, data indicate that the OSCE examination should be between 3 and 6 hours and 8 to 10 stations in length in order to obtain reliable \( r = 0.7 \) communication, history, and physical examination skills. A guideline was that at least seven cases are needed in any domain to achieve reliability. The testing period may be spread over several sessions making up an aggregate score in order to maintain validity. Many medical schools prepare their students for clinical practice with OSCE-style examinations throughout the year, which, taken together, are summative of a high-quality, multistation, valid OSCE. Checklists are typically the standard scoring tools; however, checklists alone may not be as reliable as a more comprehensive review by more experienced clinicians—particularly when assessing more advanced students and residents.119 All of the licensure examinations, discussed previously, meet the criteria for a well-designed OSCE based on number of stations and time duration.

Beginning in 2003, the ACGME mandated the use of OSCEs within residency programs. At the time, residents were wary of its adoption, particularly fearing its use as a tool for determining promotion. Residents’ perceptions of the examination, over time, did change to reflect an acceptance of its use for grading both technical and nontechnical skills.120-123

In the United States, the OSCE assesses technical and nontechnical skills in an accurate and valid fashion. The OSCE demonstrates a rapid progression of technical skills highly correlated to a postgraduate year, whereas clinical skills improve at a more moderate rate124 (Fig. 1-12).
By using simulated patients, patient-centered models, and intensive and immersive training, nontechnical skills including communication can improve interview techniques.\textsuperscript{137,138} Post communication skills training at the 12-month follow-up demonstrated that the training was effective, and with real clinical practice after the training communication skills had improved even more.\textsuperscript{139}

Lastly, there appears to be a positive feedback loop tying nontechnical leadership skills with self-perceived operating room prowess. Those surgeons who rate their own technical skills highly are also more likely to engage in positive leadership skills, including teaching in the operating room, handle difficult situations, and provide more clear instructions.\textsuperscript{140}

Simulation may be particularly critical for preventing technical skill decline in residents on dedicated research time or for attending surgeons whose research, clinic, or administrative duties decrease the amount of time they can spend in the operating room. Simulation represents the future of medicine and an excellent opportunity for research and development. Medicine, including surgery, has much ground to make up in regards to simulation training compared to other high-risk fields, such as the military, space, and aeronautics. Modern surgical leaders should recognize surgical simulation as critical to their organization’s success.

### Table 1-4

<table>
<thead>
<tr>
<th>Organization Name</th>
<th>Description</th>
<th>Main Skills, Conditions, or Qualities Evaluated</th>
<th>Evaluation of Technical Skills?</th>
<th>Evaluation of Nontechnical Skills?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS)</td>
<td>A public reporting initiative that measures patient perspectives on and satisfaction with hospital care based on qualities of healthcare that patients view as important.</td>
<td>Communication with nurses, communication with doctors, responsiveness of hospital staff, pain management, communication about medicines, discharge information, care transition</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Clinical and Group Consumer Assessment of Healthcare Providers and Systems (CGCAHPS)</td>
<td>A public reporting initiative that measures patient perspectives on and satisfaction with care provided in an office setting based on qualities of healthcare that patients view as important.</td>
<td>Access to care, provider communication, test results, office staff, overall provider rating</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Datix, Incident Reporting</td>
<td>A database of incidents that improves reliability of physicians by improving rates of reporting, promoting ownership of mistakes, and improving patient safety.</td>
<td>System issues, patient safety and quality issues, provider behavior, leadership style</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Patient Advocacy Reporting System (PARS)</td>
<td>A system that compiles patient complaints into a complaint index for each physician for comparison with other medical group members and to help identify high-malpractice-risk physicians who may benefit from peer intervention.</td>
<td>Unprofessional behavior deemed as disrespectful and rude</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Co-worker Observation Reporting System</td>
<td>A system in which physicians document coworker unprofessional conduct in order to provide nonjudgmental and timely feedback and to encourage self-reflection and change.</td>
<td>Unprofessional behavior deemed as disrespectful and unsafe</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

(Continued)
Multiple organizations have been created to evaluate both the technical and nontechnical skills of surgeons\textsuperscript{141}\textsuperscript{(Continued)}

<table>
<thead>
<tr>
<th>ORGANIZATION NAME</th>
<th>DESCRIPTION</th>
<th>MAIN SKILLS, CONDITIONS, OR QUALITIES EVALUATED</th>
<th>EVALUATION OF TECHNICAL SKILLS?</th>
<th>EVALUATION OF Nontechnical SKILLS?</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Board of Surgery (ABS) Maintenance of Certification (MOC) Program</td>
<td>A program that documents a surgeon’s ongoing commitment to professionalism, lifelong learning, and practice improvement through self-report.</td>
<td>Restrictions on medical license, restrictions on hospital privileges, continuing medical education, self-assessment of continuing medical education, cognitive expertise, ongoing participation in quality assessment program relevant to the surgeon’s practice</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Hospital Compare</td>
<td>A database that is part of the Centers for Medicare &amp; Medicaid Services (CMS) Hospital Quality Initiative and provides information on hospital performance and quality of care based on consumer perspectives so that patients can assess and compare hospitals.</td>
<td>Hospital Compare is based on data from HCAHPS and evaluates hospitals by the same guidelines as HCAHPS</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Federation of State Medical Boards (FSMB)</td>
<td>An organization representing all state medial and osteopathic boards in the United States that license physicians and sponsors the United States Medical Licensing Examination.</td>
<td>Medical knowledge, patient complaints, violations of the law</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Internet clinical scores</td>
<td>A database of direct patient opinions of physicians, provided through various sources, including Healthgrades.com, RateMDs. com, and Yelp.</td>
<td>Professionalism, communication, timeliness</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Hospital-Acquired Condition Reduction Program</td>
<td>A government program that provides incentives for hospitals to reduce the number of undesirable patient conditions resulting from their stay in the hospital and that could have been avoided by adjusting hospital reimbursement rates accordingly.</td>
<td>Foreign objects retained after surgery, air embolism, blood incompatibility, pressure ulcers, falls, poor glycemic control, catheter-associate infections, surgical site infections, deep vein thrombosis, pulmonary embolism, pneumothorax</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>American College of Surgeons National Surgical Quality Improvement Program (ACS NSQIP)</td>
<td>A program that collects information on and provides a risk-adjusted ranking of preventable surgical complication rates to encourage providers to improve care.</td>
<td>Surgical complications rates, surgical site infections, urinary tract infections, readmission rates, surgical outcomes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Centers for Medicare &amp; Medicaid Services Surgical Care Improvement Project (CMS SCIP)</td>
<td>A collaborative healthcare organization that collects data on surgical complication rates based on established guidelines.</td>
<td>Rates of infection, cardiac, venous thromboembolism, vascular, and respiratory, complications of surgery</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>
skills. Additionally, most have been tied to performance evaluations and even salary and discipline up to and including loss of licensure. To our knowledge, this is the first comprehensive listing of the various agencies that evaluate surgeon performance.141

MENTORING AND DEVELOPMENT

Mentoring

A formal leadership training program for surgical trainees should include mentoring. Mentoring is the active process by which an experienced, empathetic person guides another individual in the development and self-recognition of their own vision, learning, core competencies, and professional development. Halstead established the concept of a surgical mentor who directly provided the trainees with professional and technical guidance. Halstead’s concept went beyond a simple preceptorship by emphasizing clinical decision making based on scientific evidence. His goal was to develop surgeons who would go on to become outstanding leaders and innovators in the field. Although surgery has changed dramatically since Halstead’s era, mentorship remains crucial in surgical training. In addition to teaching technical skills, clinical judgment, and scientific inquiry, modern-day mentors must also model effective communication, empathy, humanism, and the prioritization of competing professional and personal activities.

The mentor must also be an experienced and trusted advisor committed to the success of the mentee. A greater level of trust and commitment distinguishes the mentor from the teacher. More than a teacher, a mentor is a coach. The goal of a teacher is to pass on a defined level of knowledge for each stage of a student’s education. The underlying premise is a limited level of advancement for the student. The coach, on the other hand, has the sole purpose to make his or her student the best at their game, with an unlimited level of advancement. Modern mentorship implies a partnership between the mentor and the mentee. Surgical residency program chairs and program directors must recruit and develop faculty “coaches” to mentor residents to optimize their potential. Emeritus Chair of the University of California, Los Angeles Head and Neck Surgery, Dr. Paul Ward, said it best: “We strive to produce graduates of our residency program who are among those who change the way we think and practice.” 142 Having more than 25 former residents become chairs of academic head and neck surgical programs, Dr. Ward embodied the role as a surgeon’s coach. The responsibilities of an effective mentor are summarized by Barondess: “Mentoring, to be effective, requires of the mentor empathy, maturity, self-confidence, resourcefulness, and willingness to commit time and energy to another. The mentor must be able to offer guidance for a new and evolving professional life, to stimulate and challenge, to encourage self-realization, to foster growth, and to make more comprehensible the landscape in which the protégé stands.”143

One of the major goals of mentors is to assess the aptitudes and abilities of mentees with regard to the appropriateness of their vision for their surgical career. Proper selection of the appropriate mentor can bring to the mentee much needed wisdom, guidance, and resources and can expand the scope of his or her vision. In addition, the mentor can refine the leadership skills taught to mentees in formal training programs. Highly successful surgeons most often have had excellent surgical mentors. It is impressive to note that more than 50% of United States’ Nobel laureates have served under other Nobel laureates in the capacity of student, postdoctoral fellow, or junior collaborator.144 In academic medicine, evidence-based studies have shown benefits to the mentees that include enhanced research productivity, higher likelihood of obtaining research grants, and greater success in obtaining desired positions in practice or at academic institutions.145 Mentoring provides benefits to the mentors themselves, including refinement of their own personal leadership skills and a strong sense of satisfaction and accomplishment.

Mentorship is essential to accomplish the successful development of surgical trainees and to help cultivate their vision. Therefore, formal leadership training programs that have a goal of training the future leaders in surgery should include mentoring.

Modeling Leadership for Medical Students and the “Hidden Curriculum”

Medical students enter school with great empathy, excitement, optimism, and an idealistic vision. They have self-selected to enter a profession of healing and achieved entry into a highly coveted graduate training program with centuries of tradition. Yet, these medical students are naive to the actual practice of medicine and its professional norms. Along the way to becoming a doctor, many medical students lose some of the optimism, empathy, and excitement, particularly during their first and third years of school. Some students come to see the patient-physician relationship as an afterthought to providing care.145,146 Through the “hidden curriculum,” formal leadership training, and modeling of professional behavior, surgical residents, and attendings can help medical students to realize their vision of becoming empathic physicians.

Traditionally, medical schools and professors have unknowingly relied on a hidden curriculum to mold these idealistic students into capable professionals. The hidden curriculum is the informal social norms learned by students implicitly, based on their observations of resident and attending behavior. The hidden curriculum has always been present in education, for better or worse, and may be unmasked and studied, but cannot be eliminated. Medical students actively engage in seeking out mentors, and naturally and subconsciously look to their mentors for cues on how to conduct themselves as physicians, the same way in which a child learns how to behave from a parent or older sibling. Whether or not the witnessed behavior is a positive example of professionalism, the student will begin to perceive that behavior as normal and acceptable. For better or worse, the professional norms of medicine (the Hippocratic oath, respect to patients and colleagues, ethical conduct, personal accountability, empathy, and altruism) are modeled in every personal encounter. It is imperative that all resident and attending surgeons understand that the medical students are observing them closely. When resident and attending surgeons model professional behavior, the hidden curriculum becomes a useful tool for professional development.147,150 This consistent modeling of professional behavior is one necessary component of leadership.

During their clinical years, medical students experience both an exponential growth in knowledge and a measurable decline in empathy towards their patients. Initially, medical students are filled with excitement and wonder during their first patient encounters. The rapid pace of clinical work, acquisition of knowledge, and intense experiences create stress for the student, both positively and negatively. Scrubbing into the operating room, witnessing the passing of a patient, helping deliver a baby, and studying for boards are impactful milestones that each student experiences in a matter of months. Due to the
challenges of their work, students naturally have doubts about their own career choices and abilities, even as they experience growth and success. However, as students gain knowledge and abilities, they also come to see commonly encountered clinical problems as routine work. As familiarity and comfort with clinical problems increases, the excitement and wonder experienced by the student decreases. It is during this time that a decline in student empathy occurs, typically in their third year of medical school. In medicine, even routine clinical work still requires extraordinary attention to detail, and compassionate care must be delivered to every patient, every time. This attention to detail and compassionate delivery of care are the hallmark of the true professional. It is important that surgical residents and attendings always model positive behavior.

Previously, medical schools instructed students in anatomy, physiology, pathology, and clinical medicine, but left the acquisition of professionalism to the informal hidden curriculum. The Carnegie Report, published in 2010 at the 100-year anniversary of the Flexner Report, called for medical education to promote “the progressive formation of the physician’s professional identity.” To this end, many medical schools nationwide emphasize early professional education and an integrated curriculum. The Liaison Committee on Medical Education (LCME) sets standards for administrative and faculty leadership that manage the curricular model and educational affairs of students; however, formal leadership education is not explicitly required at this time. However, career exploration, mentoring, and advising are instrumental responsibilities of each medical school and a requirement of the LCME. Establishing a leadership program that is perpetual and coexists within an integrated curriculum will support this endeavor. A longitudinal leadership program beginning at the onset of medical school can establish a pattern of ethical behavior, professionalism, balance, and professional identity.

Tools to Measure Leadership Outcomes in Healthcare

There is evidence that leadership training improves healthcare quality. The ACGME, via its core competencies, has recognized technical skills, surgical judgement, and nontechnical skills as qualities essential to develop in residents. The objective measurement of nontechnical skills is difficult. Table 1-5 includes a list of methods for assessing nontechnical skills currently in use by some residency programs. The Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey is in the early phases of being applied to individual physicians, but it has been applied to hospitals as a whole for several years.

<table>
<thead>
<tr>
<th>METHOD OF LEADERSHIP MEASUREMENT</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multifactor Leadership Questionnaire (MLQ)</td>
<td>The MLQ is a questionnaire based on the differences between transformation and transactional approaches of leadership. It identifies leadership qualities through the rater’s beliefs about effective leadership.</td>
</tr>
<tr>
<td>NEO Five-Factor Personality Inventory (NEO)</td>
<td>NEO explores different facets of five different personality traits—neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness—through a questionnaire.</td>
</tr>
<tr>
<td>Surgeon’s Leadership Inventory (SLI)</td>
<td>The SLI is a questionnaire based on literature on leadership in surgery and surgeon’s leadership behaviors observed in the operating room. It includes eight elements of surgeon’s leadership in the operating room, which are maintaining standards, managing resources, making decisions, directing, training, supporting others, communicating, and coping with pressure.</td>
</tr>
<tr>
<td>Patient feedback</td>
<td>Patient complaints are inversely related to leadership effectiveness and can thus be used as opportunities to improve and as a measure of leadership.</td>
</tr>
<tr>
<td>Objective Structured Clinical Examination (OSCE)</td>
<td>The OSCE can be administered in a controlled environment with attending feedback on various aspects of leadership tackled in the practice cases. Videotaped sessions provide further opportunities for improvement as residents will be able to later observe their own behaviors and reflect on ways to improve their approach to the case presented.</td>
</tr>
<tr>
<td>Consumer Assessment of Healthcare Providers and Systems (CAHPS)</td>
<td>CAHPS surveys are based on aspects of healthcare that matter most to patients, such as physician communication. The results are made public and can be used to shed light on areas of leadership physicians can improve on to work towards a patient-centered approach to care.</td>
</tr>
</tbody>
</table>

Leadership can be evaluated through instruments such as the Multifactor Leadership Questionnaire, the NEO Five–Factor Inventory, and the Surgeon’s Leadership Inventory. The Multifactor Leadership Questionnaire (MLQ) analyzes leadership aptitude as either a transactional or a transformational style.\cite{157} Leadership based on transaction focuses on completing and rewarding the tasks, whereas leadership based on transformation focuses more on the motivation for completing the tasks and emphasizes a positive and encouraging working environment for the team.\cite{158,159} In a study applying the questionnaire to five surgeons in a single hospital, surgeons who scored higher on the transformational section were more focused on promoting an open environment for all the attendings, residents, nurses and other staff in the operating room. This transformational style correlated with greater communication. These findings are important in showing that lack of communication is often a leading factor in surgical errors.

The use of an MLQ in 2008 studying surgical residents showed a significant association between transformational leadership and overall perceived team effectiveness and resident satisfaction.\cite{158,159} The questionnaire also found that the residents, as leaders, placed less value on the individual needs of their colleagues, possibly reflecting a high sense of independence and frequent changes in teams due to rotations among services. This finding helped identify an area of leadership training on which the program can focus to help further develop a more supportive team atmosphere amongst the residents. In 2011, a study administered the NEO Five–Factor Personality Inventory (NEO) to a group of surgical residents. NEO, which assesses personality on five broad strokes, including neuroticism, openness, agreeableness, extroversion, and conscientiousness, found that the surgeons scored above the national average on most of the factors tested but below average on agreeableness. This is a measure of altruism and tolerance, among other related factors. This result corresponded with the MLQ administered to the same group of residents and therefore highlighted areas of leadership that required modification.\cite{158,159}

The Surgeon’s Leadership Inventory (SLI) is a helpful guide for residency programs.\cite{160} The SLI grades surgeons on eight different elements of leadership, as listed in Table 1-6. As with the MLQ and NEO questionnaires, the SLI can be used to assess the growth of leadership ability in surgery residents. Table 6 provides a list and description of the different elements assessed by the SLI.\cite{141}

### Table 1-6

<table>
<thead>
<tr>
<th>ELEMENT</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maintaining standards</td>
<td>Practicing safe and quality patient care by following established protocols and asking for help when needed</td>
</tr>
<tr>
<td>Making decisions</td>
<td>Making informed judgments and communicating decisions with relevant personnel</td>
</tr>
<tr>
<td>Managing resources</td>
<td>Appropriately assigning resources and tasks to team members</td>
</tr>
<tr>
<td>Directing</td>
<td>Clearly communicating expectations and instructions and demonstrating confidence in leadership ability</td>
</tr>
<tr>
<td>Training</td>
<td>Educating and training team members when the opportunity arises</td>
</tr>
<tr>
<td>Supporting others</td>
<td>Offering assistance where appropriate and encouraging open communication</td>
</tr>
<tr>
<td>Communicating</td>
<td>Sharing information in a timely manner and encouraging input from others</td>
</tr>
<tr>
<td>Coping with pressure</td>
<td>Showing flexibility when required to meet goals</td>
</tr>
</tbody>
</table>

Data correlating patient complaints in a large number of hospitals show that improved leadership is associated with better hospital climate, improved performance, and a lower number of complaints.\cite{161} Reproduced with permission from Jacobs LA: Practical Ethics for the Surgeon. Philadelphia, PA: Wolters Kluwer; 2018.

Leadership training for the prospective surgeon

Prospective surgeons such as medical students and premedical students may have no better source for developing the personal attributes necessary for a successful career than current surgical attending surgeons and current residents. When surveyed, these doctors emphasized accountability, resilience, and high personal standards for oneself as critical tools. Prospective surgeons are advised to pursue perfectionism and be self-critical, cautioning against taking these traits to far towards neurotic behavior. Critical leadership skills of teamwork and learning to take initiative are mandatory in modern medicine and must be learned early. Innovation is highly desirable.\cite{162}

Residents, on the other hand, are closer to becoming independent. To some extent, they have already been selected for their leadership, innovation, and resiliency through the process of the match. During training their progression from novice to expert is necessarily rapid. A graded tool for all procedure based specialties including surgery – OpTrust – has been recently validated to facilitate the resident’s transition to leadership across five domains including questioning, planning, instruction, problem solving, and leadership.\cite{163}

As emphasized throughout this chapter, the concept of training leadership skills early applies particularly to junior faculty and residents. The resident-surgeon-manager conference is one model for integrating department members of various experience levels into a results-based leadership conference. In this conference, various stakeholders including attorneys, persons with business experience, and risk management experts are brought in as guest participants. Exercises were immersive and included case-based discussions, role-playing, simulation, and interactive lecture. Topics included teamwork, learning negotiating techniques, time management, risk management, balance, giving feedback, and creating immediate, goal-oriented action plans.\cite{164}

### EARLY CAREER DEVELOPMENT AND ESTABLISHING ONESELF

A variety of methods have been proposed for the professional development of new attending surgeons. “Speed Mentoring”—10-minute pairings of senior and junior surgeons answering preset questions—have been studied at national conferences with promising results. These sessions could be spread out over several days and integrated into a busy surgeon’s schedule.\cite{164}
A study of department chairs and award-winning surgeon-scientists identified perseverance and team leadership skills as critical factors for development in the young attending surgeon. Chairs advocated protected time for research, financial support, and mentorship as departmental level support that the surgeon scientist should actively seek out in their first position. The surgeon-scientist compared to the pure clinician faces a different set of challenges, particularly the financial challenge of funding research and clinical duties competing for time and attention with research interests.\(^{165}\)

One study addressed surgeon behavior in the operating room to assess the leadership style most associated with strong leadership. Based off of this study, surgeons who are trained to collaborate, consult others appropriately, be polite (simple “please” and “thank you”), and create a safe space for their operating room staff to voice concerns will demonstrate good leadership. However, surgeons who demonstrate nonconstructive criticism, destructive humor, steer conversation away from the current case, and express frustration will be perceived as demonstrating poor leadership. Under this system, surgeon behavior can be categorized—conductor, elucidator, delegator, engagement facilitator, tone setter, being human, and safe space maker—in order to provide individual feedback for professional development.\(^{166}\)

**SENIOR FACULTY DEVELOPMENT: TRANSITIONING TO DEPARTMENTAL LEADERSHIP AND LEGACY BUILDING**

The presence of experienced, senior academic surgeons within a department represents an opportunity. The formal development of a plan for late career transitioning through departmental leadership roles all the way to emeritus status naturally initiates a constructive process when thought out years in advance. The plan should be agreeable to the senior faculty member in question as well as departmental leadership and hospital stakeholders. Once in place, the senior academic surgeon and department will both thrive thanks to a shared vision, mutual understanding, and clear goals and transition points. Departmental leadership can use the transition plan to look ahead at the future of their department years down the line.\(^{167}\)

Recognition of senior academic surgeons with departmental leadership, promotions, and emeritus status is a privilege earned by the academician over a lifetime of work; however, for the department it represents an opportunity to shape the values and culture of the faculty body as a whole. The continued visibility, model, and influence of such leaders will have a trickle-down effect on the rest of the department. Surgical leaders are part of a large and extraordinary network facilitated by mentorship and decades of professional collaboration. Exceptional senior academic surgeons may often experience the “multiplier-effect” whereby one excellent leader trains several, who go on to train several more until the culture of surgery nationwide is influenced.\(^{168}\)

Although there are no mandatory ages for which surgeons must retire as in other professions, such as airline pilots, the issue of aging and when to cease practice has been controversial. There are some, however few, reports of physicians practicing after the decline of their skill and becoming dangerous. As a whole, the profession has been unable to prevent this. Nationwide, from 1975 to 2015, the number of physicians practicing after age 65 has increased by 374%. Some hospitals and healthcare organizations have implemented mandatory cognitive and physical evaluations as a condition of continued practice. In the absence of more robust professional initiatives our field may see legislative oversight in the future.\(^{169}\) The authors believe that a formally planned transition emphasizing the values of leadership and legacy-building offers a more palatable alternative.

**CONCLUSION**

Although there are several definitions of leadership and a variety of leadership styles, all share the common goal of improving patient care in the modern era. All forms of leadership require a vision and willingness—the willingness to assume the responsibility to lead, continue learning, practice effective communication styles, and resolve conflict. Effective leadership can change surgical departments and improve patient care through innovation. A growing body of evidence suggests the mastery of leadership requires practice through intentional curricula and reinforcement through mentorship.

Surgical leadership is bred through its training programs. Thus, innovation in surgical training programs is needed to enhance the development of leadership skills of surgical trainees, to prepare them for practice in modern healthcare systems, and to optimize patient care, as well as compliance with requirements set forth by regulatory institutions governing surgery and surgical education. A growing body of literature supports the value of effective leadership in improving patient care, productivity, and the work environment while it validates the ability to measure the impact of leadership training. Therefore, it is of paramount importance to teach modern leadership principles and skills to surgical trainees in order to create a new generation of surgeon leaders who will shape the modern era of surgery in the context of rapidly evolving science, technology, and systems of healthcare delivery.

**REFERENCES**

Entries highlighted in bright blue are key references.


50. Ochoa LN, Brunicardi FC, Brandt ML, Awad SS. Evaluation of resident training activities using a novel time-motion study prior to the implementation of the 80-hour work week. Chicago, IL: American College of Surgeons Annual Conference, 2006.

23
LEADERSHIP IN SURGERY


57. Lencioni PM. Make your values mean something. Harv Bus Rev. 2002;80(7):113-117.


This page intentionally left blank
OVERVIEW: INJURY-ASSOCIATED SYSTEMIC INFLAMMATORY RESPONSE

The inflammatory response to injury occurs as a consequence of the local or systemic release of “damage-associated” molecules to mobilize the necessary resources required for the restoration of homeostasis. Minor host insults result in a localized inflammatory response that is transient and, in most cases, beneficial. Major host insults follow a different trajectory. A subset of these patients will die within 24 hours of hospital admission, succumbing to overwhelming tissue injury and immediate organ damage. With advances in prehospital care and improved trauma management, these numbers have diminished. A second subgroup of patients who suffer a major host insult succumb to secondary organ damage remote from the injury site and die later (weeks) in their hospital course. They form an increasing percentage of the in hospital trauma-related deaths. A dysregulated, overwhelming systemic inflammatory response to the injury/hemorrhage and associated ischemia/reperfusion events has been implicated as the cause of multiple organ failure in these patients. Moreover, it has been linked to immune suppression that increases the risk of infectious complications and poor outcome. Finally, a third subgroup, characterized by extended length of stay in the ICU, complicated postdischarge courses, and failure to regain/recover to their preinjury status, has been described and also linked to persistent inflammation and suppressed host protective immunity. The term persistent inflammation, immunosuppression, and catabolism syndrome (PICS) has
been applied to this group. Recent data suggest that severely injured patients who are destined to die from their injuries, whether late in their hospital course or after discharge, differ from survivors only in the degree and duration of their dysregulated acute inflammatory response.

As trauma is the leading cause of mortality and morbidity for individuals under age 45, understanding the complex pathways that regulate the local and systemic inflammatory response following severe traumatic injury is necessary to develop appropriate and targeted therapeutic strategies that will improve outcomes for these patients.

In this chapter, we will review what is known about the soluble and cellular effectors of the injury-induced inflammatory response, how the signals are sensed, transduced, and modulated, and how their dysregulation is associated with alterations in the immune system. We will also discuss how these events are monitored regulated by the central nervous system. Finally, we will review how injury reprograms cellular metabolism, in an attempt to mobilize energy and structural stores to meet the challenge of restoring homeostasis.

**THE DETECTION OF CELLULAR INJURY**

**The Detection of Injury is Mediated by Members of the Damage-Associated Molecular Pattern Family**

Traumatic injury activates the innate immune system to produce a systemic inflammatory response (SIR) in an attempt to limit damage and to restore homeostasis. It includes two general responses: (a) an acute proinflammatory response resulting from innate immune system recognition of ligands, and (b) an anti-inflammatory response that may serve to modulate the proinflammatory phase and direct a return to homeostasis (Fig. 2-1). This is accompanied by a suppression of adaptive immunity. Rather than occurring sequentially, recent data indicate that all three responses are simultaneously and rapidly induced following severe traumatic injury.

The degree of the systemic inflammatory response following trauma is proportional to injury severity and is an independent predictor of subsequent organ dysfunction and resultant mortality. Recent work has provided insight into the mechanisms by which immune activation in this setting is triggered. The clinical features of the injury-mediated systemic inflammatory response, characterized by increased body temperature, heart rate, respirations, and white blood cell count, are similar to those observed with infection (Table 2-1). However, it is widely accepted that systemic inflammation following trauma is sterile, resulting from endogenous molecules that are produced as a consequence of tissue damage or cellular stress. Termed *damage-associated molecular patterns (DAMPs)* or *alarmins*, DAMPs interact with specific cell receptors that are located both on the cell surface and intracellularly.

Trauma DAMPs are structurally diverse endogenous molecules that are immunologically active. Table 2-2 includes a partial list of DAMPs that are released either passively from necrotic/damaged cells or actively from physiologically “stressed” cells by upregulation or overexpression. Once they are outside the cell, DAMPs promote the activation of innate immune cells, as well as the recruitment and activation of antigen-presenting cells, which are engaged in host defense. The best-characterized DAMP with significant preclinical evidence for posttrauma release, as well as a direct link to the systemic inflammatory response, is high-mobility group protein B1 (HMGB1). Additional evidence for other important DAMP molecules that participate in postinjury inflammation is also presented.

**High-Mobility Group Protein B1.** The best-characterized DAMP in the context of the injury-associated inflammatory response is high-mobility group B1 (HMGB1) protein. HMGB1 is highly conserved across species. It is a constitutively expressed, nonhistone chromosomal protein that participates in a variety of nuclear events, including DNA repair and transcription. Inflammatory signaling can redirect HMGB1 to the cytosol in both monocytes and macrophages, as a result of posttranslational modification. HMGB1 is released passively from damaged or necrotic cells and is detected rapidly in the circulation within 30 minutes post injury. It can also be actively secreted from immune-competent cells stimulated by bacterial-derived lipoproteins (e.g., endotoxin) or by inflammatory cytokines (e.g., tumor necrosis factor). For example, macrophages release HMGB1 following the activation of the inflammasomes.
Once outside the cell, HMGB1 has been shown to signal via the Toll-like receptors (TLR2, TLR4, TLR9), the receptor for advanced glycosylation end products (RAGE), CD24, and others. The activation of TLRs by HMGB1 occurs mainly in myeloid cells, whereas RAGE is thought to be the receptor target for HMGB1 in endothelial cells.

The diverse proinflammatory biological responses that result from HMGB1 signaling include: (a) the release of cytokines and chemokines from macrophage/monocytes and dendritic cells; (b) neutrophil activation and chemotaxis; (c) alterations in epithelial barrier function, including increased permeability; and (d) increased procoagulant activity on platelet surfaces; among others.10 In addition, HMGB1 binding to TLR4 triggers the proinflammatory cytokine release that mediates “sickness behavior.”11

The biologic function of HMGB1 is regulated by its redox state. For example, a thiol at C106 is required for HMGB1 to promote macrophage TNF release, while a disulfide bond between C23 and C45 confers proinflammatory properties. With all three cysteines in the thiol (reduced) state, HMGB1 loses its DAMP function, but gains the capacity to serve as a chemotactic mediator. Importantly, shifts between the redox states have been demonstrated and indicate that redox state dynamics are important regulators of HMGB1.12

### Table 2-1
**Clinical spectrum of infection and systemic inflammatory response syndrome (SIRS)**

<table>
<thead>
<tr>
<th>TERM</th>
<th>DEFINITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infection</td>
<td>Identifiable source of microbial insult</td>
</tr>
<tr>
<td>SIRS</td>
<td>Two or more of following criteria are met: Temperature ≥38°C (100.4°F) or ≤36°C (96.8°F) or Heart rate ≥90 beats per minute Respiratory rate ≥20 breaths per minute or Paco₂ ≤ 32 mmHg or mechanical ventilation Abnormal white blood cell count (≥12,000/µL or ≤4000/µL or ≥10% immature band forms)</td>
</tr>
<tr>
<td>Sepsis</td>
<td>Identifiable source of infection + SIRS</td>
</tr>
<tr>
<td>Severe sepsis</td>
<td>Sepsis + organ dysfunction</td>
</tr>
<tr>
<td>Septic shock</td>
<td>Sepsis + cardiovascular collapse (requiring vasopressor support)</td>
</tr>
</tbody>
</table>

Paco₂ = partial pressure of arterial carbon dioxide.

### Table 2-2
**Damage-associated molecular patterns (DAMPs) and their receptors**

<table>
<thead>
<tr>
<th>DAMP MOLECULE</th>
<th>PUTATIVE RECEPTOR(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMGB-1</td>
<td>TLRs (2,4,9), RAGE</td>
</tr>
<tr>
<td>Heat shock proteins</td>
<td>TLR2, TLR4, CD40, CD14, Siglec</td>
</tr>
<tr>
<td>S100 protein</td>
<td>TLR4, RAGE</td>
</tr>
<tr>
<td>Mitochondrial DNA</td>
<td>TLR9</td>
</tr>
<tr>
<td>Hyaluran</td>
<td>TLR2, TLR4, CD44</td>
</tr>
<tr>
<td>Biglycan</td>
<td>TLR2 and TLR4</td>
</tr>
<tr>
<td>Formyl peptides (mitochondrial)</td>
<td>Formyl peptide receptor 1</td>
</tr>
<tr>
<td>IL-1α</td>
<td>IL-1 receptor</td>
</tr>
</tbody>
</table>
HMGB1 levels in human subjects following injury correlate with the Injury Severity Score and complement activation, as well as with increases in circulating inflammatory mediators such as tumor necrosis factor. Exogenous administration of HMGB1 to normal animals produces fever, weight loss, epithelial barrier dysfunction, and, possibly, death. Further supporting the HMGB1 role in sterile inflammation, traumatic brain injury (TBI) induced by a cortical injury model has been shown to result in acute lung injury with increased alveolar hemorrhage, neutrophil infiltration, and poor oxygenation. This acute lung injury (ALI) was accompanied by a doubling in serum HMGB1 concentrations along with evidence that necrotic brain cells were a source of HMGB1 following TBI. More recently, in an animal model of hemorrhagic shock, HMGB1 release from intestinal epithelium was linked to acute lung injury. Finally, increased plasma levels of HMGB1 have been shown to correlate with immune suppression and increased infection risk in patients undergoing major surgical procedures. The identification of the receptor for advanced glycation end products as the receptor for HMGB1 in this setting has identified new therapeutic strategy to ameliorate ALI following TBI.

A Role for Mitochondrial DAMPs in the Injury-Mediated Inflammatory Response. Mitochondrial proteins and/or DNA can act as DAMPs by triggering an inflammatory response to cellular necrosis and stress. Specifically, mitochondrial DNA (mtDNA) released from damaged or dysfunctional mitochondria leads both to inflammasome activation and activation of the stimulator of interferon gene pathway (STING). Cell-free mtDNA (cf-mtDNA) has been shown to be thousands of times higher in trauma patients when compared to normal volunteers. In addition, direct injection of mitochondria lysates in an animal model causes remote organ damage, including liver, and lung inflammation. These data suggest that with cellular stress or tissue injury, cf-mtDNA released from damaged/stressed mitochondria contribute to the sterile inflammatory response in injured patients. From an evolutionary perspective, given that eukaryotic mitochondria derive from bacterial origin, it would make sense that they retain bacterial features capable of eliciting a strong response that is typically associated with a pathogen trigger. In addition, the mitochondrial transcription factor A (TFAM), a highly abundant mitochondrial protein, is functionally and structurally homologous to HMGB1. It has also been shown to be released in high amounts from damaged cells where it acts in conjunction with mtDNA to activate TLR9 signaling.

Following trauma, cf-mtDNA levels appear to be higher in nonsurvivors when compared to survivors and correlate with the development of both SIRS and sepsis post injury. Cf-mtDNA has also been linked both ex vivo and in vivo to the formation of neutrophil extracellular traps, which are also associated with sterile inflammation and are a possible cause of secondary tissue injury. Reducing cf-mtDNA, perhaps by targeting enzymes capable of digesting circulating mtDNA is an attractive therapeutic option to prevent development of inflammatory complications of trauma.

Heat-Shock Proteins as DAMPs. Heat shock proteins (HSPs) are a large and diverse family of intracellular proteins that are expressed during times of inflammation and oxidative stress or following tissue injury. Very highly conserved across species, HSPs function as molecular chaperones to monitor and maintain appropriate protein folding. They accomplish this task through the promotion of protein refolding, the targeting of misfolded proteins for degradation, or the sequestering of partially folded proteins for movement to appropriate membrane compartments. HSPs are also capable of binding foreign proteins and thereby function as intracellular chaperones for ligands such as bacterial DNA and endotoxin.

HSPs are presumed to protect cells from the effects of traumatic stress and, when released by damaged cells, alert the immune system of the tissue damage by activating both innate and acquired immunity. HSPs are also released from intact cells via a nonclassical secretory pathway, both via “secretory lysosomes” as well as the exosomal pathway. For example, HSP70-containing exosomes have been implicated in postshock inflammation. Once outside the cell, free HSPs can bind to pattern-recognition receptors (PRR) as well as other cell surface receptors to modulate the inflammatory response. Recently, the role of free HSP-mediated proinflammatory properties via TLR2 and TLR4 has been questioned, as it has been suggested that the presence of contaminating endotoxin in bacterially-produced HSP preparations may explain at least some of these inflammatory effect results. However, the additional evidence suggests that the immunostimulatory properties may be dependent on how HSPs arrive outside the cell. In the context of massive cell damage or large exosome release, HSPs may serve as proinflammatory DAMPs. In contrast, HSPs released by active secretion may exert anti-inflammatory immune dampening signals (Table 2-3). New receptors for HSP have been identified that are members of the sialic acid-binding immunoglobulin-like lectins (siglecs), which may explain these effects. Two members of the family, Siglec-5 and Siglec-14, with similar binding sites for HSP70, exhibit opposite intracellular events in response to HSP binding, being either pro-(Siglec-14) or anti-(Siglec-5) inflammatory.

From a clinical perspective, extracellular HSPs have been demonstrated to be elevated almost immediately post injury in polytraumatized patients (up to 10 times normal) with the degree of elevation being correlated with the severity of illness. Moreover, in the setting of polytrauma, plasma HSP70 levels have been shown to correlate inversely with HLA-DRA expression, a marker of immunosuppression.

Extracellular Matrix Molecules Act as DAMPs. Recent work has explored the role of extracellular matrix (ECM) proteins in the TLR-mediated inflammatory response that follows tissue injury. These molecules, which are sequestered under normal conditions, can be released in a soluble form with proteolytic digestion of the ECM. Proteoglycans, glycosaminoglycans, and glycoproteins such as fibronectin have all been implicated as key players in the DAMP/TLR interaction. Proteoglycans, in particular, have also been shown to activate the intracellular inflammasomes that trigger sterile inflammation. These molecules, which consist of a protein core with one or more covalently attached glycosaminoglycan chains, can be membrane-bound, secreted, or proteolytically cleaved and shed from the cell surface.

Biglycan is one of the first proteoglycans to be described as a TLR ligand. It consists of a protein core containing leucine-rich repeat regions, with two glycosaminoglycan (GAG) side-chains (chondroitin sulfate or dermatan sulfate). While biglycan typically exists in a matrix bound form, with tissue injury it is released from the ECM in a soluble form where it interacts with TLR2 or TLR4 to generate an immediate inflammatory response. Various proinflammatory cytokines and chemokines including tumor necrosis factor (TNF)-α and interleukin (IL)-1β are...
downstream effector molecules of biglycan/TLR2/4 signaling. Among these, the mechanism of biglycan-mediated autonomic synthesis and secretion of mature IL-1β is unique. Usually, release of mature IL-1β from the cell requires two signals: one that is needed to initiate synthesis (TLR2/4-mediated), and the other to process pro-IL-1β to its mature form (inflammasome-mediated). How is it possible for biglycan to provide both signals? Current evidence indicates that when soluble biglycan binds to the TLR, it simultaneously serves as a ligand for a purinergic receptor, which facilitates the inflammasome activation required for IL-1β processing. These data support the idea that DAMP-mediated signals can initiate a robust inflammatory response.  

**S100 Proteins as DAMPs.** S100 proteins are a group of calcium-binding proteins that participate in the regulation of intracellular calcium. There are at least 25 members identified to date, with diverse functions that are cell-type dependent. While regulation and management of calcium storage is a primary function of S100 proteins, additional specialized roles include cytoskeletal organization, protein trafficking and transcriptional regulation. They are loosely grouped according to their functional capability: those that work exclusively inside the cell, outside the cell, or in both locations.  

<table>
<thead>
<tr>
<th>CELL LOCATION</th>
<th>RECOGNIZED AS DAMP</th>
<th>IMMUNOMODULATORY FUNCTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSP90</td>
<td>Cytoplasm, endoplasmic reticulum Can function both inside and outside the cell</td>
<td>May act as DAMP chaperone to activate innate immune response</td>
</tr>
<tr>
<td>HSP70</td>
<td>Can function both inside and outside the cell Endoplasmic reticulum homolog is BiP</td>
<td>Exogenous HSP70 elicits cellular calcium flux, NF-κB activation, cytokine production</td>
</tr>
<tr>
<td>HSP60</td>
<td>Mitochondria</td>
<td>Exogenous HSP60 inhibits NF-κB activation</td>
</tr>
</tbody>
</table>

BiP = binding immunoglobulin protein; DAMP = damage-associated molecular pattern; IKK = IκB kinase; NF-κB, nuclear factor-κB; TLR = Toll-like receptor  

**Heme as DAMP.** Heme is the oxygen-binding moiety found in hemoglobin and other hemoproteins in the muscle and mitochondria. It is a very highly conserved molecule composed of a tetrapyrrole ring surrounding a single iron. When red blood cells are damaged, hemoglobin is released, where it is bound by plasma proteins such as haptoglobin. In turn, the hemoglobin-haptoglobin complexes are scavenged by the reticuloendothelial system in the liver and spleen to form bilirubin as the ultimate end product. When the amount of free hemoglobin exceeds the binding capacity of haptoglobin and other specialized binding proteins, it is loosely bound to other plasma proteins where it can be readily oxidized. Ultimately, this can result in the release of the prosthetic heme group from hemoglobin, generating labile heme, which is a pro-oxidant.  

In vitro experiments demonstrate that labile heme induces cell activation, via both TLR4-dependent processes and the inflammasome, resulting in cytokine release. Moreover,
heme-induced neutrophils activation leads to extracellular traps (NETs) release through a mechanism dependent on reactive oxygen species. However, unlike the other DAMPs discussed, labile heme can also have direct cytotoxic effects on cells by a direct interaction with membrane phospholipids and the catalysis of membrane lipid peroxidation, leading to programmed cell death. In macrophages, labile heme can induce necroptosis, rather than apoptosis.

**DAMPs Are Ligands for Pattern Recognition Receptors**

The inflammatory response that occurs following traumatic injury is similar to that observed with pathogen exposure. Not surprisingly, surface and cytoplasmic receptors that mediate the innate immune response to microbial infection have also been implicated in the activation of sterile inflammation. In support of this idea, genes have been identified that are dysregulated acutely both in response to a microbial ligand administered to human volunteers and in response to traumatic injury in a large patient population. The classes of receptors that are important for sensing damaged cells and cell debris are part of the larger group of germ-line encoded pattern recognition receptors (PRRs). The best described ligands for these receptors are microbial components, the pathogen-associated molecular patterns (PAMPs). The PRRs of the innate immune system are varied and include Toll-like receptors (TLRs), calcium-dependent (C-type) lectin receptors (CLRs), the nucleotide-binding domain, leucine-rich repeat–containing (NBD-LRR) proteins (NLRs); also nucleotide-binding and oligomerization domain [NOD]-like receptors), receptors for advanced glycation end-products (RAGE), and retinoic acid–inducible gene (RIG)-I-like receptors (RLRs). Following receptor ligation, intracellular signaling modulates transcriptional and posttranslational events necessary for host defense by coordinating the synthesis and release of cytokines and chemokines to either initiate or suppress the inflammatory response. The best described of these receptors, the TLRs, NLRs, CLRs and RAGE, are discussed in the following section.

**Toll-Like Receptors.** The Toll-like receptors are evolutionarily conserved type 1 transmembrane proteins that are the best-characterized PRRs in mammalian cells. They were first identified in *Drosophila*, where a mutation in the *Toll* gene led to its identification as a key component in their immune defense against fungal infection. The first human TLR, TLR4, was identified shortly thereafter. Now, more than 10 human TLR family members have been identified, with distinct ligands that include lipid, carbohydrate, peptide, and nucleic-acid components of various pathogens. TLRs are expressed by both immune and nonimmune cells. At first, the expression of TLR was thought to be isolated to professional antigen-presenting cells such as dendritic cells and macrophages. However, mRNA for TLR family members have been detected in most cells of myeloid lineage, as well as NK cells. In addition, activation of T cells increases their TLR expression and induces their survival and clonal expansion. Direct engagement of TLR in Treg cells promotes their expansion and reprograms them to differentiate into T helper cells, which in turn provides help to effector cells. In addition, B cells express a distinct subset of the TLR family that determines their ability to respond to DAMPs; however, the significance of restricted TLR expression in these cells is not yet clear.

All TLRs consist of a ligand-binding domain, characterized by multiple leucine-rich repeats (LRRs), and a carboxy-terminal, intracellular Toll/interleukin (IL) 1 receptor (TIR) domain. The LRR domains recognize bacterial and viral PAMPs in the extracellular environment (TLR1, TLR2, TLR4, TLR5, TLR6, and TLR11) or in the endolysosomes (TLR3, TLR7, TLR8, TLR9, and TLR10). While the role of TLRs in sepsis has been well described, more recent data indicate that a subset of the TLRs—TLR4 in particular—also recognize DAMPs released from injured cells and tissues. Among the DAMP ligands for surface TLR are HMGB1, HSPs, S100 proteins, and several others. Endosomal TLR ligands include mtDNA and other mitochondrial proteins.

What we know about TLR signaling events has largely been derived from the TLR-mediated response to bacterial pathogens. However, it is largely accepted that the intracellular adaptors required for signal transmission by TLRs are conserved and utilized for “damage” sensing of endogenous (“self”) ligands as well. The intracellular domain structure of TLRs is highly conserved and is characterized by a cytoplasmic Toll/IL-1R homology (TIR) domain. Binding of ligand to the receptor results in a receptor dimer, either a homodimer (e.g., TLR4/TLR4) or heterodimer (e.g., TLR2/TLR1), which recruits a number of adaptor proteins to the TIR domains through TIR-TIR interaction. With one exception (TLR3), the universal adaptor protein central to the TLR signaling complex is myeloid differentiation factor 88 (MyD88), a member of the interleukin-1 receptor subfamily. MyD88 works through the recruitment of a second TIR-containing adaptor, MyD88 adaptor-like protein (Mal), also termed Toll/interleukin-1 receptor-containing adaptor protein, or TIRAP) in the context of TLR4 and TLR2 signaling, which serves as a bridge between MyD88 and activated TLRs to initiate signal transduction. It is interesting that Mal’s adaptor function requires cleavage of the carboxy-terminal portion of the protein by caspase-1, a key effector of the inflammasome. This finding suggests an important synergy between TLRs and the inflammasome that may potentiate TLR-mediated signaling.

Signaling through the MyD88-dependent pathway occurs once the receptor is ligated at the cell surface. Receptor ligation, dimerization, and recruitment of the MyD88/Mal complex results in the activation of numerous cytoplasmic protein kinases, including IL-1 receptor–associated kinases, resulting in an interaction with tumor necrosis factor receptor–associated factor 6 (TRAF6). TRAF6, an E3 ubiquitin ligase, forms a complex with two other proteins, which together activate the complex that subsequently phosphorylates IxB kinase (IKK)-β and the MAP kinases (MAPKs). Ultimately, the phosphorylation of IxB leads to its degradation, which frees NF-κB and allows its translocation to the nucleus and the transcription of NF-κB target genes. Simultaneously, MAP kinase activation is critical for activation of the activator protein-1 (AP-1) transcription factor, and thus production of inflammatory cytokines.

Two other TIR domain-containing adaptor proteins, TIR-domain-containing adapter-inducing interferon-β (TRIF) and TRIF-related adaptor molecule (TRAM), are important to TLR-signaling events that are involved in the MyD88-independent signaling pathways, activated by TLR3 and TLR4. One distinction of MyD88-dependent and -independent TLR signaling is that TLR4/TRIF transduction begins after the signaling complex is internalized into endosomes. The MyD88-independent pathway acts through TRIF to activate NF-κB, similar to the MyD88-dependent pathway. However, TRIF can also recruit other signaling molecules to phosphorylate interferon-regulatory factor 3 (IRF3), which induces expression of type I IFN genes.
The initiation of transcription by TLR activation leads to the upregulation of a large cohort of target genes that include interferons α and β (IFNα/β), nitric oxide synthase 2 (NOS2A), and tumor necrosis factor (TNF), which play critical roles in initiating innate immune responses to cellular injury and stress. Given the importance of TLR triggering of the innate immune response to immune homeostasis, it is no surprise that the process is tightly regulated. TLR signaling is controlled at multiple levels, both posttranslationally via ubiquitination, phosphorylation, and micro RNA actions that affect mRNA stability, and by the localization of the TLRs and their signaling complexes within the cell.

TLR expression is significantly increased following blunt traumatic injury. A recent study of patients undergoing “high-risk” surgical procedures examined immune parameters, including TLR expression, that were associated with the development of SIRS. The investigators demonstrated that patients who developed postoperative SIRS exhibited increased TLR and TLR expression on a subgroup of CD14+ monocytes when compared to those patients with an uneventful recovery. Moreover, the upregulation of TLR in these patients was associated with increased expression of IL-6. Interestingly, the authors hypothesize that preoperatively, a subset of monocytes may already be primed to act in this way and thus may identify a vulnerable patient group.

**Nucleotide-Binding Oligomerization Domain (NOD)-like Receptor (NLR) Family.** The nucleotide-binding oligomerization domain-like receptors (NLRs) are a large family of proteins composed of intracellular PRRs that sense both endogenous (DAMPs) and exogenous (PAMPs) molecules to trigger innate immune activation. The best characterized of the NLRs is the NLR family pyrin domain-containing 3 (NLRP3), which is highly expressed in peripheral blood leukocytes. It forms the key “sensing” component of the larger, multiprotein inflammasome complex, which is composed of NLRP3; the adapter protein apoptosis-associated speck-like protein containing a CARD (ASC); and the effector protein, caspase 1. Activation of the NLRP3 inflammasome is tightly regulated, both transcriptionally and at the posttranslational level. An initial priming event (typically via TLR/nuclear factor [NF]-kB signals) upregulates NLRP3 expression. The receptor then resides in the cytoplasm in an inactive form due to an internal interaction between two adjacent domains. When phagocytosed DAMPs are sensed by NLRP3, this second event releases the self-repression. The protein can then oligomerize and recruit other complex members. The net result is the auto-activation of pro-caspase 1 to caspase 1. This event is pivotal to all known inflammasome signaling pathways. The caspase-1 products assemble to form the IL-1 converting enzyme (ICE), which cleaves the proforms of IL-1β, IL-18, and IL-33 to form their active, mature forms required for secretion from the cell.

The inflammasome-activated cytokines, IL-1β and IL-18, are potent proinflammatory molecules that promote key immune responses essential to host defense. Both IL-1β and IL-18 lack a signal sequence, which is usually necessary for the secretion of cellular proteins. More than 20 proteins in addition to IL-1β and IL-18 undergo unconventional protein secretion independent of the ER and Golgi, including a number of the DAMP molecules. Currently, the mechanisms responsible for unconventional protein secretion are not understood; however, the process is also evident in yeast under conditions of cellular stress. It makes evolutionary sense that a mechanism for rapid secretion of stored proteins essential to the stress response is highly conserved.

Evidence suggests that genetic variations in the NLRP3 gene might affect the magnitude of immune inflammatory responses following trauma. Single nucleotide polymorphisms within the NLRP3 gene were found to be associated with increased risk of sepsis and MODS in patients with major trauma. In an animal model of burn injury, early inflammasome activation has been detected in a variety of immune cells (NK cells, CD4/CD8 T cells, and B cells), as determined by the assessment of caspase 1 cleavage by flow cytometry. Further, inhibition of caspase 1 activity in vivo results in increased burn mortality, suggesting that inflammasome activation may play an unanticipated protective role in the host response to injury that may be linked to increased production of specific cytokines.

CNS trauma induces inflammasome activation in the nervous system. Moreover, exosomes containing inflammasome protein cargo are secreted into cerebral spinal fluid and can be detected in patients with TBI. In an animal model of TBI, controlled cortical impact, exosomes containing inflammasome proteins are detected in the serum and appear to be linked to TBI-related acute lung injury.

**C-Type Lectin and Lectin-Like Receptors.** Macrophages and dendritic cells possess receptors that detect molecules released from damaged or dying cells in order to retrieve and process antigens for T cell presentation. A key family of receptors that directs this process is the C-type lectin (CLR) and C-type lectin-like (CTLR) receptor family that includes the selectin and the mannose receptor families. CLR and CTLR bind carbohydrates in both a calcium-dependent (CLR) and -independent (CTLR) fashion. Best described for their sensing of PAMPs, the CLRs can also act to promote the endocytosis and clearance of cell debris, which can be processed and presented to T cells. CTLR receptor recognition of DAMPs of intracellular origin, such as F-actin and the ribonucleoprotein SAP-130, can trigger multiple signaling pathways leading to NF-kB, type I interferon (IFN), and/or inflammasome activation. Expression of the CTLR, MINCLE (macrophage-inducible C-type lectin), is increased after exposure to proinflammatory stimuli or cell stress. When MINCLE senses self-damage in association with ischemia-reperfusion injury, it promotes proinflammatory cytokine, chemokine, and nitric oxide production.

**Receptor for Advanced Glycation End Products (RAGE).** Another key player in the sterile inflammatory response to injury is the transmembrane receptor, the receptor for advanced glycation endproducts, or RAGE. Highly conserved across species, RAGE is a member of the immunoglobulin superfamily that is constitutively expressed at high levels in the lung, with low/absent expression in other adult cell types. However, proinflammatory stimuli and the presence of RAGE ligands can increase RAGE expression on immune cells such as neutrophils, macrophages, and lymphocytes. RAGE also exists as a soluble form (sRAGE) composed only of the extracellular domain, which can bind to and sequester RAGE ligands, without consequent signaling events. RAGE binds diverse ligands, including HMGB1 and S100, as well as components of the extracellular matrix such as collagen. As a receptor, RAGE recognizes the three-dimensional structure of its ligands that allow it to bind a diverse repertoire of molecules, independent of their amino acid sequence.
Signaling via RAGE is mediated via multiple pathways leading to transcriptional activation and release of proinflammatory mediators. Animal models have linked RAGE to acute lung injury in ischemia-reperfusion models. In clinical studies, high sRAGE levels have been linked to prolonged mechanical ventilation post lung transplant as well as worse outcomes following TBI-associated acute lung injury. These events likely represent a role for an HMGB1-RAGE axis in these pathologic processes.

**Soluble Pattern Recognition Molecules: The Pentraxins.**

Soluble pattern recognition molecules (PRMs) are a molecularly diverse group of molecules that share a conserved mode of action defined by complement activation, agglutination and neutralization, and opsonization. The best described of the PRMs are the pentraxins. PRMs can be synthesized at sites of injury and inflammation by macrophages and dendritic cells, while neutrophils can store PRMs and release them rapidly following activation. In addition, epithelial tissues (the liver in particular) serve as a reservoir source for systemic mass release. The short pentraxin, C-reactive protein (CRP), was the first PRM to be identified. Serum amyloid protein (SAP), which has 51% sequence similarity to human CRP, also contains the pentraxin molecular signature. CRP and SAP plasma levels are low (≤3 mg/L) under normal circumstances. However, CRP is synthesized by the liver in response to interleukin-6, increasing serum levels more than a 1000-fold. Thus, CRP is considered part of the **acute-phase protein response** in humans. For this reason, C-reactive protein has been studied as a marker of the proinflammatory response in many clinical settings, including appendicitis, vasculitis, and ulcerative colitis. CRP and SAP are ancient immune molecules that share many functional properties with antibodies: they bind bacterial polysaccharides, ECM components, apoptotic cells, and nuclear materials, as well as all three classes of Fcγ receptors (FcγR). Both molecules also participate in the activation and regulation of complement pathways. In this way, short pentraxins can link immune cells to the complement system.

Finally, there is significant data to support a role for pentraxin 3 (PTX3), a long pentraxin family member, in the “sterile” inflammatory response associated with cellular stress. While CRP is produced solely in the liver, PTX3 is produced by various cells in peripheral tissues, including immune cells. PTX3 plasma concentrations increase rapidly in various inflammatory conditions, including sepsis. Further, in a recent prospective study of polytraumatized patients, serum PTX3 concentrations were highly elevated, peaking at 24 hours. Further, PTX3 concentrations at admission were associated with injury severity, while higher PTX3 serum concentrations 24 hours after admission correlated with lower probability for survival.

**CENTRAL NERVOUS SYSTEM REGULATION OF INFLAMMATION IN RESPONSE TO INJURY**

The central nervous system (CNS) communicates with the body through ordered systems of sensory and motor neurons, which receive and integrate information to generate a coordinated response. Rather than being an immune-privileged organ, recent work indicates that the CNS receives information with regard to injury-induced inflammation both via soluble mediators as well as direct neural projections that transmit information to regulatory areas in the brain (Fig. 2-2). How does the

---

**Figure 2-2.** Neural circuit relaying messages of localized injury to the brain (nucleus tractus solitarius). The brain follows with a hormone release (adrenocorticotrophic hormone [ACTH], glucocorticoids) into the systemic circulation and by sympathetic response. The vagal response rapidly induces acetylcholine release directed at the site of injury to curtail the inflammatory response elicited by the activated immunocytes. This vagal response occurs in real time and is site specific. EPI = epinephrine; IL-1 = interleukin-1; NOREPI = norepinephrine; TNF = tumor necrosis factor. (Adapted with permission from Tracey KJ: The inflammatory reflex, Nature. 2002 Dec 19-26;420(6917):853-859.)
CNS sense inflammation? DAMPs and inflammatory molecules convey stimulatory signals to the CNS via multiple routes. For example, soluble inflammatory signaling molecules from the periphery can reach neurons and glial cells directly through the fenestrated endothelium of the circumventricular organs (CVO) or via a leaky blood-brain barrier in pathological settings following a traumatic brain injury. In addition, inflammatory stimuli can interact with receptors located on the brain endothelial cells to generate a variety of proinflammatory mediators (cytokines, chemokines, adhesion molecules, proteins of the complement system, and immune receptors) that directly impact the brain parenchyma. Not surprising, this response is countered by potent anti-inflammatory signaling, a portion of which is provided by the HPA axis and the release of systemic glucocorticoids. Inflammatory stimuli in the CNS result in behavioral changes, such as increased sleep, lethargy, reduced appetite, and the most common feature of infection, fever.

Information regarding peripheral inflammation and tissue damage can also be signaled to the brain via afferent neural fibers, particularly those of the vagus nerve. These afferent fibers can interconnect with neurons that project to the hypothalamus to modulate the HPA axis. In addition, afferent vagal nerve impulses modulate cells in the brain stem, at the dorsal motor nucleus of the vagus, from which efferent preganglionic parasympathetic originate. Axons from these cells, which comprise the visceromotor component of the vagus nerve, form an “inflammatory reflex” that feeds back to the periphery to regulate inflammatory signaling events. Mechanistic insight into the “inflammatory reflex” was provided by the observation in several experimental model systems, that vagal stimulation reduced proinflammatory cytokine production from the spleen. This effect was dependent on both vagal efferent signals and on splenic catecholaminergic nerve fibers that originated in the celiac plexus and terminated in the T cell–rich area of the spleen. The vagal efferent fibers that terminated within the celiac ganglion were found to synapse on the cell bodies of the catecholaminergic splenic nerves. Vagal stimulation resulted in the firing of these adrenergic nerves, resulting in the activation of β2-adrenergic receptors on a subset of acetylcholine (ACH)-producing T cells. The ACh released from this T cell population targets α7 nicotinic ACh receptors (a7nACHR) expressed by splenic macrophages. Macrophage ACh receptor ligation blocks cell activation, inhibiting cytokine production and shifting the macrophages towards an M2 anti-inflammatory phenotype. Moreover, ACh-receptor binding inhibits intracellular signaling including the nuclear translocation of NF-κB and the activation of the inflammasome. In a rat model of hemorrhagic shock with reperfusion, vagal nerve stimulation post injury resulted in a decrease in the inflammatory response to hemorrhage.

Neuroendocrine Response to Injury

Traumatic injury results in complex neuroendocrine signaling from the brain that serves to enhance immune defense and rapidly mobilize substrates necessary to meet essential energy and structural needs. The two principle neuroendocrine pathways that orchestrate the host response are the hypothalamic-pituitary-adrenal (HPA) axis, which results in the release of glucocorticoid hormones, and the sympathetic nervous system, which results in release of the catecholamines, epinephrine (EPI), and norepinephrine (NEP). Virtually every hormone of the HPA axis influences the physiologic response to injury and stress (Table 2-4), but some with direct influence on the inflammatory response or immediate clinical impact are highlighted here, including growth hormone (GH), macrophage inhibitory factor (MIF), aldosterone, and insulin.

The Hypothalamic-Pituitary-Adrenal Axis. One of the main mechanisms by which the brain responds to injury-associated stress is through activation of the hypothalamic-pituitary-adrenal (HPA) axis. Following injury, corticotrophin-releasing hormone (CRH) is secreted from the paraventricular nucleus (PVN) of the hypothalamus. This action is mediated in part by circulating cytokines produced as a result of the innate immune response to injury. These include tumor necrosis factor-α (TNF-α) IL-1β, IL-6, and the type I interferons (IFN-α/β). Cytokines that are produced as a result of the adaptive immune response (IL-2 and IFN-γ) are also capable of increasing cortisol release. Direct neural input via afferent vagal fibers that interconnect with neurons projecting to the hypothalamus can also trigger CRH release. CRH acts on the anterior pituitary to stimulate the secretion of adrenocorticotropic hormone (ACTH) into the systemic circulation. Interestingly, the cytokines that act on the hypothalamus are also capable of stimulating ACTH release from the anterior pituitary so that marked elevations in ACTH and in cortisol can occur that are proportional in magnitude to the injury severity. Additionally, pain, anxiety, vasopressin, angiotensin II, cholecystokinin, vasoactive intestinal peptide, and catecholamines all contribute to ACTH release in the injured patient.

### Table 2-4

<table>
<thead>
<tr>
<th>Hormones regulated by the hypothalamus, pituitary, and autonomic system</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hypothalamic Regulation</strong></td>
</tr>
<tr>
<td>Corticotropin-releasing hormone</td>
</tr>
<tr>
<td>Thyrotropin-releasing hormone</td>
</tr>
<tr>
<td>Growth hormone–releasing hormone</td>
</tr>
<tr>
<td>Luteinizing hormone–releasing hormone</td>
</tr>
<tr>
<td><strong>Anterior Pituitary Regulation</strong></td>
</tr>
<tr>
<td>Adrenocorticotropic hormone</td>
</tr>
<tr>
<td>Cortisol</td>
</tr>
<tr>
<td>Thyroid-stimulating hormone</td>
</tr>
<tr>
<td>Triiodothyronine</td>
</tr>
<tr>
<td>Growth hormone</td>
</tr>
<tr>
<td>Gonadotrophins</td>
</tr>
<tr>
<td>Sex hormones</td>
</tr>
<tr>
<td>Insulin-like growth factor</td>
</tr>
<tr>
<td>Somatostatin</td>
</tr>
<tr>
<td>Prolactin</td>
</tr>
<tr>
<td>Endorphins</td>
</tr>
<tr>
<td><strong>Posterior Pituitary Regulation</strong></td>
</tr>
<tr>
<td>Vasopressin</td>
</tr>
<tr>
<td>Oxytocin</td>
</tr>
<tr>
<td><strong>Autonomic System</strong></td>
</tr>
<tr>
<td>Norepinephrine</td>
</tr>
<tr>
<td>Epinephrine</td>
</tr>
<tr>
<td>Aldosterone</td>
</tr>
<tr>
<td><strong>Renin-Angiotensin System</strong></td>
</tr>
<tr>
<td>Insulin</td>
</tr>
<tr>
<td>Glucagon</td>
</tr>
<tr>
<td>Enkephalins</td>
</tr>
</tbody>
</table>
ACTH acts on the zona fasciculate of the adrenal glands to synthesize and secrete glucocorticoids (Fig. 2-3). Cortisol is the major glucocorticoid in humans and is essential for survival during significant physiologic stress. The resulting increase in cortisol levels following trauma have several important anti-inflammatory actions.

Cortisol elicits its many actions through a cytosolic receptor, the glucocorticoid receptor (GR). Because it is lipid soluble, cortisol can diffuse through the plasma membrane to interact with its receptor, which is sequestered in the cytoplasm in a complex with heat shock proteins (Fig. 2-4). Upon ligand binding, the GR is activated and can employ a number of mechanisms to modulate proinflammatory gene transcription and signaling events, with a “net” anti-inflammatory effect. For example, the activated GR complex can interact with transcription factors to sequester them in the cytoplasm, promote their degradation, or inhibit them through other mechanisms. Affected target genes include proinflammatory cytokines, growth factors, adhesion molecules, and nitric oxide. In addition, glucocorticoids can negatively affect the access of the transcription factor, nuclear factor-κB (NF-κB), to the promoter regions of its target genes via a mechanism that involves histone deacetylase 2. In this way, glucocorticoids can inhibit a major mechanism by which TLR ligation induces proinflammatory gene expression. The GR complex can also bind to specific nucleotide sequences (termed glucocorticoid response elements) to promote the

Figure 2-3. Steroid synthesis from cholesterol. Adrenocorticotropic hormone (ACTH) is a principal regulator of steroid synthesis. The end products are mineralocorticoids, glucocorticoids, and sex steroids.

Figure 2-4. Simplified schematic of steroid transport into the nucleus. Steroid molecules (S) diffuse readily across cytoplasmic membranes. Intracellularly, the receptors (R) are rendered inactive by being coupled to heat shock protein (HSP). When S and R bind, HSP dissociates, and the S-R complex enters the nucleus, where the S-R complex induces DNA transcription, resulting in protein synthesis. mRNA = messenger RNA.
transcription of genes, which have anti-inflammatory functions. These include interleukin-10 and interleukin-1-receptor antagonists. Further, GR complex activation can indirectly influence TLR activity via an interaction with signaling pathways such as the mitogen-activated protein kinase and transforming growth factor–activated kinase-1 (TAK1) pathways. Finally, a recent report demonstrated that GCs target suppressor of cytokine signaling 1 (SOCS1) and type 1 interferons to regulate TLR-induced signaling events.81

Adrenal insufficiency represents a clinical syndrome highlighted largely by inadequate amounts of circulating cortisol and aldosterone. Classically, adrenal insufficiency is described in patients with atrophic adrenal glands caused by exogenous steroid administration who undergo a stressor such as surgery. These patients subsequently manifest signs and symptoms such as tachycardia, hypotension, weakness, nausea, vomiting, and fever. However, it is now apparent that severe traumatic injury associated with an extended proinflammatory response can increase the risk of critical illness–related corticosteroid insufficiency, or CIRCI.

In the postinjury setting, CIRCI describes a phenomenon in which an exaggerated proinflammatory response is associated with a blunted adrenocortical response.82 Factors that have been linked to CIRCI include dysregulation of the HPA axis with altered adrenal synthesis of cortisol, altered cortisol metabolism, and tissue resistance to corticosteroids with inadequate glucocorticoid receptor activity. As a consequence, cortisol levels prove insufficient for the severity of stress. Investigators have determined that CIRCI in trauma patients occurs more frequently than previously thought.83 In one recent study, CIRCI occurred in 38 of 70 patients with multiple injuries. In most cases, the diagnosis was made within the first 48 hours following injury.84 Laboratory findings in adrenal insufficiency include hypoglycemia from decreased gluconeogenesis, hyponatremia from impaired renal tubular sodium resorption, and hyperkalemia from diminished kaliuresis. Recommended guidelines to diagnose CIRCI include measuring delta cortisol (change in baseline cortisol at 60 min of <9 µg/dL) after cosyntropin (250 µg) administration and a random plasma cortisol of <10 µg/dL. Treatment strategies remain controversial in the setting of trauma.85,86

Macrophage Migration Inhibitory Factor Modulates Cortisol Function. Macrophage migration inhibitory factor (MIF) is a proinflammatory cytokine expressed by a variety of cells and tissues, including the anterior pituitary, macrophages and T lymphocytes. MIF is also classified as an atypical chemokine that binds to the CXCR4 receptor.87 Several important functions of MIF in innate and adaptive immune responses and in inflammation have been described supporting the idea that MIF may function to counteract the anti-inflammatory activity of glucocorticoids.88 For example, MIF has been reported to play a central role in the exacerbation of inflammation associated with acute lung injury, where it has been detected in the affected lungs and in alveolar macrophages. MIF has also been reported to upregulate the expression of Toll-like receptor 4 (TLR4) in macrophages,89 and an early increase in plasma MIF has been detected in severely injured patients and was found to correlate with NF-κB translocation and respiratory burst in PMNs derived from severely injured patients. Further, nonsurvivors were shown to have higher serum MIF concentrations early after injury than survivors.90 These data suggest that targeting MIF with available small molecule inhibitors may be a novel therapeutic strategy for preventing early PMN activation and subsequent organ failure in severely injured patients.

Growth Hormone, Insulin-Like Growth Factor, and Ghrelin. Growth hormone (GH) is a neurohormone expressed primarily by the pituitary gland that has both metabolic and immunomodulatory effects. GH promotes both protein synthesis and insulin resistance while enhancing the mobilization of fat stores. GH secretion is upregulated by hypothalamic GH–releasing hormone and downregulated by somatostatin. GH primarily exerts its downstream effects through direct interaction with GH receptors and through the enhanced hepatic synthesis of insulin-like growth factor-1 (IGF-1), an anabolic growth factor that is known to improve the metabolic rate, gut mucosal function, and protein loss after traumatic injury. Less than 5% of IGF-1 circulates free in the plasma, with the remainder bound principally to one of six IGF-binding proteins (IGFBPs), the majority to IGFBP-3. In the liver, IGF-1 stimulates protein synthesis and glycosogenesis; in adipose tissue, it increases glucose uptake and lipid utilization; and in skeletal muscles, it mediates glucose uptake and protein synthesis. In addition to its effects on cellular metabolism, GH enhances phagocytic activity of immune cells and tissues, including the anterior pituitary, macrophages and T lymphocytes. MIF is also classified as an atypical chemokine expressed by a variety of cells and tissues, including the anterior pituitary, macrophages and T lymphocytes. MIF is also classified as an atypical chemokine that binds to the CXCR4 receptor.87 Several important functions of MIF in innate and adaptive immune responses and in inflammation have been described supporting the idea that MIF may function to counteract the anti-inflammatory activity of glucocorticoids.88 For example, MIF has been reported to play a central role in the exacerbation of inflammation associated with acute lung injury, where it has been detected in the affected lungs and in alveolar macrophages. MIF has also been reported to upregulate the expression of Toll-like receptor 4 (TLR4) in macrophages,89 and an early increase in plasma MIF has been detected in severely injured patients and was found to correlate with NF-κB translocation and respiratory burst in PMNs derived from severely injured patients. Further, nonsurvivors were shown to have higher serum MIF concentrations early after injury than survivors.90 These data suggest that targeting
inflammatory markers. Moreover, the high ghrelin levels were a positive predictor of ICU-survival in septic patients, matching previous results from animal models. Based on these data, ghrelin seems to exert anti-inflammatory effects that are mediated by diverse pathways. Recent work has linked ghrelin to a novel pathway mediated by upregulation of uncoupling protein 2 (UCP2) particularly in the setting of traumatic brain injury.95

**The Role of Catecholamines in Postinjury Inflammation.**
Injury-induced activation of the sympathetic nervous system results in secretion of acetylcholine from the preganglionic sympathetic fibers innervating the adrenal medulla. The adrenal medulla is a special case of autonomic innervation and is considered a modified postganglionic neuron. Thus, acetylcholine signaling to the resident chromaffin cells ensures that a surge of epinephrine (EPI) and norepinephrine (NE) release into the circulation takes place in a ratio that is tightly regulated by both central and peripheral mechanisms. Circulating levels of EPI and NE are three- to fourfold elevated, an effect that persists for an extended time. The release of EPI can be modulated by transcriptional regulation of phenylethanolamine N-methyltransferase (PNMT), which catalyzes the last step of the catecholamine biosynthesis pathway methylating NE to form EPI. PNMT transcription, a key step in the regulation of epinephrine production, is activated in response to stress and tissue hypoxia by hypoxia-inducible factor 1α (HIF1A).

Catecholamine release almost immediately prepares the body for the “fight or flight” response with well-described effects on the cardiovascular and pulmonary systems, and on metabolism. These include increased heart rate, myocardial contractility, conduction velocity, and blood pressure; the redirection of blood flow to skeletal muscle; increased cellular metabolism throughout the body; and mobilization of glucose from the liver via glycogenolysis, gluconeogenesis, lipolysis, and ketogenesis. To compound the resulting hyperglycemia, insulin release is decreased mainly through the stimulation of α-adrenergic pancreatic receptors. Hyperglycemia, as will be discussed, contributes to the proinflammatory response and to further mitochondrial dysfunction.

The goal of this well-orchestrated catecholamine response is to reestablish and maintain the systems’ homeostasis, including the innate immune system. Circulating catecholamines can directly influence inflammatory cytokine production.96 Data indicate that basal EPI levels condition the activity and responsiveness of cytokine-secreting cells, which may explain large inter-individual variability in innate cytokine profiles observed following injury. Epinephrine infusion at higher doses has been found to inhibit production of tumor necrosis factor (TNF) alpha in vivo and to enhance the production of the anti-inflammatory cytokine interleukin IL-10.97 Additionally, in vitro studies indicate that stress levels of glucocorticoids and epinephrine, acting in concert, can inhibit production of IL-12, a potent stimulator of Th1 responses. Further, they have been shown in vitro to decrease Th1 cytokine production and increase Th2 cytokine production to a significantly greater degree compared to either adrenal hormone alone. Thus, catecholamines secreted from the adrenal, specifically epinephrine, play a role in both innate proinflammatory cytokine regulation, as well as adaptive Th responses, and may act in concert with cortisol during the injury response to modulate cytokine activity.98

How are these effects explained? It is well established that a variety of human immune cells (e.g., mononuclear cells, macrophages, and granulocytes) express adrenergic receptors that are members of the family of G-protein coupled receptors that act through the activation of intracellular second messengers such as cAMP and calcium ions influx (discussed in more detail in the following section). These second messengers can regulate a variety of immune cell functions, including the release of inflammatory cytokines and chemokines.

The sympathetic nervous system also has direct immunomodulatory properties via its innervation of lymphoid tissues that contain resting and activated immune cells. The close proximity of sympathetic nerve terminals to immune cells responding to antigens (e.g., in the spleen) allows for a high concentration of norepinephrine to be localized within the microenvironment of antigen-activated immune cells. Norepinephrine can then interact with β,-adrenergic receptors expressed by CD4+ T and B lymphocytes, many of which also express α,-adrenergic receptors. Additionally, endogenous catecholamine expression has been detected in these cells (both CD4+ CD25+ T cells and phagocytes) as has the machinery for catecholamine synthesis. For example, monocytes contain inducible mRNA for the catecholamine-generating enzymes, tyrosine-hydroxylase, and dopamine-β-hydroxylase, and there is data to suggest that cells can regulate their own catecholamine synthesis in response to extracellular cues. Immune cell release of NE provides a way in which cells may exert additional regulation of inflammatory cell activation. For example, mature dendritic cells express both functional α- and β-adrenergic receptor (AR) types, as do monocytes and monocyte-derived macrophages, whereas B cells and Th1 cells express β,-AR exclusively.99 Exposure of PBMCs to NE triggers a distinct genetic profile that indicates a modulation of Th cell function. Thus, stimulation of AR results in varied signaling events to regulate both immune cell phenotype as well as mature cell function.100

**Aldosterone.** Aldosterone is a mineralocorticoid released by the zona glomerulosa of the adrenal cortex. It binds to the mineralocorticoid receptor (MR) of principal cells in the collecting duct of the kidney where it can stimulate expression of genes involved in sodium reabsorption and potassium excretion to regulate extracellular volume and blood pressure. Mineralocorticoid receptors (MR) have also been shown to have effects on cell metabolism and immunity. For example, recent studies show aldosterone interferes with insulin signaling pathways and reduces expression of the insulin-sensitizing factors, adiponectin and peroxisome proliferator activated receptor–γ (PPAR–γ), which contribute to insulin resistance. In the immune system, monocytes, lymphocytes, dendritic cells, and neutrophils have all been shown to possess a MR that binds aldosterone with high specificity, regulating sodium and potassium flux, as well as plasminogen activator inhibitor-1 and p22 phox expression in these cells.101 In dendritic cells, MR activation by aldosterone induces the secretion of proinflammatory cytokines. Further, aldosterone inhibits cytokine-mediated NF-κB activation in neutrophils, which also possess a functional MR.

**Insulin.** Hyperglycemia and insulin resistance are hallmarks of injury and critical illness due to the catabolic effects of circulating mediators, including catecholamines, cortisol, glucagon, and growth hormone. The increase in these circulating proglycemic factors, particularly epinephrine, induces glycogenolysis, lipolysis, and increased lactate production independent of available oxygen in a process that is termed “aerobic glycolysis.” Although there is an increase in insulin production at the
same time, severe stress is frequently associated with insulin resistance, leading to decreased glucose uptake in the liver and the periphery contributing to acute hyperglycemia. Insulin is a hormone secreted by the pancreas, which mediates an overall host anabolic state through hepatic glycogenesis and glycolysis, peripheral glucose uptake, lipogenesis, and protein synthesis.102

The insulin receptor (IR) is widely expressed and consists of two isoforms, which can form homo- or heterodimers with insulin binding. Dimerization leads to receptor autophosphorylation and activation of intrinsic tyrosine kinase activity. Downstream signaling events are dependent on the recruitment of the adaptor proteins, insulin receptor substrate (IRS-1), and Shc to the IR. Systemic insulin resistance likely results from proinflammatory signals, which modulate the phosphorylation of IRS-1 to affect its function.

Hyperglycemia during critical illness is predictive of increased mortality in critically ill trauma patients.103 It can modulate the inflammatory response by altering leukocyte functions and the resulting decreases in phagocytosis, chemotaxis, adhesion, and respiratory burst activities are associated with an increased risk for infection. In addition, glucose administration results in a rapid increase in NF-κB activation and proinflammatory cytokine production. Insulin therapy to manage hyperglycemia has grown in favor and has been shown to be associated with both decreased mortality and a reduction in infectious complications in select patient populations. However, the trend towards tight glycemic control in the intensive care unit failed to show benefit when examined in several reviews.104 Thus, the ideal blood glucose range within which to maintain critically ill patients and to avoid hypoglycemia has yet to be determined.

THE CELLULAR STRESS RESPONSES

Reactive Oxygen Species and the Oxidative Stress Response

Reactive oxygen and nitrogen species (ROS, RNS, respectively) are small molecules that are highly reactive due to the presence of unpaired outer orbit electrons. They can cause cellular injury to host cells and invading pathogens through the oxidation of cell membrane substrates, cellular proteins, and DNA. ROS has also been shown to have important roles as signaling messengers, particularly in the immune system.105,106

Oxygen radicals (superoxide anion, hydroxyl radical, hydrogen peroxide) are produced as a by-product of oxygen metabolism. The main areas of ROS production are oxidative processes involving the mitochondrial electron transport chain as well as those mediated by NADPH oxidases (NOX), a large class of ROS producing enzymes. Additional metabolic enzymes such as lipoxygenases, cytochrome P-450 and b5, and cyclooxygenases also produce ROS as by-products of their reactions.107 The synthesis of ROS is regulated at several checkpoints and via complex signaling mechanisms, including Ca2+ signaling, phosphorylation, and small G protein activation, which influence both the recruitment of the molecules required for NOX function and the synthesis of ROS in the mitochondria. Not surprisingly, NOX activation is triggered by a number of inflammatory mediators (e.g., TNF, chemokines, lysophospholipids, complement, and leukotrienes).

Host cells are protected from the damaging effects of ROS through a number of mechanisms. The best described of these is via the upregulation and/or activation of endogenous antioxidant enzymes such as superoxide dismutases, catalases, and glutaredoxins. Pyruvate kinase also provides negative feedback for ROS synthesis as do molecules that react nonenzymatically with ROS. Under normal physiologic conditions, ROS production is balanced effectively by these antioxidative strategies. As a consequence, ROS can act as signaling molecules through their ability to modulate cysteine residues by oxidation, and thus influence the functionality of target proteins.108 ROS can also contribute to transcriptional activity both indirectly through its effects on transcription factor lifespan, and directly through the oxidation of DNA.

The role for ROS has been well described in phagocytes, which utilize these small molecules for pathogen killing. A second important role for ROS is in the regulation of the inflammasome. As discussed previously, the inflammasome mediates the activation of inflammatory caspases leading to the production and secretion of mature cytokines in macrophages.109 Importantly, the best described inflammasome, NLRP3, is redox sensitive. Increased intracellular ROS enables the assembly of the protein complex.110 ROS also appears to be involved in adaptive immunity by influencing immune cell response.110 ROS can alter thiol group oxidative states on the cell surface and, in turn, affect cell signaling. Moreover, intracellular ROS can inhibit DNA transcription. ROS has been described as a prime source of phosphatase activation in both B and T lymphocytes, which can regulate the function of key receptors and intracellular signaling molecules in these cells by affecting phosphorylation events. Finally, large amounts of ROS cannot only suppress cell function, but also can result in cell death.111

The Unfolded Protein Response

Secreted, membrane-bound, and organelle-specific proteins fold in the lumen of the endoplasmic reticulum (ER) where they also receive their posttranslational modifications. Cellular stress disrupts the quality control required for this process leading to the accumulation of misfolded or unfolded proteins. These occurrences are sensed by a highly conserved array of signaling proteins in the ER that try to reestablish appropriate folding, while at the same time decreasing protein synthesis.112 The important proteins involved in this process include inositol requiring enzyme 1 (IRE1), protein kinase RNA (PKR)–like ER kinase (PERK), and activating transcription factor 6 (ATF6). Together, these proteins form a complex that generates the unfolded protein response (UPR). The UPR is a mechanism by which ER distress signals are sent to the nucleus to modulate transcription in an attempt to restore homeostasis. While obviously important to secretory epithelial cells, the UPR is also important to cells of the immune system.113

Significant protein misfolding results in an alarm signal that, if not addressed, can result in cell death. Genes activated in the UPR result not only in the inhibition of translation, but also other potentially immunomodulatory events including induction of the acute phase response, activation of NF-κB, and the generation of antibody-producing B cells.114 Activation of the UPR is also an alternative mechanism for activation of the inflammasome115 and can increase proinflammatory cytokine production.116

Markers of ER stress during critical illness have been demonstrated conclusively in burn patients,114,117 and in animal models they have been detected following hemorrhagic shock, correlating with the degree of organ dysfunction. Burn injury in particular leads to the marked reduction in ER calcium levels
and activation of UPR sensing proteins. Moreover, recent data in a series of burn patients strongly links the UPR to insulin resistance and hyperglycemia in these patients. Thus, a better understanding of the UPR, which is triggered by severe inflammation, may allow the identification of novel therapeutic targets for injury-associated insulin resistance.

Fibroblast growth factor-21 (FGF21), a recently identified hormone that regulates systemic metabolic homeostasis, is upregulated following mitochondrial damage and may be part of an integrated stress response that includes ER stress and the UPR. In animal models, induction of ER stress with chemical ER stressors results in increased FGF21 expression. A recent study examining FGF21 in critically ill patients demonstrated that serum FGF21 concentrations were eightfold higher in the critically ill patients as compared with the matched controls, regardless of the presence of sepsis. While FGF21 concentrations gradually decreased over time, they remained highly elevated at all studied time points and correlated with patient mortality. These data support the idea that the UPR may play an important role in the response to severe injury.

**Autophagy**

Under normal circumstances, cells need to have a way of disposing of damaged organelles and debris aggregates that are too large to be managed by proteosomal degradation. In order to accomplish this housekeeping task, cells utilize a process referred to as “macroautophagy” (autophagy), which is thought to have originated as a stress response. The steps of autophagy include the engulfment of cytoplasm/organelle by an “isolation membrane,” which is also called a phagophore. The edges of the phagophore then fuse to form the autophagosome, a double-membraned vesicle that sequesters the cytoplasmic material and is a characteristic feature of autophagy. The autophagosome then fuses with a lysosome to form an autolysosome, where the contents, together with the inner membrane, are degraded. This process is controlled by numerous autophagy-specific genes and by the specific kinase, mammalian target of rapamycin (mTOR).

As noted previously, autophagy is a normal cellular process that occurs in quiescent cells for cellular maintenance. However, under conditions of hypoxia and low cellular energy, autophagy is induced in an attempt to provide additional nutrients for energy production. The induction of autophagy promotes a shift from aerobic respiration to glycolysis and allows cellular components of the autophagosome to be hydrolyzed to energy substrates. Increased levels of autophagy are typical in activated immune cells and are a mechanism for the disposal of ROS and phagocytosed debris.

Recent data support the idea that autophagy plays an important role in the immune response. Autophagy is stimulated by Th1 cytokines and with activation of TLR in macrophages but is inhibited by Th2 cytokines. It is also recognized as an important regulator of cytokine secretion, particularly those cytokines of the IL-1 family that are dependent on inflammatory processing for activation. For example, autophagosomes can sequester and degrade pro-IL-1β and inflammasome components. In animal models of sepsis, inhibition of autophagy results in increased proinflammatory cytokine levels that correlate with increased mortality. These data suggest that autophagy is a protective mechanism whereby the cell can regulate the levels of cytokine production.

**Apoptosis**

Apoptosis (regulated cell death) is an energy-dependent, organized mechanism for clearing senescent or dysfunctional cells, including macrophages, neutrophils, and lymphocytes, without promoting an inflammatory response. This contrasts with cellular necrosis, which results in a disorganized sequence of intracellular molecular releases with subsequent immune activation and inflammatory response. Systemic inflammation modulates apoptotic signaling in active immunocytes, which subsequently influences the inflammatory response through the loss of effector cells.

Apoptosis proceeds primarily through two pathways: the extrinsic pathway and the intrinsic pathway. The extrinsic pathway is activated through the binding of death receptors (e.g., Fas, TNFR), which leads to the recruitment of Fas-associated death domain protein and subsequent activation of caspase 3 (Fig. 2-5). On activation, caspases are the effectors of apoptotic signaling because they mediate the organized breakdown of nuclear DNA. The intrinsic pathway proceeds through protein mediators (e.g., Bcl-2, Bcl-2-associated death promoter, Bcl-2–associated X protein, Bim) that influence mitochondrial membrane permeability. Increased membrane permeability leads to the release of mitochondrial cytochrome C, which ultimately activates caspase 3 and thus induces apoptosis. These pathways do not function in a completely autonomous manner because there is significant interaction and crosstalk between mediators of both extrinsic and intrinsic pathways. Apoptosis is modulated by several regulatory factors, including inhibitor of apoptosis proteins and regulatory caspases (e.g., caspases 1, 8, 10).

Apoptosis during sepsis may influence the ultimate competency of the acquired immune response. In a murine model of peritoneal sepsis, increased lymphocyte apoptosis was associated with mortality, which may be due to a resultant decrease in IFN-γ release. In postmortem analysis of patients who expired from overwhelming sepsis, there was an increase in lymphocyte apoptosis, whereas macrophage apoptosis did not appear to be affected. Clinical trials have observed an association between the degree of lymphopenia and disease severity in sepsis. In addition, after the phagocytosis of apoptotic cells by macrophages, anti-inflammatory mediators such as IL-10 are released that may exacerbate immune suppression during sepsis. Neutrophil apoptosis is inhibited by inflammatory products, including TNF, IL-1, IL-3, IL-6, GM-CSF, and IFN-γ. This retardation in regulated cell death may prolong and exacerbate secondary injury through neutrophil free radical release as the clearance of senescent cells is delayed.

**Necroptosis**

Cellular necrosis refers to the premature uncontrolled death of cells in living tissue typically caused by accidental exposure to external factors, such as ischemia, inflammation or trauma, which result in extreme cellular stress. Necrosis is characterized by the loss of plasma membrane integrity and cellular collapse with extrusion of cytoplasmic contents, but the cell nuclei typically remain intact. Recent data have defined a process by which necrosis occurs through a series of well-described steps that are dependent on a signaling pathway that involves the receptor-interacting protein kinase (RIPK) complex. Termed “necroptosis,” it occurs in response to specific stimuli, such as TNF- and TLR-mediated signals. For example, ligation of the tumor necrosis factor receptor 1 (TNFR1) under conditions in which caspase-8 is inactivated (e.g., by pharmacological
agents) results in the over-generation of ROS and a metabolic collapse. The net result is programmed necrosis (necroptosis). The effect of cell death by necroptosis on the immune response is not yet known. However, it is likely that the “DAMP” signature that occurs in response to necrotic cell death is an important contributor to the systemic inflammatory response. Evidence to support this concept was provided by investigators who examined the role of necroptosis in murine model of sepsis. They demonstrated that Ripk3−/− mice were capable of recovering body temperature better, exhibited lower circulating DAMP levels, and survived at higher rates than their WT littermates.126 These data suggest that the cellular damage that occurs with programmed necrosis exacerbates the sepsis-associated systemic inflammatory response.

**Pyroptosis**

Pyroptosis is a form of regulated cell death that is dependent on the activity of the proinflammatory caspase enzymes associated with the inflammasome and is thus an inflammatory form of cell death.127 Pyroptosis shares some features with apoptosis, including DNA fragmentation and positive annexin V staining, among others. However, it is associated with the activation of caspase-1 and the formation of caspase-1–dependent pores that allow early permeabilization of the cell membrane, electrolyte movement into the cells, and, finally, osmotic lysis of the cell.128 As a form of cell death, pyroptosis seems to be largely observed in macrophages, dendritic cells, and neutrophils, although it has been documented in other cells as well, especially if they express high levels of caspase-1. As noted, pyroptosis is linked to activation of the inflammasome, which can occur in response to diverse cell alarm signals, including DAMPs. Not surprising, the mechanism of cell death leads to the release of additional intracellular DAMPs, including HMGB1 and S100 proteins.

A recent study examined pyroptosis in peripheral blood mononuclear cells in a cohort of 60 trauma patients.129 The investigators found that the percentages of pyroptotic PBMCs were significantly higher in trauma patients than those in healthy

---

**Figure 2-5.** Signaling pathway for tumor necrosis factor receptor 1 (TNFR-1) (55 kDa) and TNFR-2 (75 kDa) occurs by the recruitment of several adapter proteins to the intracellular receptor complex. Optimal signaling activity requires receptor trimerization. TNFR-1 initially recruits TNFR-associated death domain (TRADD) and induces apoptosis through the actions of proteolytic enzymes known as caspases, a pathway shared by another receptor known as CD95 (Fas). CD95 and TNFR-1 possess similar intracellular sequences known as death domains (DDs), and both recruit the same adapter proteins known as Fas-associated death domains (FADDs) before activating caspase 8. TNFR-1 also induces apoptosis by activating caspase 2 through the recruitment of receptor-interacting protein (RIP). RIP also has a functional component that can initiate nuclear factor-kB (NF-kB) and c-Jun activation, both favoring cell survival and proinflammatory functions. TNFR-2 lacks a DD component but recruits adapter proteins known as TNFR-associated factors 1 and 2 (TRAF1, TRAF2) that interact with RIP to mediate NF-kB and c-Jun activation. TRAF2 also recruits additional proteins that are antiapoptotic, known as inhibitor of apoptosis proteins (IAPs). DED = death effector domain; I-kB = inhibitor of kB; I-kB/NF-kB = inactive complex of NF-kB that becomes activated when the I-kB portion is cleaved; JNK = c-Jun N-terminal kinase; MEKK1 = mitogen-activated protein/extracellular regulatory protein kinase kinase-1; NIK = NF-kB–inducing kinase; RAIDD = RIP-associated interleukin-1b-converting enzyme and ced-homologue-1–like protein with death domain, which activates proapoptotic caspases. (Adapted with permission from Vincent JL: Marshall JC, Cohen J: Update in Intensive Care and Emergency Medicine: Vol. 31: Immune Response in Critical Illness. Berlin: Springer-Verlag; 2002.)
controls and correlated with injury severity. Moreover, increase in pyroptotic PBMCs significantly correlated with elevated cytokine levels (IL-10, IL-18, and MCP-1) and was a strong predictor for the development of sepsis.

**MEDIATORS OF INFLAMMATION**

**Cytokines**

Cytokines are a class of protein signaling compounds that are essential for both innate and adaptive immune responses. Cytokines mediate a broad sequence of cellular responses, including cell migration, DNA replication, cell turnover, and immunocyte proliferation (Table 2-5). When functioning locally at the site of injury and infection, cytokines mediate the eradication of invading microorganisms and also promote wound healing. However, an exaggerated proinflammatory cytokine response to inflammatory stimuli may result in hemodynamic instability (i.e., septic shock) and metabolic derangements (i.e., muscle wasting). Anti-inflammatory cytokines also are released, at least in part, as an opposing influence on the proinflammatory cascade. These anti-inflammatory mediators may also result in immunocyte dysfunction and host immunosuppression. Cytokine signaling after an inflammatory stimulus can best be represented as a finely tuned balance of opposing influences and should not be oversimplified as a “black and white” proinflammatory/anti-inflammatory response. A brief discussion of the important cytokine molecules is included below.

**Tumor Necrosis Factor-α**. Tumor necrosis factor-α (TNF-α) is a potent inflammatory mediator that is rapidly mobilized in response to stressors such as injury and infection. It is primarily synthesized by immune cells, such as macrophages, dendritic cells, and T lymphocytes, and is generated in a precursor form that is expressed as a trimer on the surface of activated cells. After being processed by the metalloproteinase, TNF-α-converting enzyme (TACE, also known as ADAMS 17), a smaller, soluble form of TNF is released, which mediates its biological activities through types 1 and 2 TNF receptors (TNFR-1; TNFR-2). Transmembrane TNF-α also binds to TNFR-1 and -2, but its biological activities are likely mediated through TNFR-2. While the two receptors share homology in their ligand-binding regions, there are distinct differences that regulate their biologic function. For example, TNFR-1 is expressed by a wide variety of cells, but it is typically sequestered in the Golgi. Following appropriate cell signaling, TNFR-1 is mobilized to the cell surface, where it sensitizes cells to TNF or it can be cleaved from the surface in the form of a soluble receptor that can neutralize TNF. In contrast, TNFR-2 expression is confined principally to immune cells where it resides in the plasma membrane. Both TNF receptors are capable of binding intracellular adaptor proteins that lead to activation of complex signaling processes and mediate the effects of TNF.

Although the circulating half-life of soluble TNF is brief, it acts upon almost every differentiated cell type, eliciting a wide range of cellular responses. Moreover, it is one of the first cytokines to be released following trauma. In particular, TNF elicits many metabolic and immunomodulatory activities. It stimulates muscle breakdown and cachexia through increased catabolism, insulin resistance, and redistribution of amino acids to hepatic circulation as fuel substrates. TNF also mediates coagulation activation, cell migration, and macrophage phagocytosis and enhances the expression of adhesion molecules, prostaglandin E₂, platelet-activating factor, glucocorticoids, and eicosanoids. TNF-α increases endothelial cell permeability and activates macrophages, NK cells and lymphocytes to induce the secretion of various cytokines. While TNF is clearly playing a role in injury-induced inflammation, reports are conflicting whether postinjury TNF concentrations correlated with the development of multiple organ dysfunction syndrome.

**Interleukin-1 Family.** The IL-1 family of proteins contains 11 members. The best-studied of these are IL-1α and IL-1β and IL-1 receptor antagonist (IL-1Ra), but member cytokines also include IL-18, IL-33, IL-36, IL-3,7 and IL-38. IL-1α and IL-1β, which are encoded by two distinct IL-1 genes, share similar biologic functions despite limited sequence homology. They utilize the same cell surface receptor, termed IL-1 receptor type 1 (IL-1RI), which is present on nearly all cells. Once bound to its receptor, IL-1 initiates signaling events that result in the synthesis and release of a variety of inflammatory mediators. Members of the IL-1 family are expressed as proforms (pIL-1) that are matured through enzymatic cleavage. The IL-1α precursor is constitutively expressed and stored in a variety of healthy cells, including epithelium and endothelium, and its expression can be increased in response to proinflammatory or stress-associated stimuli.

Both the precursor and mature forms of IL-α have nearly identical biologic activities as measured by their ability to trigger IL-6 and TNF release. With appropriate signals, IL-1α can move both to the cell membrane, where it can act on adjacent cells bearing the IL-1R and to the nucleus where it can stimulate gene transcription. Pro-IL-1α can also be released passively from damaged injured cells in its active form. In this way, IL-1α is believed to function as a DAMP, which promotes the synthesis of inflammatory mediators, such as chemokines and eicosanoids. These mediators attract neutrophils to the injured site, facilitate their exit from the vasculature, and promote their activation. Once they have reached their target, neutrophil lifespan is extended by the presence of IL-1α.

IL-1β is a multifunctional proinflammatory cytokine whose expression and synthesis is tightly regulated and confined to activated cells, such as monocytes, tissue macrophages, and dendritic cells. In contrast to IL-1α, IL-1β is synthesized as an inactive precursor, pro-IL-1β, which is processed by the inflammasome in response to various stimuli, including cytokines and foreign pathogens, via pattern recognition receptors such as TLR4 as well as ROS. Mature IL-1β is then released from the cell via an unconventional secretory pathway. IL-1β has a spectrum of proinflammatory effects that are largely similar to those induced by TNF, and injection of IL-1β alone is sufficient to induce an acute inflammatory response. High doses of either IL-1β or TNF are associated with profound hemodynamic compromise. Interestingly, low doses of both IL-1β and TNF administered together elicit hemodynamic events similar to those elicited by high doses of either mediator, which suggests a synergistic effect.

There are two primary receptor types for IL-1: IL-1R1 and IL-1R2. IL-1R1 is widely expressed and mediates inflammatory signaling on ligand binding. IL-1R2 is proteolytically cleaved from the membrane surface to soluble form on activation and thus serves as another mechanism for competition and regulation of IL-1 activity. IL-1α or IL-1β bind first to the IL-1R1, which is considered the ligand-binding chain. This is followed by recruitment of a transmembrane co-receptor, termed the
**Table 2-5**  
Cytokines and their sources

<table>
<thead>
<tr>
<th>CYTOKINE</th>
<th>SOURCE</th>
<th>COMMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNF</td>
<td>Macrophages/monocytes, Kupffer cells, Neutrophils, NK cells, Astrocytes, Endothelial cells, T lymphocytes, Adrenal cortical cells, Adipocytes, Keratinocytes, Osteoblasts, Mast cells, Dendritic cells</td>
<td>Among earliest responders after injury; half-life &lt;20 min; activates TNF receptors 1 and 2; induces significant shock and catabolism</td>
</tr>
<tr>
<td>IL-1</td>
<td>Macrophages/monocytes, B and T lymphocytes, NK cells, Endothelial cells, Epithelial cells, Keratinocytes, Fibroblasts, Osteoblasts, Dendritic cells, Astrocytes, Adrenal cortical cells, Megakaryocytes, Platelets, Neutrophils, Neuronal cells</td>
<td>Two forms (IL-1α and IL-1β); similar physiologic effects as TNF; induces fevers through prostaglandin activity in anterior hypothalamus; promotes β-endorphin release from pituitary; half-life &lt;6 min</td>
</tr>
<tr>
<td>IL-2</td>
<td>T lymphocytes</td>
<td>Promotes lymphocyte proliferation, immunoglobulin production, gut barrier integrity; half-life &lt;10 min; attenuated production after major blood loss leads to immunocompromise; regulates lymphocyte apoptosis</td>
</tr>
<tr>
<td>IL-3</td>
<td>T lymphocytes, Macrophages, Eosinophils, Mast cells</td>
<td></td>
</tr>
<tr>
<td>IL-4</td>
<td>T lymphocytes, Mast cells, Basophils, Macrophages, B lymphocytes, Eosinophils, Stromal cells</td>
<td>Induces B-lymphocyte production of IgG4 and IgE, mediators of allergic and anthelmintic response; downregulates TNF, IL-1, IL-6, IL-8</td>
</tr>
<tr>
<td>IL-5</td>
<td>T lymphocytes, Eosinophils, Mast cells, Basophils</td>
<td>Promotes eosinophil proliferation and airway inflammation</td>
</tr>
<tr>
<td>IL-6</td>
<td>Macrophages, B lymphocytes, Neutrophils, Basophils, Mast cells, Fibroblasts, Endothelial cells</td>
<td>Elicited by virtually all immunogenic cells; long half-life; circulating levels proportional to injury severity; prolongs activated neutrophil survival</td>
</tr>
</tbody>
</table>

(Continued)
Table 2-5
Cytokines and their sources (Continued)

<table>
<thead>
<tr>
<th>CYTOKINE</th>
<th>SOURCE</th>
<th>COMMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-8</td>
<td>Macrophages/monocytes</td>
<td>Chemoattractant for neutrophils, basophils, eosinophils, lymphocytes</td>
</tr>
<tr>
<td></td>
<td>T lymphocytes Basophils Mast cells Epithelial cells Platelets</td>
<td></td>
</tr>
<tr>
<td>IL-10</td>
<td>T lymphocytes B lymphocytes Macrophages Basophils Mast cells Keratinocytes</td>
<td>Prominent anti-inflammatory cytokine; reduces mortality in animal sepsis and ARDS models</td>
</tr>
<tr>
<td>IL-12</td>
<td>Macrophages/monocytes Neutrophils Keratinocytes Dendritic cells B lymphocytes</td>
<td>Promotes Th1 differentiation; synergistic activity with IL-2</td>
</tr>
<tr>
<td>IL-13</td>
<td>T lymphocytes</td>
<td>Promotes B-lymphocyte function; structurally similar to IL-4; inhibits nitric oxide and endothelial activation</td>
</tr>
<tr>
<td>IL-15</td>
<td>Macrophages/monocytes Epithelial cells</td>
<td>Anti-inflammatory effect; promotes lymphocyte activation; promotes neutrophil phagocytosis in fungal infections</td>
</tr>
<tr>
<td>IL-18</td>
<td>Macrophages Kupffer cells Keratinocytes Adrenal cortical cells Osteoblasts</td>
<td>Similar to IL-12 in function; levels elevated in sepsis, particularly gram-positive infections; high levels found in cardiac deaths</td>
</tr>
<tr>
<td>IFN-γ</td>
<td>T lymphocytes NK cells Macrophages</td>
<td>Mediates IL-12 and IL-18 function; half-life of days; found in wounds 5–7 d after injury; promotes ARDS</td>
</tr>
<tr>
<td>GM-CSF</td>
<td>T lymphocytes Fibroblasts Endothelial cells Stromal cells</td>
<td>Promotes wound healing and inflammation through activation of leukocytes</td>
</tr>
<tr>
<td>IL-21</td>
<td>T lymphocytes</td>
<td>Preferentially secreted by Th2 cells; structurally similar to IL-2 and IL-15; activates NK cells, B and T lymphocytes; influences adaptive immunity</td>
</tr>
<tr>
<td>HMGB1</td>
<td>Monocytes/lymphocytes</td>
<td>High mobility group box chromosomal protein; DNA transcription factor; late (downstream) mediator of inflammation (ARDS, gut barrier disruption); induces “sickness behavior”</td>
</tr>
</tbody>
</table>

ARDS = acute respiratory distress syndrome; GM-CSF = granulocyte-macrophage colony-stimulating factor; IFN = interferon; Ig = immunoglobulin; IL = interleukin; NK = natural killer; Th1 = helper T cell subtype 1; Th2 = helper T cell subtype 2; TNF = tumor necrosis factor.

accessory protein (IL-1RαcP). A complex is formed of IL-1RI plus IL-1 plus the coreceptor. The signal is initiated with recruitment of the adaptor protein MyD88 to the Toll-IL-1 receptor (TIR) domains of the receptor complex and signal transduction via intermediates, which are homologous to the signal cascade initiated by TLRs. These events culminate in the activation of NF-κB and its nuclear translocation. Recent animal studies have implicated postinjury IL-1β in the exacerbation of traumatic brain injury. In a mouse model of polytrauma, which included both cortical brain injury and tibial fracture, mice that received both injuries demonstrated increased neuroinflammation, brain damage, and behavioral deficits compared to mice given an isolated-TBI. These changes correlated with increased IL-1β levels in the brain. Treatment...
with IL-1R antagonist post injury reduced volume loss in the injured cortex as well as markers of axonal injury, resulting in improved outcome in these animals.

IL-18 is also a member of the IL-1 superfamily of cytokines. \(^{138}\) First noted as an IFN-\(\gamma\)-inducing factor produced by LPS-stimulated macrophages, IL-18 expression is found both in immune cells and nonimmune cells at low to intermediate levels. However, activated macrophages and Kupffer cells produce large amounts of mature IL-18. Similar to IL-1\(\beta\), IL-18 is synthesized and stored as an inactive precursor form (pro-IL-18), and activation requires activation of the inflammasome resulting in the processing of pro-IL-18 by caspase-1. \(^{139}\) It then exits the cell through a nontraditional secretory pathway. The IL-18 receptor (IL-18R) is composed of two subunits, IL-18R\(\alpha\) and IL-18R\(\beta\), and is a member of the IL-1R superfamily that is structurally similar in their cytoplasmic domains to the TLR. One unique biological property of IL-18 is the potential, in conjunction with IL-12, to promote the Th1 response.

IL-18 induces IFN\(\gamma\) production by CD4\(^+\) T cells. IFN\(\gamma\), in turn, activates macrophages to produce inflammatory cytokines. Independent of its ability to induce interferon, IL-18 can act similarly to other proinflammatory cytokines by acting directly to increase in cell adhesion molecule expression, nitric oxide synthesis, and chemokine production by macrophages. \(^{140}\)

In a cohort of critically ill patients with acute lung injury and ARDS, inflammasome-related mRNA transcripts (CASP1, IL1B, and IL18) were increased in peripheral blood. Moreover, plasma IL-18 were also elevated and served as a marker of mortality risk. \(^{141}\) Recent studies suggest that IL-18 therapy may hold promise as effective therapy in promoting immune recovery after severe surgical stress. \(^{142}\)

IL-33, a second important IL-1 family member, is mainly expressed in surface epithelium and endothelium, where it is normally bound via an N-terminal chromatin-binding motif. \(^{143}\) Nuclear localization is important for its function and perhaps its regulation. Expression in mice of an IL-33 that lacks the nuclear localization sequence, results lethal inflammation, suggesting that nuclear localization acts to prevent unregulated extracellular release. \(^{144}\) Similar to HMGB1 and other IL-1 family members, IL-33 lacks a signal sequence for active secretion so that its release is injury-dependent. Once released from damaged cells, full length IL-33 is biologically active, but can be further processed by inflammatory proteases to a mature form that exhibits ten- to thirtyfold higher activity. IL-33 can binds to a member of the IL-1R family, ST2, leading to activation of NF-κB-mediated transcriptional events. ST2+ cells include macrophages, mast cells, Th2 cells, and tissue regulatory T cells (Tregs) which are important controllers of immune homeostasis.

**Interleukin-2 Family.** Interleukin-2 (IL-2) is a multifunctional cytokine produced primarily by CD4\(^+\) T cells after antigen activation, which plays pivotal roles in the immune response. Other cellular sources for IL-2 include CD8\(^+\) and NK T cells, mast cells, and activated dendritic cells. Discovered as a T cell growth factor, IL-2 also promotes CD8\(^+\) T cell and natural killer cell cytolytic activity and modulates T cell differentiation programs in response to antigen. Thus, IL-2 promotes naive CD4\(^+\) T cell differentiation into T helper 1 (Th1) and T helper 2 (Th2) cells while inhibiting T helper 17 (Th17) and T follicular helper (Tfh) cell differentiation. Moreover, IL-2 is essential for the development and maintenance of T regulatory (Treg) cells and for activation-induced cell death, thereby mediating tolerance and limiting inappropriate immune reactions. The upregulation of IL-2 requires calcium as well as protein kinase C signaling, which leads to the activation of transcription factors such as nuclear factor of activated T cells (NFAT) and NF-κB. MicroRNAs also play a role in the regulation of IL-2 expression. \(^{145}\)

IL-2 binds to IL-2 receptors (IL-2R), which are expressed on leukocytes. IL-2Rs are formed from various combinations of three receptor subunits: IL-2R\(\alpha\), IL-2R\(\beta\), and IL-2R\(\gamma\). These subunits form in low, medium, and high affinity forms of the receptor depending on the subunit combination. IL-2R\(\gamma\) has been renamed the common cytokine receptor γ chain (γc), which is now known to be shared by IL-2, IL-4, IL-7, IL-9, IL-15, and IL-21. Constitutive IL-2 receptor expression is low and is inducible by T cell receptor ligation and cytokine stimulation. Importantly, the transcription of each receptor subunit is individually regulated via a complex process to effect tight control of surface expression. Once the receptor is ligated, the major IL-2 signaling pathways that are engaged include Janus Kinase (JAK)-signal transducer and activator of transcription (STAT), Shc-Ras-MAPK, and phosphoinositol-3-kinase (PI3-K)-AKT. Partly due to its short half-life of <10 minutes, IL-2 is not readily detectable after acute injury. IL-2 receptor blockade induces immunosuppressive effects and can be pharmacologically used for organ transplantation. Attenuated IL-2 expression observed during major injury or blood transfusion may contribute to the relatively immunosuppressed state of the surgical patient. \(^{146}\)

**Interleukin-6 Family.** Following burn or traumatic injury, damage-associated molecular patterns (DAMPs) from damaged or dying cells stimulate TLRs to produce IL-6, a proinflammatory cytokine that plays a central role in host defense. IL-6 levels in the circulation are detectable by 60 minutes post injury, peak between 4 and 6 hours, and can persist for as long as 10 days. Further, plasma levels of IL-6 are proportional to the degree of injury. In the liver, IL-6 strongly induces a broad spectrum of acute-phase proteins such as C-reactive protein (CRP) and fibrinogen, among others, while it reduces expression of albumin, cytochrome P 450, and transferrin. In lymphocytes, IL-6 induces B cell maturation into immunoglobulin-producing cells and regulates Th17/Treg balance. IL-6 modulates T cell behavior by inducing the development of Th17 cells and inhibiting Treg cell differentiation in conjunction with transforming growth factor-β. IL-6 also promotes angiogenesis and increased vascular permeability, which are associated with local inflammatory responses. To date, ten IL-6 family cytokines have been identified, including IL-6, oncostatin M, neuropoietin, IL-11, IL-27, and IL-31. \(^{147}\)

The interleukin-6 receptor (IL-6R, gp80) is expressed on hepatocytes, monocytes, B cells, and neutrophils in humans. However, many other cells respond to IL-6 through a process known as trans-signaling. \(^{148}\) In this case, soluble IL-6Rs (sIL-6R) exist in the serum and bind to IL-6, forming an IL-6/sIL-6R complex. The soluble receptor is produced by proteolytic cleavage from the surface of neutrophils in a process that is stimulated by C-reactive protein, complement factors, and leukotrienes. The IL6/sIL6R complex can then bind to the gp130 receptor, which is expressed ubiquitously on cells. Upon IL-6 stimulation, gp130 transduces two major signaling pathways: the JAK-STAT3 pathway and the SHP2-Gab-Ras-Erk-MAPK pathway, which is regulated by cytoplasmic suppressor of cytokine signaling (SOCS3). These signaling events can lead to increased expression of adhesion molecules as well as proinflammatory chemokines and cytokines. High plasma IL-6 levels have been associated with mortality during
intra-abdominal sepsis. Moreover, prolonged (more than 3 days) elevation of IL-6 concentrations has been reported to correlate with the occurrence of complications and mortality following severe traumatic injury. More recently, a meta-analysis analyzed the predictive value of IL-6 for the development of complications and mortality after trauma and found that the concentration of IL-6 in the first 24 hours after trauma was predictive for the development of multiple organ failure and death.

**Interleukin-10 Family.** We have talked almost exclusively about the factors that initiate the proinflammatory response following cellular stress or injury. The reestablishment of immune homeostasis following these events requires the resolution of inflammation and the initiation of tissue repair processes. Interleukin-10 (IL-10) plays a central role in this anti-inflammatory response by regulating the duration and magnitude of inflammation in the host.

The IL-10 family currently has six members, including IL-10, IL-19, IL-20, IL-22, IL-24, and IL-26. IL-10 is produced by a variety of immune cells of both myeloid and lymphoid origin. Its synthesis is up-regulated during times of stress and systemic inflammation; however, each cell type that produces IL-10 does so in response to different stimuli, allowing for tight control of its expression. IL-10 exerts effects by binding to the IL-10 receptor (IL-10R), which is a tetramer formed from two distinct subunits, IL-10R1 and IL-10R2. Specifically, IL-10 binds first to the IL10R1 subunit, which then recruits IL-10R2, allowing the receptor complex to form. While IL-10R2 is widely expressed, IL-10R1 expression is confined to leukocytes so that the effects of IL-10 are confined to the immune system. Once receptor ligation occurs, signaling proceeds by the activation of JAK1 and STAT3. In particular, STAT3 in conjunction with IL-10 is absolutely required for the transcription of genes responsible for the anti-inflammatory response (AIR). IL-10 inhibits the secretion of proinflammatory cytokines, including TNF and IL-1, partly through the downregulation of NF-κB and thereby functions as a negative feedback regulator of the inflammatory cascade. In macrophages, IL-10 suppresses the transcription of 20% of all LPS-induced genes. Further, experimental models of inflammation have shown that neutralization of IL-10 increases TNF production and mortality, whereas restitution of circulating IL-10 reduces TNF levels and subsequent deleterious effects. Increased plasma levels of IL-10 also have been associated with mortality and disease severity after traumatic injury.

**Interleukin-12 Family.** Interleukin-12 (IL-12) is unique among the cytokines in being the only heterodimeric cytokine. This family, which includes IL-12, IL-23, IL-27, and IL-35, consists of an α-chain that is structurally similar to the IL-6 cytokine and a β-chain that is similar to the class I receptor for cytokines. The individual IL-12 family members are formed from various combinations of the α and β subunits. Despite the sharing of individual subunits, and the similarities of their receptors, the IL-12 cytokines have different biological functions. IL-12 and IL-23 are considered proinflammatory, stimulatory cytokines with key roles in the development of Th1 and Th17 subsets of helper T cells. In contrast, both IL-27 and IL-35 appear to have immunoregulatory functions that are associated with cytokine inhibition in specific Treg cell populations, particularly the Th17 cells. The effects of these cytokines require specific receptor chains that are also shared among the cytokines. The complexity of signaling is evidenced by the fact that these receptor chains can function both as dimers and as monomers. Ligation of the IL-12 receptors initiate signaling events via the JAK-STAT pathway.

IL-12 synthesis and release is increased during endotoxemia and sepsis. Together with IL-18, it stimulates lymphocytes to increase secretion of IFN-γ. IL-12 also stimulates NK cell cytotoxicity and helper T cell differentiation in this setting. IL-12 release is inhibited by IL-10, and its deficiency inhibits phagocytosis in neutrophils. In experimental models of inflammatory stress, IL-12 neutralization conferred a mortality benefit in mice during endotoxemia.

IL-23, an important IL-12 family member, is a heterodimeric cytokine comprised of a unique p19 subunit linked to a p40 subunit that is common with IL-12. IL-23 appears to be an important survival signal for a specific subset of T helper (Th) cells, Th-17 cells, where it provides a secondary stimulus for Th-17 differentiation. The Th-17 population of cells has recently been demonstrated to expand following traumatic injury and may mark an early phenotypic shift in cell population that has prognostic significance.

**Interleukin-17 Family.** IL-17A (also called IL-17) is the major effector cytokine predominantly produced by a subset of helper T cells, the T helper (Th)-17 cells. It is the founding member of the IL-17 family of cytokines, which includes IL-17A through F. The original described activity for IL-17A was to promote the differentiation of bone marrow progenitor cells along the granulopoietic lineage. Subsequent studies have confirmed that IL-17A is required for increasing circulating neutrophil numbers following stress. In the setting of infection, it is now known that IL-17 acts in conjunction with IL-23 to upregulate granulocyte-colony stimulating factor to promote granulopoiesis. IL-17A has also been shown to regulate the production of specific chemokines in both gut and lung epithelial cells and thus can modulate both the emigration of neutrophils into these tissues and their activation at the site. IL-17 also induces the expression of matrix metalloproteinases, which can make the extracellular matrix more accessible for immune cell recruitment.

IL-17 has the ability to induce the expression of important proinflammatory cytokines, including IL-1β, IL-6, and TNF from macrophages and other cells, and in this way, creates a self-sustaining loop that enhances its own production and strengthens its overall effects. Recent data supports a pivotal role for IL-17 in the posttrauma immune response and has identified associations between increased IL-17 expression associated with Th17- immune response outcomes following blunt trauma.

**Interferons.** Interferons were first recognized as soluble mediators that inhibited viral replication through the activation of specific antiviral genes in infected cells. Interferons are categorized into three types based on receptor specificity and sequence homology. The two major types, type I and type II are discussed in the following section.

Type I interferon family is composed of twenty distinct proteins. These include IFN-α, IFN-β, and IFN-ω, which are structurally related and bind to a common receptor. They are likely produced by most cell types and tissues after the detection of PAMPs/ DAMPs by cytosolic or membrane receptors, including TLR in macrophages and dendritic cells. Type 1 IFNs bind to a heterodimeric transmembrane receptor interferon (α and β) receptor 1, resulting in STAT activation and nuclear translocation. In the nucleus, dimeric STATs recruit an additional
transcriptional factor to form a complex capable of binding to interferon-stimulated response elements, inducing hundreds of IFN-stimulated genes.

Type I interferons influence adaptive immune responses by inducing the maturation of dendritic cells and by stimulating class I MHC expression. IFN-α and IFN-β also enhance immune responses by increasing the cytotoxicity of natural killer cells both in culture and in vivo. Further, they have been implicated in the enhancement of chemokine synthesis, particularly those that recruit myeloid cells and lymphoid cells. Thus, IFN/STAT signaling has important effects on the mobilization, tissue recruitment, and activation of immune cells that compose the inflammatory infiltrate. In contrast, type I IFNs appear to inhibit inflammasome activity, possibly via IL-10.159

The single type II interferon, IFN-γ is secreted by various T cells, NK cells, and antigen-presenting cells in response to bacterial antigens and cytokines. It functions as a key regulator of macrophage activation toward the “M1” proinflammatory phenotype.160 In response to IFN-γ, macrophages produce high levels of proinflammatory cytokines such as IL-1β, IL-12, IL-23, and TNF-α as well as reactive nitrogen and oxygen species. As a consequence, macrophages demonstrate enhanced phagocytosis and killing. In addition, IFN-γ signaling generates additional cytokines and inflammatory factors to sustain inflammation and help to maintain Th1 responses. IFN-γ regulation of macrophage activity may contribute to acute lung injury after major surgery or trauma. A diminished IFN-γ level, as seen in knockout mice, is associated with increased susceptibility to both viral and bacterial pathogens. In addition, IFN-γ promotes differentiation of T cells to the helper T cell subtype 1 and also enhances B-cell isotype switching to immunoglobulin G.158

Receptors of all IFN subtypes belong to the class II of cytokine receptors and utilize JAK-STAT signaling pathway for nuclear signaling, although different STAT activation (e.g., STAT1 and STAT2) is favored by individual receptors.

**Granulocyte-Macrophage Colony-Stimulating Factor/Interleukin-3/Interleukin-5 Family.** Granulocyte-macrophage colony-stimulating factor (GM-CSF), IL-3, and IL-5 compose a small family of cytokines that regulates the growth and activation of immune cells. They are largely the products of activated T cells, which when released, stimulate the behavior of myeloid cells by inducing cytokine expression and antigen presentation. In this way, GM-CSF, IL-3, and IL-5 are able to link the innate and acquired immune responses. With the exception of eosinophils, GM-CSF/IL-3/IL-5 are not essential for constitutive hematopoietic cell function. Rather, they play an important role when the host is stressed by serving to increase the numbers of activated and sensitized cells required to bolster host defense.161 Currently, GM-CSF is in clinical trials for administration to children with an injury severity score >10 following blunt or penetrating trauma. The goal of the study is to provide evidence of the effectiveness of GM-CSF as an agent that can ameliorate posttraumatic immune suppression.

Receptors for the GM-CSF/IL-3/IL-5 family of cytokines are expressed at very low level on hematopoietic cells. Similar to the other cytokine receptors discussed, they are heterodimers composed of a cytokine-specific α subunit and a common β subunit (βc), which is shared by all three receptors and is required for high affinity signal transduction. The binding of cytokine to its receptor activates JAK2-STAT, MAPK, and PI3-K—mediated signaling events to regulate a variety of important cell behaviors, including effector function in mature cells.

**Eicosanoids**

**Omega-6 Polyunsaturated Fat Metabolites: Arachidonic Acid.** Eicosanoids are derived primarily by oxidation of the membrane phospholipid, arachidonic acid (AA), which is relatively abundant in the membrane lipids of inflammatory cells. The major precursor of arachidonic acid is the omega-6 (n-6) polyunsaturated fatty acid (PUFA) linolenic acid, a major source of which is soybean oil. Not surprisingly, an excess of linolenic acid is thought to promote inflammation via increased availability of AA, and in turn, eicosanoids.

Eicosanoids generated from AA include prostaglandins, thromboxanes, and leukotrienes. When a cell senses the proper stimulus, AA is released from phospholipids or diacylglycerols by the enzymatic activation of phospholipase A2 (Fig. 2-6A). Prostanoids, which include all of the prostaglandins (PG) and the thromboxanes, result from the sequential action of the cyclooxygenase (COX) enzyme and terminal synthetases on arachidonic acid. In contrast, arachidonic acid may be oxidized along the lipoxygenase pathway via the central enzyme 5-lipoxygenase, to produce several classes of leukotrienes and lipoxins, which have anti-inflammatory functions. In general, the effects of eicosanoids are mediated via specific receptors, which are members of a superfamily of G protein-coupled receptors.

Eicosanoids are not stored within cells but are instead generated rapidly in response to many proinflammatory stimuli, including hypoxic injury, direct tissue injury, endotoxin (lipopolysaccharide), norepinephrine, vasopressin, angiotensin II, bradykinin, serotonin, acetylcholine, cytokines, and histamine. They have a broad range of physiologic roles, including neurotransmission, and vasomotor regulation. Eicosanoids are also involved in immune cell regulation (Table 2-6), by modulating the intensity and duration of inflammatory responses.

Glucocorticoids, NSAIDs, and leukotriene inhibitors can successfully block the end products of eicosanoid pathways to modulated inflammation.

The production of eicosanoids is cell- and stimulus-specific. Therefore, the signaling events that are initiated will depend on the concentrations and types of eicosanoids generated, as well as the unique complement of receptors expressed by their target cells. For example, prostaglandin E2 (PGE2) suppresses the effector function of macrophages (i.e., phagocytosis and intracellular pathogen killing) via a mechanism that is dependent on increased cAMP levels. PGE2 also modulates chemokine production and enhances local accumulation of regulatory T cells and myeloid-derived suppressor cells. Prostacyclin (PGI2) has an inhibitory effect on Th1 and Th2-mediated immune responses, while enhancing Th17 differentiation and cytokine production. Leukotrienes are potent mediators of capillary leakage as well as leukocyte adherence, neutrophil activation, bronchoconstriction, and vasoconstriction. Leukotriene B4 is synthesized from arachidonic acid in response to acute Ca2+ signaling induced by inflammatory mediators.162 High affinity leukotriene receptors (BLT1) are expressed primarily in leukocytes, including granulocytes, eosinophils, macrophages, and differentiated T cells, whereas the low affinity receptor is expressed in many cell types. Leukotrienes, most notably leukotriene B4 (LTB4), has been implicated in the development of both acute lung injury and acute kidney injury following hemorrhagic shock in animal models.163,164
Omega-3 Polyunsaturated Fat Metabolites: All-cis-5, 8, 11, 14, 17-Eicosapentaenoic Acid. The second major family of PUFAs is the omega-3 fatty acid, α-linolenic acid, which is found primarily in cold water fish. α-Linolenic acid is the metabolic precursor of eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA). Omega-3 PUFAs are also substrates for the cyclooxygenase and lipoxygenase enzymes that produce eicosanoids, but the mediators produced have a different structure from the AA-derived mediators, and this influences their actions (Fig. 2-6B). For example, omega-3 fatty acids are reported to have specific anti-inflammatory effects, including inhibition of NF-κB activity, TNF release from hepatic Kupffer cells, and leukocyte adhesion and migration. Key derivatives of omega-3 PUFAs have also been identified and synthesized. These include resolvins, protectins, and maresins. In a variety of model systems, resolvins have been shown to attenuate the inflammatory phenotypes of a number of immune cells by decreasing neutrophil recruitment, reducing synthesis of pro-inflammatory cytokines and regulating transcription factor activation.165,166

The ratio of dietary omega-6 to omega-3 PUFA is reflected in the membrane composition of various cells, including cells of the immune system, which has potential implications for the inflammatory response. For example, a diet that is rich in omega-6 PUFA will result in cells whose membranes are “omega-6 PUFA rich.” When omega-6 PUFAs are the main plasma membrane lipid available for phospholipase activity, more proinflammatory PUFAs (i.e., 2-series prostaglandins) are generated. Many lipid preparations are soy-based and thus primarily composed of omega-6 fatty acids. These are thought to be “inflammation-enhancing.” Nutritional supplementation with omega-3 fatty acid has the potential to dampen inflammation by shifting the cell membrane composition in favor of omega 3-PUFAs. In a study of surgical patients, preoperative supplementation with omega-3 fatty acid was associated with reduced need for mechanical ventilation, decreased hospital length of stay, and decreased mortality with a good safety profile.167

Plasma Contact System

Complement. Following traumatic injury, there is almost immediate activation of the complement system, which is a major effector mechanism of the innate immune system. The complement system was thought to act initially as the required “first line of defense” for the host against pathogens, by binding and clearing them from the circulation. Recent data indicate...
Systemic stimulatory and inhibitory actions of eicosanoids

<table>
<thead>
<tr>
<th>ORGAN/FUNCTION</th>
<th>STIMULATOR</th>
<th>INHIBITOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pancreas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucose-stimulated insulin secretion</td>
<td>12-HPETE</td>
<td>PGE₂</td>
</tr>
<tr>
<td>Glucagon secretion</td>
<td>PGD₂, PGE₂</td>
<td></td>
</tr>
<tr>
<td>Liver</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucagon-stimulated glucose production</td>
<td>PGE₂</td>
<td></td>
</tr>
<tr>
<td>Fat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hormone-stimulated lipolysis</td>
<td>PGE₂</td>
<td></td>
</tr>
<tr>
<td>Bone</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resorption</td>
<td>PGE₂, PGE-m, 6-K-PGE₁, PGF₁₀, PGI₂</td>
<td></td>
</tr>
<tr>
<td>Parathyroid</td>
<td>PGE₂</td>
<td>PGF₂</td>
</tr>
<tr>
<td>Lung</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bronchoconstriction</td>
<td>PGF₂, TXA₂, LTC₄, LTD₄, LTE₄</td>
<td>PGE₂</td>
</tr>
<tr>
<td>Kidney</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stimulation of renin secretion</td>
<td>PGE₂, PGI₂</td>
<td></td>
</tr>
<tr>
<td>Gastrointestinal system</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cytoprotective effect</td>
<td>PGE₂</td>
<td></td>
</tr>
<tr>
<td>Immune response</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suppression of lymphocyte activity</td>
<td>PGE₂</td>
<td></td>
</tr>
<tr>
<td>Hematologic system</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Platelet aggregation</td>
<td>TXA₂</td>
<td>PGI₂</td>
</tr>
</tbody>
</table>

5-HETE = 5-hydroxyeicosatetraenoic acid; 12-HPETE = 12-hydroperoxyl-HEs6-ketoprostaglandin E₂; LT = leukotriene; PG = prostaglandin; PGE-m = 13,14-dihydro-15-keto-PGE₁ (major urine metabolite of PGE₂); TXA₂ = thromboxane A₂.

that complement is also participates in the elimination of immune complexes as well as damaged and dead cells. In addition, complement is recognized as contributing to mobilization of hematopoietic stem/progenitor cells (HSPC) and lipid metabolism. Although complement activation is typically depicted as a linear process in which parallel pathways are activated, it actually functions more like a central node that is tightly networked with other systems. Then, depending on the activating signal, several initiation and regulatory events act in concert to heighten immune surveillance.

Complement activation proceeds via distinct pathways. Pathway initiation occurs by the binding and activation of a specific recognition unit to its designated ligand. The classical pathway, which is often referred to as “antibody-dependent” is initiated by direct binding of C1q to its common ligands, which include IgM/IgG aggregates. Alternately, C1q can activate complement by binding to soluble pattern recognition molecules such as pentraxins (e.g., C-reactive protein [CRP]). In a series of subsequent activation and amplification steps, the pathway ultimately leads to the generation of C3a and C5a, which are potent anaphylatoxins, in addition to C3b, which acts as an opsonin. An additional product, C5b, initiates the formation of the membrane attack complex, which becomes inserted into cell membrane to form a lytic pore. The subsequent effect of complement signaling is neutrophil activation leading to ROS production, as well as protease and vasoactive mediator secretion. The complement cascade also results in the release of proinflammatory cytokines synergistically with TLR-signaling, which contributes to systemic inflammation and generalized capillary leak following severe injury.

An additional means of complement activation via the lectin pathway is initiated by mannose-binding lectins (MBL) or ficolins, which bind specific carbohydrate structures. This pathway has been speculated to be a major mechanism for complement activation post injury by interactions with matrix fragments and mitochondrial DAMPs. In addition, both low pH and serine proteases of the coagulation cascade may contribute to complement activation. As a consequence, high levels of activated complement components may help to continue to drive systemic inflammation post injury.

Kallikrein-Kinin System. The kallikrein-kinin system, also referred to as the “contact” system, is a group of proteins that contribute to both coagulation and inflammation. Prekallikrein circulates in the plasma bound to high molecular weight kininogen (HK). A variety of stimuli lead to the binding of prekallikrein-HK complex to Hageman factor (factor XII) to initiate the intrinsic clotting cascade. This results in formation of the serine protease kallikrein, which is both proinflammatory and procoagulant. HK is cleaved by kallikrein to form bradykinin (BK).

The kinins (e.g., BK) mediate several physiologic processes, including vasodilation, increased capillary permeability, tissue edema, and neutrophil chemotaxis. They also increase renal vasodilation and consequently reduce renal perfusion pressure. Kinin receptors are members of the rhodopsin family of G-protein-coupled receptors and are located on vascular endothelium and smooth muscle cells. Kinin receptors are rapidly upregulated following TLR4 and cytokine signaling and appear to have important effects on both immune cell behavior and on immune mediators. For example, activation of the kinin receptor, B1, results in increased neutrophil chemotaxis, while increased B2 receptor expression causes activation of arachidonic-prostaglandin pathways. Bradykinin and kallikrein levels are increased following hemorrhagic shock and tissue injury. The degree of elevation in the levels of these mediators has been associated with the magnitude of injury and mortality.
Serotonin

Serotonin is a monoamine neurotransmitter (5-hydroxytryptamine; 5-HT) derived from tryptophan. Serotonin is synthesized by neurons in the CNS as well as by intestinal enterochromaffin cells, which are the major source of plasma 5-HT. Once in the plasma, 5-HT is taken up rapidly into platelets via the serotonin transporter (SERT), where it is either stored in the dense granules in millimolar concentrations or targeted for degradation. It is interesting that the surface expression of SERT on platelets is sensitive to plasma 5-HT levels, which in turn modulates platelet 5-HT content. Receptors for serotonin are widely distributed in the periphery and are found in the GI tract, cardiovascular system, and some immune cells.172 Serotonin is a potent vasoconstrictor and also modulates cardiac inotropy and chronotropy through nonadrenergic cyclic adenosine monophosphate (cAMP) pathways. Serotonin is released at sites of injury, primarily by platelets. Recent work has demonstrated an important role for platelet 5-HT in the local inflammatory response to injury. Using mice that lack the nonneuronal isoform of tryptophan hydroyxylase (Tph1), the rate-limiting step for 5-HT synthesis in the periphery, investigators demonstrated fewer neutrophils rolling on mesenteric venules.173 Tph1−/− mice, in response to an inflammatory stimulus, also showed decreased neutrophil extravasation. Together, these data indicate an important role for nonneuronal 5-HT in neutrophil recruitment to sites of inflammation and injury.

Histamine

Histamine is a short-acting endogenous amine that is widely distributed throughout the body. It is synthesized by histidine decarboxylase (HDC), which decarboxylates the amino acid histidine. Histamine is either rapidly released or stored in neurons, skin, gastric mucosa, mast cells, basophils, and platelets and plasma levels are increased with hemorrhagic shock, trauma, thermal injury, and sepsis.174 Not surprisingly, circulating cytokines can increase immune cell expression of HDC to further contribute to histamine synthesis. There are four histamine receptor (HR) subtypes with varying physiologic roles, but they are all members of the rhodopsin family of G-protein coupled receptors. H1R binding mediates vasodilation, bronchoconstriction, intestinal motility, and myocardial contractility. H1R knockout mice demonstrate significant immunologic defects, including impaired B and T cell responses.

H2R binding is best described for its stimulation of gastric parietal cell acid secretion. However, H2R can also modulate a range of immune system activities, such as mast cell degranulation, antibody synthesis, Th1 cytokine production, and T-cell proliferation. H3R was initially classified as a presynaptic autoreceptor in the peripheral and central nervous system (CNS). However, data using H3R knockout mice demonstrates that it also participates in inflammation in the CNS. H3R knockout mice display increased severity of neuroinflammatory diseases, which correlates with dysregulation of blood-brain barrier permeability and increased expression of macrophage inflammatory protein 2, IFN-inducible protein 10, and CXCR3 by peripheral T cells. H4R is expressed primarily in bone marrow, but it has also been detected in leukocytes, including neutrophils, eosinophils, mast cells, dendritic cells, T cells, and basophils. H4R is emerging as an important modulator of chemotraction and cytokine production in these cells. Thus, it is clear that cells of both the innate and adaptive immune response can be regulated by histamine, which is up-regulated following injury.175

CELLULAR RESPONSE TO INJURY

Cytokine Receptor Families and Their Signaling Pathways

Cytokines act on their target cells by binding to specific membrane receptors. These receptor families have been organized by structural motifs and include type I cytokine receptors, type II cytokine receptors, chemokine receptors, tumor necrosis factor receptors (TNFR), and transforming growth factor receptors (TGFβ). In addition, there are cytokine receptors that belong to the immunoglobulin receptor superfamilies. Several of these receptors have characteristic signaling pathways that are associated with them. These will be briefly reviewed in the following section.

JAK-STAT Signaling

A major subgroup of cytokines, comprising roughly 60 factors, bind to receptors termed type I/II cytokine receptors. Cytokines that bind these receptors include type I IFNs, IFN-γ, many interleukins (e.g., IL-6, IL-10, IL-12, and IL-13), and hematopoietic growth factors. These cytokines play essential roles in the initiation, maintenance, and modulation of innate and adaptive immunity for host defense. All type I/II cytokine receptors selectively associate with the Janus kinases (JAK1, JAK2, JAK3, TYK2), which represent a family of tyrosine kinases that mediate the signal transduction for these receptors. As such, the JAK-STAT signaling pathway is considered a central communication hub for the immune system.176

JAKs are constitutively bound to the cytokine receptors, and on ligand binding and receptor dimerization, activated JAKs phosphorylate the receptor to recruit signal transducer and activator of transcription (STAT) molecules (Fig. 2-7). Activated STAT proteins further dimerize and translocate into the nucleus where they modulate the transcription of target genes. Rather than being a strictly linear pathway, it is likely that individual cytokines activate more than one JAK-STAT combination. The molecular implications for this in terms of cytokine signaling are still being unraveled, but the development of JAK-specific inhibitors (jakinibs) is moving the field forward quickly.177 Interestingly, STAT-DNA binding can be observed within minutes of cytokine binding. STATs have also been shown to modulate gene transcription via epigenetic mechanisms. Thus, JAKs and STATs are central players in the regulation of key immune cell function, by providing a signaling platform for proinflammatory cytokines (IL-6 via JAK1 and STAT3); anti-inflammatory cytokines (IL-10 via STAT3) and integrating signals required for helper and regulatory T cell development and differentiation. The JAK/STAT pathway is inhibited by the action of phosphatase, the export of STATs from the nucleus, as well the interaction of antagonistic proteins.178 JAK/STAT signaling has also been implicated in the secondary muscle wasting that occurs with chronic, persistent inflammation.179

Suppressors of Cytokine Signaling

Suppressor of cytokine signaling (SOCS) molecules are a family of proteins that function as a negative feedback loop for types I and II cytokine receptors by terminating JAK/STAT signaling. There are currently eight family members (SOCS1-7 and CIS [cytokine-inducible SH2-containing protein]) that are associated with cytokine receptor signaling. Pattern recognition receptors, including both TLR and C-type lectin receptors, also activate SOCS.180 Interestingly, induction of SOCs proteins is also achieved through activators of JAK/STAT signaling, creating an inhibitory feedback loop through which cytokines...
Figure 2-7. The Janus kinase/signal transducer and activator of transcription (JAK/STAT) signaling pathway also requires dimerization of monomeric units. STAT molecules possess “docking” sites that allow for STAT dimerization. The STAT complexes translocate into the nucleus and serve as gene transcription factors. JAK/STAT activation occurs in response to cytokines (e.g., interleukin-6) and cell stressors, and has been found to induce cell proliferation and inflammatory function. Intracellular molecules that inhibit STAT function, known as suppressors of cytokine signaling (SOCSs), have been identified. P = phosphate.

SOCS molecules can effectively self-regulate by extinguishing their own signal. SOCS molecules can positively and negatively influence the activation of macrophages and dendritic cells and are crucial for T-cell development and differentiation. All SOCS proteins are able to regulate receptor signaling through the recruitment of proteasomal degradation components to their target proteins, whether the target is a specific receptor or an associated adaptor molecule. Once associated with the SOCS complex, target proteins are readily ubiquinated and targeted for proteasomal degradation. SOCS1 and SOCS3 can also exert an inhibitory effect on JAK-STAT signaling via their N-terminal kinase inhibitory region (KIR) domain, which acts as a pseudo-substrate for JAK.

SOCS3 has been shown to be a positive regulator of TLR4 responses in macrophages via inhibition of IL-6 receptor-mediated STAT3 activation. A deficiency of SOCS activity may render a cell hypersensitive to certain stimuli, such as inflammatory cytokines and growth hormones. Interestingly, in a murine model, SOCS knockout resulted in a lethal phenotype in part because of unregulated interferon signaling.

Chemokine Receptors Are Members of the G-Protein-Coupled Receptor Family All chemokine receptors are members of the G-protein-coupled seven transmembrane family of receptors (GPCR), which is one of the largest and most diverse of the membrane protein families. GPCRs function by detecting a wide spectrum of extracellular signals, including photons, ions, small organic molecules, and entire proteins. After ligand binding, GPCRs undergo conformational changes, causing the recruitment of heterotrimeric G proteins to the cytoplasmic surface (Fig. 2-8). Heterotrimeric G proteins are composed of three subunits, Gα, Gβ, and Gγ, each of which have numerous members, adding to the complexity of the signaling. When signaling, however, G proteins perform functionally as dimers because the signal is communicated either by the Gα subunit or the Gβγ complex. The GPCR family includes the receptors for catecholamines, bradykinins, and leukotrienes, in addition to a variety of other ligands important to the inflammatory response. In general, GPCRs can be classified according to their pharmacological properties into four main families: class A rhodopsin-like, class B secretin-like, class C metabotropic glutamate/pheromone, and frizzled receptors. As noted above, GPCR activation by ligand binding results in an extracellular domain shift, which is then transmitted to cytoplasmic portion of the receptor to facilitate coupling to its principle effector molecules, the heterotrimeric G proteins. Although there are more than 20 known Gα subunits, they have been divided into four families based on sequence similarity, which has served to define both receptor and effector coupling. These include Gαq and Gα12, which signal through the activation (Gαq) or inhibition (Gα12) of adenylate cyclase to increase or decrease cyclic adenosine monophosphate (cAMP) levels, respectively. Increased intracellular cAMP can activate gene transcription through the activity of intracellular signal transducers such as protein kinase A. The Gα subunits also include the Gq pathway, which stimulates phospholipase C-β to produce the intracellular messengers inositol triphosphate and diacylglycerol. Inositol triphosphate triggers the release of calcium from intracellular stores, while diacylglycerol recruits protein kinase C to the plasma membrane for activation. Finally, Gα12/13 appears to act through Rh and Ras-mediated signaling.

Tumor Necrosis Factor Superfamily

Signaling pathway for tumor necrosis factor receptor 1 (TNFR-1) (55 kDa) and TNFR-2 (75 kDa) occurs by the recruitment of several adapter proteins to the intracellular receptor complex. Optimal signaling activity requires receptor trimerization. TNFR-1 initially recruits TNFR-associated death domain (TRADD) and induces apoptosis through the actions of proteolytic enzymes known as caspases, a pathway shared by another receptor, CD95 (Fas). CD95 and TNFR-1 possess similar intracellular sequences known as death domains (DDs), and both recruit the same adapter proteins (Fas-associated death domains [FADDs]) before activating caspase 8. TNFR-1 also induces apoptosis by activating caspase 2 through the recruitment of receptor-interacting protein (RIP). RIP also has a functional component that can initiate nuclear factor κB (NF-κB) and c-Jun activation, both favoring cell survival and proinflammatory functions. TNFR-2 lacks a DD component but recruits adapter proteins known as TNFR-associated factors 1 and 2 (TRAF1, TRAF2) that interact with RIP to mediate NF-κB and c-Jun activation. TRAF2 also recruits additional proteins that are antiapoptotic, known as inhibitors of apoptosis proteins (IAPs).

Transforming Growth Factor-β Family of Receptors

Transforming growth factor-β1 (TGF-β1) is a pleiotropic cytokine expressed by immune cells that has potent immunoregulatory activities. Specifically, recent data indicate that TGF-β is
essential for T cell homeostasis, as mice deficient in TGF-β1 develop a multiorgan autoimmune inflammatory disease and die a few weeks after birth, an effect that is dependent upon the presence of mature T cells. The receptors for TGF-β ligands are the TGF-β superfamily of receptors, which are type I transmembrane proteins that contain intrinsic serine/threonine kinase activity. These receptors comprise two subfamilies, the type I and the type II receptors that are distinguished by the presence of a glycine/serine-rich membrane domain found in the type I receptors. Each TGF-β ligand binds a characteristic combination of type I and type II receptors, both of which are required for signaling. Whether the type I or the type II receptor binds first is ligand-dependent, and the second type I or type II receptor is then recruited to form a heteromeric signaling complex. When TGF-β binds to the TGF-βR, heterodimerization activates the receptor which then directly recruits and activates a receptor-associated Smad (Smad 2 or 3) through phosphorylation. An additional “common” Smad is then recruited. The activated Smad-complex translocates into the nucleus and, with other nuclear cofactors, regulates the transcription of target genes. TGF-β can also induce the rapid activation of the Ras-extracellular signal-regulated kinase (ERK) signaling pathway in addition to other MAPK pathways (JNK, p38MAPK). How does TGF-β inhibit immune responses? One of the most important effects is the suppression of interleukin-2 production by T cells. It also inhibits T cell proliferation. More recently, it was noted that TGF-β can regulate the maturation of differentiated dendritic cells and dendritic cell-mediated T-cell responses. Importantly, TGF-β can induce “alternative activation” macrophages, designated M2 macrophages, which express a wide array of anti-inflammatory molecules, including IL-10 and arginase1.

Figure 2-8. G-protein–coupled receptors are transmembrane proteins. The G-protein receptors respond to ligands such as adrenaline and serotonin. On ligand binding to the receptor (R), the G protein (G) undergoes a conformational change through guanosine triphosphate–guanosine diphosphate conversion and in turn activates the effector (E) component. The E component subsequently activates second messengers. The role of inositol triphosphate (IP₃) is to induce release of calcium from the endoplasmic reticulum (ER). cAMP = cyclic adenosine triphosphate.

**TRANSCRIPTIONAL AND TRANSLATIONAL REGULATION OF THE INJURY RESPONSE**

Transcriptional Events Following Blunt Trauma

Investigators have examined the transcriptional response in circulating leukocytes in a large series of patients who suffered severe blunt trauma. This work identified an overwhelming shift in the leukocyte transcriptome, with more than 80% of the cellular functions and pathways demonstrating some alteration in gene expression. In particular, changes in gene expression for pathways involved in the systemic inflammatory, innate immune, compensatory anti-inflammatory, and adaptive immune responses were simultaneous and marked. Moreover, they occurred rapidly (within 4–12 hours), and were prolonged for days and weeks. When different injuries (i.e., blunt trauma, burn injury, human model of endotoxemia) were compared, the patterns of gene expression were surprisingly similar, suggesting that the stress response to both injury and inflammation is highly conserved and may follow a universal pathway that includes common denominators. Finally, delayed clinical recovery and organ injury were not associated with a distinct pattern of transcriptional response elements. These data describe a new paradigm based on the observation of a rapid and coordinated transcriptional response to severe traumatic injury that involves both the innate and adaptive immune systems. Further, the data support the idea that individuals who are destined to die from their injuries are characterized primarily by the degree and duration of their deregulated inflammatory response rather than a “unique signature” indicative of a “second hit.”

**Transcriptional Regulation of Gene Expression**

Many genes are regulated at the point of DNA transcription and thus influence whether messenger RNA (mRNA) and its subsequent product are expressed (Fig. 2-9). Gene expression relies on the coordinated action of transcription factors and coactivators (i.e., regulatory proteins), which are complexes that bind to highly specific DNA sequences upstream of the target gene known as the promoter region. Enhancer sequences of DNA mediate gene expression, whereas repressor sequences are non-coding regions that bind proteins to inhibit gene expression. For example, nuclear factor kB (NF-kB), one of the best-described transcription factors, has a central role in regulating the gene products expressed after inflammatory stimuli (Fig. 2-10). The NF-kB family of transcription factors is composed of five members that share a common domain. They form numerous homo or heterodimers that are normally retained in the cytosol through...
the inhibitory binding of inhibitor of κB (I-κB). In response to an inflammatory stimulus (e.g., TNF, IL-1, or DAMP) a sequence of intracellular mediator phosphorylation reactions leads to the degradation of I-κB and subsequent release of NF-κB to allow nuclear translocation and the initiation of transcription.

Epigenetic Regulation of Transcription
The DNA access of protein machineries involved in transcription processes is tightly regulated by histones, which are a family of basic proteins that associate with DNA in the nucleus. Histone proteins help to condense the DNA into tightly packed nucleosomes that limit transcription. Emerging evidence indicates that transcriptional activation of many proinflammatory genes requires nucleosome remodeling, a process that is regulated by the histone modifying enzymes. There are at least seven identified chromatin modifications, including acetylation, methylation, phosphorylation, ubiquitinylation, sumoylation, ADP ribosylation, deamination, and proline isomerization. Alteration of chromatin packing in this way makes the DNA more or less accessible for transcription. Recently, the development of chromatin immunoprecipitation (ChIP) coupled to massively parallel DNA sequencing technology (ChIP-Seq) has enabled the mapping of histone modifications in living cells. In this way, it has allowed the identification of the large number of posttranslational histone modifications that are “written” and “erased” by histone-modifying enzymes. The role of histone modifications in the regulation of gene expression is referred to as “epigenetic” control.

The addition of an acetyl group to lysine residues on histones is an epigenetic mark associated with gene activation. These acetyl groups are reversibly maintained by histone acetyltransferases (HATs) and histone deacetylases (HDACs). Hemorrhagic shock has been shown to alter the acetylation of histone proteins via an imbalance of HDAC/HAT activity in the heart, lung, and liver in a tissue-specific pattern. In animal models, postshock administration of the HDAC inhibitor valproic acid improved overall survival. Valproic acid is currently in Phase 2 clinical trials for the treatment of hemorrhagic shock.

Translation Regulation of Inflammatory Gene Expression
Once mRNA transcripts are generated, they can also be regulated by a variety of mechanisms, including (a) splicing, which can cleave mRNA and remove noncoding regions; (b) capping, which modifies the 5' ends of the mRNA sequence to inhibit breakdown by exonucleases; (c) and the addition of a polyadenylated tail, which adds a noncoding sequence to the mRNA, to regulate the half-life of the transcript. Recent data has identified microRNAs (miRNAs) as important translational regulators of gene expression via their binding to partially complementary sequences in the 3'-untranslated region (3'-UTR) of target mRNA transcripts. Binding of miRNA to the
mRNA usually results in gene silencing. MicroRNAs are endogenous, single-stranded RNAs of approximately 22 nucleotides in length that are highly conserved in eukaryotes. MicroRNAs are encoded either singly or can be transcribed in a “polycistronic” clusters and produced by an elaborate expression and processing mechanism. After a primary miRNA transcript is generated by RNA polymerase II or III, it is processed in the nucleus to produce a short hairpin precursor miRNA transcript. The precursor is then transported into the cytoplasm where the final mature miRNA is generated by a protein termed Dicer. The mature double-stranded miRNA is then incorporated into the RNA-induced silencing complex (RISC) in the cytoplasm. Once programmed with a small RNA, RISC can silence targeted genes by one of several distinct mechanisms, working at (a) the level of protein synthesis through translation inhibition, (b) the transcript level through mRNA degradation, or (c) the level of the genome itself through the formation of heterochromatin or by DNA elimination. MiRNAs are involved in TLR signaling in the innate immune system by targeting multiple molecules in the TLR signaling pathways.196 Traumatic brain injury alters serum miRNA profiles that may be useful both as biomarkers for severe TBI and as therapeutic targets.189

**CELL-MEDIATED INFLAMMATORY RESPONSE**

**Neutrophils**

Neutrophils (PMNs) are among the first responders to sites of infection and injury and as such are potent mediators of acute inflammation.191 Mobilization of PMNs from the bone marrow is facilitated by reduction in bone-marrow expression of stromal cell-derived factor-1 (SDF1, also CXCL12) and subsequent expression of both SDF1 and its receptor CXCR4 in target tissues.192 This and other chemotactic mediators induce PMN adherence to the vascular endothelium and promote eventual cell migration into the injured tissue. Early signals for PMN recruitment include endogenous “self” molecules released from damaged tissues, like the DAMPs described previously and also include histone proteins as well as adenosine triphosphate. DAMP molecules can also induce secretion of powerful chemokines such as IL-8 (CXCL8), which can bind to tissue glycosaminoglycans, creating a gradient for PMN migration. PMNs generally have short half-lives (4 to 10 hours). However, inflammatory signals may promote their longevity in target tissues, which can contribute to their potential detrimental effects and subsequent bystander injury. In addition, following sterile trauma, large numbers of immature PMNs are recruited from the bone marrow into the circulation.193

Once primed and activated by inflammatory stimuli, including TNF, IL-1, and microbial pathogens, PMNs are capable of amplifying the inflammatory response as well as releasing toxic effectors such as ROS and proteolytic enzymes into the extracellular space.194 Neutrophils can also dump their granule contents into the extracellular space, and many of these proteins also have important effects on the innate and adaptive immune responses. When highly activated, neutrophils can extrude a meshwork of chromatin fibers, composed of DNA and histones that are decorated with granule contents. Termed neutrophils extracellular traps or NETs, they were first described as effective mechanism whereby neutrophils can immobilize bacteria to facilitate their killing. In the setting of tissue injury, NETs may allow continued presentation of auto-antigens to the host immune system, which can contribute to further tissue injury.195 NETs may also serve to prime T cells, making their threshold for activation lower.

Neutrophils do facilitate the recruitment of monocytes into inflamed tissues. These recruited cells are capable of phagocytosing apoptotic neutrophils to contribute to resolution of the inflammatory response and to promote tissue repair.196 However, at least some portion of the neutrophil population from the injury site is capable of reentering the bloodstream and returning to the bone marrow in a process regulated by chemokine CXC receptor 4.196 Whether reverse migration of neutrophils is beneficial to the host or likely to cause distant organ injury needs further investigation.

**Monocyte/Macrophages**

Monocytes and macrophages are mononuclear phagocytes that play a critical role in inflammation and the injury response.197 Monocytes are leukocytes derived from bone marrow progenitors that circulate in the bloodstream and given the right stimuli, exit the vasculature, and differentiate into monocyte-derived macrophages (e.g., alveolar macrophages or Kupffer cells) upon migrating into appropriate tissues. Macrophages represent the large number of phagocytes that are resident in tissues under resting conditions. Distinct from monocytes, they are derived from embryonic precursors and can repopulate their numbers either by self-renewal or from monocytes derived from the bone marrow.198 Together, monocytes/macrophages are the main effector cells that sense and respond to “danger signals,” primarily through mechanisms that include phagocytosis of cellular debris, release of inflammatory mediators, and recruitment of additional immune cells to injury sites. Moreover, these cells fulfill homeostatic roles beyond host defense by performing important functions in the remodeling of tissues, both during development and in the adult animal. SDF1 has also been implicated in the recruitment of monocytes to sites of tissue injury.199 Importantly, SDF1 forms a complex with HMGB1, a DAMP molecule, which potently increases its chemotactic function. In conjunction with CXCR4, the SDF1-HMGB1 complex induces early monocyte migration into injured tissues, where they play an important role coordinating between innate and adaptive immunity.

In tissues, mononuclear phagocytes are quiescent. However, they respond to external cues (e.g., PAMPs, DAMPs, activated lymphocytes) by changing their phenotype.200 In response to various signals, macrophages may undergo classical M1 activation (stimulated by TLR ligands and IFN-γ) or alternative M2 activation (stimulated by type II cytokines IL-4/IL-13); these states mirror the Th1–Th2 polarization of T cells described in the following section. The M1 phenotype is characterized by the expression of high levels of proinflammatory cytokines, like TNF-α, IL-1 and IL-6, in addition to the synthesis of ROS and RNS. Activated macrophages can also secrete HMGB1 and in this way, can recruit additional macrophages to form a self-activating loop.

In contrast, M2 macrophages are considered to be involved in the promotion of wound repair and the restoration of immune homeostasis through their expression of arginase-1 and IL-10, in addition to a variety of PRR (e.g., scavenging molecules).201 In truth, this classification system is overly simplistic. In fact, macrophages are highly heterogeneous and possess specialized properties that are precisely adapted to individual tissues. Thus, they are likely to also possess individualized response to local tissue damage.197
In a mouse model of hemorrhagic shock/reperfusion, macrophages play a key role in the recruitment of hematopoietic stem cells from the bone marrow by secreting granulocyte-macrophage colony-stimulating factor (GM-CSF) in response to circulating HMGB1.202 In the lung, alveolar macrophages sense DAMPs and extracellular matrix fragments via pattern recognition receptors. In response, they upregulate their expression of TLR4, which primes the cell for response against potential infection.203 At the same time, they release proinflammatory cytokines and ROS, which contribute injury to the alveolar epithelial cells. More recently, data indicate that an imbalance of M1/M2 macrophage populations in the lung contribute to acute lung injury following hemorrhagic shock (HS).204 In this study, investigators demonstrated that HS/resuscitation resulted in a significant decrease in M2 phenotype macrophages, with a delayed increase in M1. Augmenting the M2 population prior to injury lessened the degree of lung injury as assessed histologically.

**Lymphocytes and T-Cell Immunity**

The expression of genes associated with the adaptive immune response is rapidly altered following severe blunt trauma.3 In fact, significant injury is associated with adaptive immune suppression that is characterized by altered cell–mediated immunity. This correlates with both a decrease in the overall number of lymphocytes as well as the balance between the NK and T cell populations.205

CD4+ T cells (helper) play central roles in the function of the immune system through their effects on B cell antibody production, their enhancement of specific T

Reg cell functions, and their assistance with macrophage activation. CD4+ Th cells are functionally divided into subsets, which include Th1, Th2, and Th17 cells. Each of these groups produces specific effector cytokines that are under unique transcriptional control. The specific functions of these cells include the recognition and killing of intracellular pathogens (cellular immunity, Th1 cells); regulation of antibody production (humoral immunity, Th2 cells); and maintenance of mucosal immunity and barrier integrity (Th17 cells). Historically, activities have been characterized as proinflammatory (Th1) and anti-inflammatory (Th2) respectively, as determined by their distinct cytokine signatures (Fig. 2-11). Given the proinflammatory action of IL-17A produced by Th17 cells, they could also be placed in this category. However, it is clear that the Th17 differentiation is more complex and may involve the two distinct phenotypes, a pathogenic phenotype characterized by increased IL-17 production and a more regulatory phenotype in which IL-10 expression is increased.206

Recent evidence suggests that the population of Th17 cells is altered following severe traumatic injury. Mass cytometry by time-of-flight (CyTOF) was used to collect single cell phenotyping data on circulating peripheral blood mononuclear cells from a cohort of severely injured trauma patients.153 The investigators identified an expansion of Th17 cells at all time points following injury and was associated with an increase in the cytokine profile associated with a Th17 phenotype. This supports prior work also demonstrating a robust type 17 immune response early (within the first 24 hours) among nonsurvivors, which also identified a Th17 profile more consistent with “pathogenic” Th17 cells.154

Successful recovery from injury also depends upon a balanced Th1/Th2 response. Following injury, however, there is a reduction in Th1 cell differentiation and cytokine production in favor of an increased population of Th2 lymphocytes and their signaling products. As a consequence, both macrophage activation and proinflammatory cytokine synthesis are inhibited. This imbalance, which may be associated with decreased IL-12 production by activated monocytes/macrophages, has been associated with increased risk of infectious complications following surgery and trauma. What are the systemic mechanisms responsible for this shift? Several events have been implicated, including the direct effect of glucocorticoids on monocyte IL-12 production and T cell IL-12 receptor expression. In addition, sympathoadrenal catecholamine production has also been demonstrated to reduce IL-12 production and proinflammatory cytokine synthesis. Finally, more recent work has implicated circulating immature myeloid cells, termed myeloid-derived suppressor cells, that have immune suppressive activity particularly through their increased expression of arginase.208 These cells have the potential to deplete the microenvironment of arginine, leading to further T cell dysfunction.

**Dendritic Cells**

Recent studies have focused on the cellular components of the immune system in the context of polytrauma. While the
activation of granulocytes and monocyte/macrophages following trauma has been well described, more recent work has demonstrated that dendritic cells (DC) are also activated in response to damage signals, to stimulate both the innate and the adaptive immune responses. Dendritic cells are the most important antigen-presenting cells (APCs) for initiating T-cell responses against protein antigens. Primary “danger signals” that are recognized and activated by DC include debris from damaged or dying cells (e.g., HMGB1, nucleic acids including single nucleotides, and degradation products of the extracellular matrix). DC are frequently referred to as “professional APCs” since their principal function is to capture, process, and present both endogenous and exogenous antigens, which, along with their co-stimulatory molecules, are capable of inducing a primary immune response in resting naïve T lymphocytes. In addition, they have the capacity to further regulate the immune response, both positively and negatively, through the upregulation and release of immunomodulatory molecules such as the chemokine CCL5 (RANTES) and the CXC chemokine CXCL5. Finally, they have been implicated both in the induction and maintenance of immune tolerance as well as in the acquisition of immune memory. There are distinct classes and subsets of DC, which are functionally heterogeneous. Different levels of damage-sensing receptors (e.g., TLR) that dictate a preferential response to DAMPs at that site. While relatively small in number relative to the total leukocyte population, the diverse distribution of DC in virtually all body tissues underlines their potential for a collaborative role in the initiation of the trauma-induced sterile systemic inflammatory response. Data support a phenotypic alteration in these cells following traumatic injury.

Platelets
Platelets are small (2 μm), circulating fragments of a larger cell precursor, the megakaryocyte that is located chiefly within the bone marrow. Although platelets lack a nucleus, they contain both mRNA and a large number of cytoplasmic and surface proteins that equip them for diverse functionality. While their role in hemostasis is well described, more recent work suggest that platelets play a role in both local and systemic inflammatory responses, particularly following ischemia reperfusion. Platelets express functional scavenger and Toll-like receptors (TLR) that are important detectors of both pathogens and “damage”-associated molecules. At the site of tissue injury, complex interactions between platelets, endothelial cells, and circulating leukocytes facilitate cellular activation by the numerous local alarmins and immune mediators. For example, platelet-specific TLR4 activation can cause thrombocytes to bind to and activate neutrophils to extrude their DNA to form neutrophil extracellular traps or NETs, an action that facilitates the capacity of the innate immune system to trap bacteria but also leads to local endothelial cell damage.

Once activated, platelets adopt an initial proinflammatory phenotype by expressing and releasing a variety of adhesion molecules, cytokines, and other immune modulators, including high mobility group 1 protein (HMGB1), interleukin (IL)-1β, and CD40 ligand (CD40L, CD154). However, activated platelets also express large amounts of the immunosuppressive factor, transforming growth factor-β (TGF-β), that has been implicated in Treg cell homeostasis. Recently, in a large animal model of hemorrhage, TGF-β levels were shown to be significantly increased 2 hours post injury, suggesting a possible mechanism for injury-related immune dysfunction. And, while soluble CD154 was not increased following hemorrhage and traumatic brain injury in that study, in a murine model of mesenteric ischemia-reperfusion injury, platelet expression of CD40 and CD154 was linked to remote organ damage.

Mast Cells
Mast cells are important in the primary response to injury because they are located in tissues. TNF release from mast cells has been found to be crucial for neutrophil recruitment and pathogen clearance. Mast cells are also known to play an important role in the anaphylactic response to allergens. On activation from stimuli including allergen binding, infection, and trauma, mast cells produce histamine, cytokines, eicosanoids, proteases, and chemokines, which leads to vasodilatation, capillary leakage, and immunocyte recruitment. Mast cells are thought to be important cosignalng effector cells of the immune system via the release of IL-3, IL-4, IL-5, IL-6, IL-10, IL-13, and IL-14, as well as macrophage migration–inhibiting factor.

ENDOTHELIUM-MEDIATED INJURY
Vascular Endothelium
Under physiologic conditions, the vascular endothelium has important anticoagulant properties and forms a critical barrier to regulate the tissue migration of circulating cells. Following injury, endothelial cells are differentially modulated, resulting in a procoagulant shift that may lead to microthrombosis and organ injury. Recent work has associated postinjury vascular dysfunction (traumatic endotheliopathy) with circulating levels of syndecan-1, a surrogate marker for disruption of the endothelial cell glycocalyx. In a cohort of over 400 severely injured patients, higher syndecan-1 measurements correlated with ISS and plasma catecholamine levels and, ultimately, with mortality in this group. The authors’ hypothesis that the increased disruption of the endothelial glycocalyx results in endothelial cell injury and an altered phenotype resulting in a prothrombotic state that leads to microvascular thrombosis and ensuing organ dysfunction.

Neutrophil-Endothelium Interaction
The regulated inflammatory response to infection facilitates neutrophil and other immunocyte migration to compromised regions through the actions of increased vascular permeability, chemoattractants, and increased endothelial adhesion factors referred to as selectins that are elaborated on cell surfaces (Table 2-7). In response to inflammatory stimuli released from sentinel leukocytes in the tissues, including chemokines, thrombin, leukotrienes, histamine, and TNF, vascular endothelium are activated and their surface protein expression is altered. Within 10 to 20 minutes, prestored reservoirs of the adhesion molecule P-selectin are mobilized to the cell surface where it can mediate neutrophil recruitment (Fig. 2-12). After 2 hours, endothelial cell transcriptional processes provide additional surface expression of E-selectin. E-selectin and P-selectin bind P-selectin glycoprotein ligand-1 (PSGL-1) on the neutrophils to orchestrate the capture and rolling of these leukocytes and allow targeted immunocyte extravasation. Immobilized chemokines on the endothelial surface create a chemotactic gradient to further enhance immune cell recruitment. Also important are secondary leukocyte-leukocyte interactions in which PGSL-1 and L-selectin binding facilitates further leukocyte tethering.
Table 2-7

Molecules that mediate leukocyte-endothelial adhesion, categorized by family

<table>
<thead>
<tr>
<th>ADHESION MOLECULE</th>
<th>ACTION</th>
<th>ORIGIN</th>
<th>INDUCERS OF EXPRESSION</th>
<th>TARGET CELLS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Selectins</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-selectin</td>
<td>Fast rolling</td>
<td>Leukocytes</td>
<td>Native Thrombin, histamine</td>
<td>Endothelium, platelets, eosinophils</td>
</tr>
<tr>
<td>P-selectin</td>
<td>Slow rolling</td>
<td>Platelets and endothelium</td>
<td>Cytokines</td>
<td>Neutrophils, monocytes</td>
</tr>
<tr>
<td>E-selectin</td>
<td>Very slow rolling</td>
<td>Endothelium</td>
<td>Cytokines</td>
<td>Neutrophils, monocytes, lymphocytes</td>
</tr>
</tbody>
</table>

| Immunoglobulins    |        |        |                        |              |
| ICAM-1             | Firm adhesion/transmigration | Endothelium, leukocytes, fibroblasts, epithelium | Cytokines | Leukocytes |
| ICAM-2             | Firm adhesion | Endothelium, platelets | Native | Leukocytes |
| VCAM-1             | Firm adhesion/transmigration | Endothelium, platelets, leukocytes | Cytokines | Leukocytes |
| PECAM-1            | Adhesion | Endothelium | Cytokines, Monocytes | Leukocytes, lymphocytes |

| β₃-(CD18) Integrins |        |        |                        |              |
| CD18/11a           | Firm adhesion/transmigration | Leukocytes | Leukocyte activation | Endothelium |
| CD18/11b (Mac-1)   | Firm adhesion/transmigration | Neutrophils, monocytes, natural killer cells | Leukocyte activation | Endothelium |
| CD18/11c           | Adhesion | Neutrophils, monocytes, natural killer cells | Leukocyte activation | Endothelium |

| β₃-(CD29) Integrins |        |        |                        |              |
| VLA-4              | Firm adhesion/transmigration | Lymphocytes, monocytes | Leukocyte activation | Monocytes, endothelium, epithelium |

ICAM-1 = intercellular adhesion molecule-1; ICAM-2 = intercellular adhesion molecule-2; Mac-1 = macrophage antigen 1; PECAM-1 = platelet-endothelial cell adhesion molecule-1; VCAM-1 = vascular cell adhesion molecule-1; VLA-4 = very late antigen-4.

Although there are distinguishable properties among individual selectins in leukocyte rolling, effective rolling most likely involves a significant degree of functional overlap.²¹⁷

Chemokines

Chemokines are a family of small proteins (8–13 kDa) that were first identified through their chemotactic and activating effects on inflammatory cells. They are produced at high levels following nearly all forms of injury in all tissues, where they are key attractants for immune cell extravasation. There are more than 50 different chemokines and 20 chemokine receptors that have been identified. Chemokines are released from endothelial cells, mast cells, platelets, macrophages, and lymphocytes. They are soluble proteins, which when secreted, bind to glycosaminoglycans on the cell surface or in the extracellular matrix. In this way, the chemokines can form a fixed chemical gradient that promotes immune cell exit to target areas. Supporting the idea of their importance in leukocyte recruitment post injury, a subset of chemokines are elevated early following traumatic injury in both survivors and nonsurvivors.²¹⁸

Chemokines are distinguished (in general) from cytokines by virtue of their receptors, which are members of the G-protein–coupled receptor superfamily. Most chemokine receptors recognize more than one chemokine ligand leading to redundancy in chemokine signaling.

The chemokines are subdivided into families based on their amino acid sequences at their N-terminus. For example, CC chemokines contain two N-terminal cysteine residues that are immediately adjacent (hence the “C-C” designation) while the N-terminal cysteines in CXC chemokines are separated by a single amino acid. The CXC chemokines are particularly important for neutrophil (PMN) proinflammatory function. Members of the CXC chemokine family, which include IL-8 (CXCL8), induce neutrophil migration and secretion of cytotoxic granular contents and metabolites. Additional chemokine families include the C- and CX3C-chemokines.²¹⁶

Recent studies support the idea that a subset of chemokines, monokine induced by γ-interferon (MIG), monocyte chemotactic protein 1 (MCP-1), and interferon γ-induced proteins 10 (IP-10) may work in concert to regulate the inflammatory response post injury and may serve as biomarkers for clinical outcome in trauma patients.²¹⁸,²¹⁹ These investigators propose that MIG, MCP-1, and IP-10 function as a “chemokine switch” in which the relative levels of each chemokine may promote its own expression, which suppresses the expression of the other two according to the severity and type of injury.²²⁰ In this way, the authors propose that the balance between these three chemokines, by regulating inflammatory mediator production (e.g., IL-6) may help to correlate with long-term outcomes.
Nitric Oxide

Nitric oxide (NO) was initially known as endothelium-derived relaxing factor due to its effect on vascular smooth muscle. Normal vascular smooth muscle cell relaxation is maintained by a constant output of NO that is regulated in the endothelium by both flow- and receptor-mediated events. NO can also reduce microthrombosis by reducing platelet adhesion and aggregation (Fig. 2-13) and interfering with leukocyte adhesion to the endothelium. NO easily traverses cell membranes and has a short half-life of a few seconds. Endogenous NO formation is derived largely from the action of NO synthase (NOS), which is constitutively expressed in endothelial cells (NOS3, eNOS). Nitric oxide synthase generates NO by catalyzing the degradation of L-arginine to L-citrulline and NO, in the presence of oxygen and NADPH. There are two additional isoforms of NOS: neuronal NOS (NOS1, nNOS) and inducible NOS (iNOS/NOS2), which is expressed in response to cytokines and bacterial products. The vasodilatory effects of NO are mediated by guanylyl cyclase, an enzyme that is found in vascular smooth muscle cells and most other cells of the body. When NO is formed by endothelium, it rapidly diffuses into adjacent cells where it binds to and activates guanylyl cyclase. This enzyme catalyzes the dephosphorylation of GTP to cGMP, which serves as a second messenger for many important cellular functions, particularly for signaling smooth muscle relaxation.

NO synthesis is increased due to the upregulation of iNOS expression in response to proinflammatory mediators such as TNF-α, and IL-1β, as well as microbial products. In fact, studies in both animal models and humans have shown that severe systemic injury and associated hemorrhage produce an early upregulation of iNOS in the liver, lung, spleen, and vascular system. In these circumstances, NO is reported to function as an immunoregulator, which is capable of modulating cytokine production and immune cell development. In particular, the formation of S-nitrosothiols, which can serve as a molecular switch to regulate protein functions, may explain many signaling effects of both iNOS- and eNOS-derived NO in the immune system with regard to T-cell activation and signaling through the T cell receptor. In T cells, NO effects have been implicated in the regulation of the immune synapse as well as the regulation of mitochondrial bioenergetics indicating that NO may play an important role as a link between innate and adaptive immunity.

Inhibition of NO production seemed initially to be a promising strategy in patients with severe sepsis. However, a randomized clinical trial in patients with septic shock determined that treatment with a nonselective NOS inhibitor was associated with an increase in mortality compared with placebo. More recent data utilizing an ovine model of peritonitis demonstrated that selective iNOS inhibition reduced pulmonary artery hypertension and gas exchange impairment and promoted higher visceral organ blood flow, coinciding with lower plasma cytokine concentrations. These data suggest that specific targeting of iNOS in the setting of sepsis may remain a viable therapeutic option.

Recent work using an animal model of traumatic brain injury (TBI) showed that acute TBI results in endothelial dysfunction in a remote vascular bed. The investigators linked the effect of TBI with impaired nitric oxide (NO) production and also with an increase in arterial arginase activity, implicating the depletion of L-arginine by arginase with the decreased NO production.

Prostacyclin

The immune effects of prostacyclin (PGI2) have been discussed previously. The best-described effects of PGI2 are in the cardiovascular system, however, where it is produced by vascular endothelial cells. Prostacyclin is a potent vasodilator that also inhibits platelet aggregation. In the pulmonary system, PGI2 reduces pulmonary blood pressure as well as bronchial hyperresponsiveness. In the kidneys, PGI2 modulates renal blood flow and glomerular filtration rate. Prostacyclin acts through its receptor (a G-protein–coupled receptor of the rhodopsin family) to stimulate the enzyme, adenylate cyclase, allowing the synthesis of cyclic adenosine monophosphate (cAMP) from ATP.
This leads to a cAMP-mediated decrease in intracellular calcium and subsequent smooth muscle relaxation.

During systemic inflammation, endothelial prostacyclin expression is impaired, and thus the endothelium favors a more procoagulant profile. Exogenous prostacyclin analogues, both intravenous and inhaled, have been utilized to improve oxygenation in patients with acute lung injury. Early clinical studies with prostacyclin have delivered some encouraging results. However, a recent study examining the administration of epoprostenol in the setting of severe injury and TBI demonstrated and attenuation of the inflammatory response as measured by serologic markers had no effect on long-term outcome.

**Endothelins**

Endothelins (ETs) are potent mediators of vasoconstriction ET-1, synthesized primarily by endothelial cells, is the most potent endogenous vasoconstrictor, and is estimated to be 10 times more potent than angiotensin II. ET release is upregulated in response to hypotension, LPS, injury, thrombin, TGF-β, IL-1, angiotensin II, vasopressin, catecholamines, and anoxia. ETs release is transcriptionally regulated and occurs at the abluminal side of endothelial cells. Very little is stored in cells; thus, a plasma increase in ET is associated with a marked increase in production. Three endothelin receptors have been identified and function via the G-protein–coupled receptor mechanism. ET_B receptors are associated with increased NO and prostacyclin production, which may serve as a feedback mechanism. Atrial ET_A receptor activation has been associated with increased inotropy and chronotropy. ET-1 infusion is associated with increased pulmonary vascular resistance and pulmonary edema and may contribute to pulmonary abnormalities during sepsis. At low levels, in conjunction with NO, ETs regulate vascular tone. However, at increased concentrations, ETs can disrupt the normal blood flow and distribution and may compromise oxygen delivery to the tissue. Recent data links endothelin expression in pulmonary vasculature with persistent inflammation associated with the development of pulmonary hypertension. Endothelin expression is linked to posttranslational and transcriptional initiation of the unfolded protein response in the affected cells, which results in the production of inflammatory cytokines. Persistent endothelin-1 stimulation may play a role in decreased vascular reactivity that is evident following hemorrhagic shock.

**Platelet Activating Factor**

Phosphotidylcholine is a major lipid constituent of the plasma membrane. Its enzymatic processing by cytosolic phospholipase A 2 (cPLA 2) or calcium-independent phospholipase A 2 (iPLA 2) generates powerful small lipid molecules, which function as intracellular second messengers. One of these is arachidonic acid, the precursor molecule for eicosanoids. Another is platelet-activating factor (PAF). During acute inflammation, PAF is released by immune cells following the activation of PLA2. The receptor for PAF (PAFR), which is constitutively expressed by platelets, leukocytes, and endothelial cells, is a G-protein–coupled receptor of the rhodopsin family. Ligand binding to the PAFR promotes the activation and aggregation of platelets and leukocytes, leukocyte adherence, motility, chemotaxis, and invasion, as well as ROS generation. Additionally, PAF activation of human PMNs induces extrusion of neutrophil extracellular traps (NETs), while platelet activation induces IL-1 via a novel posttranscriptional mechanism. Finally, PAFR ligation results not only in the upregulation of numerous proinflammatory genes, including COX-2, iNOS, and IL-6, but also in the generation of lipid intermediates such as arachidonic acid and lysophospholipids through the
activation of Phospholipase A2. Antagonists to PAF receptors have been experimentally shown to mitigate the effects of ischemia and reperfusion injury. Of note, human sepsis is associated with a reduction in the levels of PAF-acetylhydrolase, which inactivates PAF by removing an acetyl group. Indeed, PAF-acetylhydrolase administration in patients with severe sepsis has yielded some reduction in multiple organ dysfunction and mortality; however, larger phase III clinical trials failed to show benefit.

Natriuretic Peptides

The natriuretic peptides, atrial natriuretic factor (ANP) and brain natriuretic peptide (BNP), are a family of peptides that are released primarily by atrial and ventricular tissue respectively, but are also synthesized by the gut, kidney, brain, adrenal glands, and endothelium. The functionally active forms of the peptides are C-terminal fragments of a larger pro-hormone, and both N- and C-terminal fragments are detectable in the blood (referred to as N-terminal pro-BNP and pro-ANF, respectively). ANF and BNP share most biological properties, including diuretic, natriuretic, vasorelaxant, and cardiac remodeling properties that are affected by signaling through a common receptor: the guanylyl cyclase- (GC-) A receptor. They are both increased in the setting of cardiac disorders; however, evidence indicates some distinctions in the setting of inflammation. For example, elevated proBNP has been detected in septic patients in the absence of myocardial dysfunction and appears to have prognostic significance. More recently, investigators examined changes in N-terminal pro-BNP (NT-proBNP) in a cohort of severely injured patients and determined that persistently high level of NT-proBNP in major trauma patients is indicative of poor outcome.

**SURGICAL METABOLISM**

The initial hours after surgical or traumatic injury are metabolically associated with a reduced total body energy expenditure and urinary nitrogen wasting. With adequate resuscitation and stabilization of the injured patient, a reprioritization of substrate use ensues to preserve vital organ function and to support repair of injured tissue. This phase of recovery also is characterized by functions that participate in the restoration of homeostasis, such as augmented metabolic rates and oxygen consumption, enzymatic preference for readily oxidizable substrates such as glucose, and stimulation of the immune system. Understanding of the collective alterations in amino acid (protein), carbohydrate, and lipid metabolism characteristic of the surgical patient lays the foundation upon which metabolic and nutritional support can be implemented.

**Metabolism During Fasting**

Fuel metabolism during unstressed fasting states has historically served as the standard to which metabolic alterations after acute injury and critical illness are compared (Fig. 2-14). To maintain basal metabolic needs (i.e., at rest and fasting), a normal healthy adult requires approximately 22 to 25 kcal/kg per day drawn from carbohydrate, lipid, and protein sources. This requirement can substantially increase during severe stress states, such as those seen in patients with burn injuries.

In the healthy adult, principal sources of fuel during short-term fasting (<5 days) are derived from muscle protein and body fat, with fat being the most abundant source of energy (Table 2-8). The normal adult body contains 300 to 400 g of carbohydrates in the form of glycogen, of which 75 to 100 g are stored in the liver. Approximately 200 to 250 g of glycogen are stored within skeletal, cardiac, and smooth muscle cells. The greater glycogen stores within the muscle are not readily available for systemic use due to a deficiency in glucose-6-phosphatase but are available for the energy needs of muscle cells. Therefore, in the fasting state, hepatic glycogen stores are rapidly and preferentially depleted, which results in a fall of serum glucose concentration within hours (<16 hours).

During fasting, a healthy 70-kg adult will utilize 180 g of glucose per day to support the metabolism of obligate glycolytic cells such as neurons, leukocytes, erythrocytes, and the renal medullae. Other tissues that use glucose for fuel are skeletal muscle, intestinal mucosa, fetal tissues, and solid tumors.
Glucagon, norepinephrine, vasopressin, and angiotensin II can promote the utilization of glycogen stores (glycogenolysis) during fasting. Although glucagon, epinephrine, and cortisol directly promote gluconeogenesis, epinephrine and cortisol also promote pyruvate shunting to the liver for gluconeogenesis. Precursors for hepatic gluconeogenesis include lactate, glycerol, and amino acids such as alanine and glutamine. Lactate is released by glycolysis within skeletal muscles, as well as by erythrocytes and leukocytes. The recycling of lactate and pyruvate for gluconeogenesis is commonly referred to as the Cori cycle, which can provide up to 40% of plasma glucose during starvation (Fig. 2-15).

Lactate production from skeletal muscle is insufficient to maintain systemic glucose needs during short-term fasting (simple starvation). Therefore, significant amounts of protein must be degraded daily (75 g/d for a 70-kg adult) to provide the amino acid substrate for hepatic gluconeogenesis. Proteolysis during starvation, which results primarily from decreased insulin and increased cortisol release, is associated with elevated urinary nitrogen excretion from the normal 7 to 10 g/d up to 30 g or more per day. Although proteolysis during starvation occurs mainly within skeletal muscles, protein degradation in solid organs also occurs.

In prolonged starvation, systemic proteolysis is reduced to approximately 20 g/d and urinary nitrogen excretion stabilizes at 2 to 5 g/d (Fig. 2-16). This reduction in proteolysis reflects the adaptation by vital organs (e.g., myocardium, brain, renal cortex, and skeletal muscle) to using ketone bodies as their principal fuel source. In extended fasting, ketone bodies become an important fuel source for the brain after 2 days and gradually become the principal fuel source by 24 days.

Enhanced deamination of amino acids for gluconeogenesis during starvation consequently increases renal excretion of ammonium ions. The kidneys also participate in gluconeogenesis by the use of glutamine and glutamate and can become the primary source of gluconeogenesis during prolonged starvation, accounting for up to one-half of systemic glucose production.

**Table 2-8**

<table>
<thead>
<tr>
<th>A. COMPONENT</th>
<th>MASS (kg)</th>
<th>ENERGY (kcal)</th>
<th>DAYS AVAILABLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water and minerals</td>
<td>49</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Protein</td>
<td>6.0</td>
<td>24,000</td>
<td>13.0</td>
</tr>
<tr>
<td>Glycogen</td>
<td>0.2</td>
<td>800</td>
<td>0.4</td>
</tr>
<tr>
<td>Fat</td>
<td>15.0</td>
<td>140,000</td>
<td>78.0</td>
</tr>
<tr>
<td>Total</td>
<td>70.2</td>
<td>164,800</td>
<td>91.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. SUBSTRATE</th>
<th>O₂ CONSUMED (L/g)</th>
<th>CO₂ PRODUCED (L/g)</th>
<th>RESPIRATORY QUOTIENT</th>
<th>kcal/g</th>
<th>RECOMMENDED DAILY REQUIREMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>0.75</td>
<td>0.75</td>
<td>1.0</td>
<td>4.0</td>
<td>7.2 g/kg per day</td>
</tr>
<tr>
<td>Dextrose</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>3.4</td>
<td>—</td>
</tr>
<tr>
<td>Lipid</td>
<td>2.0</td>
<td>1.4</td>
<td>0.7</td>
<td>9.0</td>
<td>1.0 g/kg per day</td>
</tr>
<tr>
<td>Protein</td>
<td>1.0</td>
<td>0.8</td>
<td>0.8</td>
<td>4.0</td>
<td>0.8 g/kg per day</td>
</tr>
</tbody>
</table>

**Figure 2-15.** The recycling of peripheral lactate and pyruvate for hepatic gluconeogenesis is accomplished by the Cori cycle. Alanine within skeletal muscles can also be used as a precursor for hepatic gluconeogenesis. During starvation, such fatty acid provides fuel sources for basal hepatic enzymatic function. RBC = red blood cell; WBC = white blood cell.
Fuel utilization in long-term fasting man (70 kg)

Figure 2-16. Fuel utilization in extended starvation. Liver glycogen stores are depleted, and there is adaptive reduction in proteolysis as a source of fuel. The brain uses ketones for fuel. The kidneys become important participants in gluconeogenesis. RBC = red blood cell; WBC = white blood cell.

Lipid stores within adipose tissue provide 40% or more of caloric expenditure during starvation. Energy requirements for basal enzymatic and muscular functions (e.g., gluconeogenesis, neural transmission, and cardiac contraction) are met by the mobilization of triglycerides from adipose tissue. In a resting, fasting, 70-kg person, approximately 160 g of free fatty acids and glycerol can be mobilized from adipose tissue per day. Free fatty acid release is stimulated in part by a reduction in serum insulin levels and in part by the increase in circulating glucagon and catecholamine. Such free fatty acids, like ketone bodies, are used as fuel by tissues such as the heart, kidney (renal cortex), muscle, and liver. The mobilization of lipid stores for energy importantly decreases the rate of glycolysis, gluconeogenesis, and proteolysis, as well as the overall glucose requirement to sustain the host. Furthermore, ketone bodies spare glucose utilization by inhibiting the enzyme pyruvate dehydrogenase.

Metabolism After Injury

Injuries or infections induce unique neuroendocrine and immunologic responses that differentiate injury metabolism from that of unstressed fasting (Fig. 2-17). The magnitude of metabolic expenditure over time appears to be directly proportional to the severity of insult, with thermal injuries and severe infections having the highest energy demands (Fig. 2-18). Of note, the first few days following both sepsis and trauma are not hypermetabolic states, with the more severe insults associated with increased “metabolic hibernation.” However, by week 2, the total energy expenditure increases dramatically.238 The increase

Figure 2-17. Acute injury is associated with significant alterations in substrate utilization. There is enhanced nitrogen loss, indicative of catabolism. Fat remains the primary fuel source under these circumstances. RBC = red blood cell; WBC = white blood cell.
in energy expenditure is mediated in part by sympathetic activation and catecholamine release, which has been replicated by the administration of catecholamines to healthy human subjects. Lipid metabolism after injury is intentionally discussed first because this macronutrient becomes the primary source of energy during stressed states.239

**Lipid Metabolism After Injury**

Lipids are not merely nonprotein, noncarbohydrate fuel sources that minimize protein catabolism in the injured patient. Lipid metabolism potentially influences the structural integrity of cell membranes as well as the immune response during systemic inflammation. Adipose stores within the body (triglycerides) are the predominant energy source (50% to 80%) during critical illness and after injury. Fat mobilization (lipolysis) occurs mainly in response to catecholamine stimulus of the hormone-sensitive triglyceride lipase. Other hormonal influences that potentiate lipolysis include adrenocorticotropic hormone (ACTH), catecholamines, thyroid hormone, cortisol, glucagon, growth hormone release, and reduction in insulin levels.240

**Lipid Absorption.** Although the process is poorly understood, adipose tissue provides fuel for the host in the form of free fatty acids and glycerol during critical illness and injury. Oxidation of 1 g of fat yields approximately 9 kcal of energy. Although the liver is capable of synthesizing triglycerides from carbohydrates and amino acids, dietary and exogenous sources provide the major source of triglycerides. Dietary lipids are not readily absorbable in the gut but require pancreatic lipase and phospholipase within the duodenum to hydrolyze the triglycerides into free fatty acids and monoglycerides. The free fatty acids and monoglycerides are then readily absorbed by gut enterocytes, which resynthesize triglycerides by esterification of the monoglycerides with fatty acyl coenzyme A (acyl-CoA) (Fig. 2-19). Long-chain triglycerides (LCTs), defined as those with 12 carbons or more, generally undergo this process of esterification

---

**Figure 2-18.** Influence of injury severity on resting metabolism (resting energy expenditure, or REE). The shaded area indicates normal REE. (Reproduced with permission from Long CL, Schaffel N, Geiger JW, et al: Metabolic response to injury and illness: estimation of energy and protein needs from indirect calorimetry and nitrogen balance, JPEN J Parenter Enteral Nutr. 1979 Nov-Dec;3(6):452-456.)

**Figure 2-19.** Pancreatic lipase within the small intestinal brush borders hydrolyzes triglycerides into monoglycerides and fatty acids. These components readily diffuse into the gut enterocytes, where they are re-esterified into triglycerides. The resynthesized triglycerides bind carrier proteins to form chylomicrons, which are transported by the lymphatic system. Shorter triglycerides (those with <10 carbon atoms) can bypass this process and directly enter the portal circulation for transport to the liver. CoA = coenzyme A.
and enter the circulation through the lymphatic system as chylomicrons. Shorter fatty acid chains directly enter the portal circulation and are transported to the liver by albumin carriers. Hepatocytes use free fatty acids as a fuel source during stress states but also can synthesize phospholipids or triglycerides (i.e., very-low-density lipoproteins) during fed states. Systemic tissue (e.g., muscle and the heart) can use chylomicrons and triglycerides as fuel by hydrolysis with lipoprotein lipase at the luminal surface of capillary endothelium. Trauma or sepsis suppresses lipoprotein lipase activity in both adipose tissue and muscle, presumably mediated by TNF.

Lipolysis and Fatty Acid Oxidation. Periods of energy demand are accompanied by free fatty acid mobilization from adipose stores. This is mediated by hormonal influences (e.g., catecholamines, ACTH, thyroid hormones, growth hormone, and glucagon) on triglyceride lipase through a cAMP pathway (Fig. 2-20). In adipose tissues, triglyceride lipase hydrolyzes triglycerides into free fatty acids and glycerol. Free fatty acids enter the capillary circulation and are transported by albumin to tissues requiring this fuel source (e.g., heart and skeletal muscle). Insulin inhibits lipolysis and favors triglyceride synthesis by augmenting lipoprotein lipase activity as well as intracellular levels of glycerol-3-phosphate. The use of glycerol for fuel depends on the availability of tissue glycerokinase, which is abundant in the liver and kidneys.

Free fatty acids absorbed by cells conjugate with acyl-CoA within the cytoplasm. The transport of fatty acyl-CoA from the outer mitochondrial membrane across the inner mitochondrial membrane occurs via the carnitine shuttle (Fig. 2-21). Medium-chain triglycerides (MCTs), defined as those 6 to 12 carbons in length, bypass the carnitine shuttle and readily cross the mitochondrial membranes. This accounts in part for the fact that MCTs are more efficiently oxidized than LCTs. Ideally, the rapid oxidation of MCTs makes them less prone to fat deposition, particularly within immune cells and the reticuloendothelial system—a common finding with lipid infusion in parenteral nutrition. However, exclusive use of MCTs as fuel in animal studies has been associated with higher metabolic demands and toxicity, as well as essential fatty acid deficiency.

Within the mitochondria, fatty acyl-CoA undergoes β-oxidation, which produces acetyl-CoA with each pass through the cycle. Each acetyl-CoA molecule subsequently enters the tricarboxylic acid (TCA) cycle for further oxidation to yield 12 adenosine triphosphate (ATP) molecules, carbon dioxide, and water. Excess acetyl-CoA molecules serve as precursors for ketogenesis. Unlike glucose metabolism, oxidation of fatty acids requires proportionally less oxygen and produces less carbon dioxide. This is frequently quantified as the ratio of carbon dioxide produced to oxygen consumed for the reaction and is known as the respiratory quotient (RQ). An RQ of 0.7 would imply greater fatty acid oxidation for fuel, whereas an RQ of 1 indicates greater carbohydrate oxidation (overfeeding). An RQ of 0.85 suggests the oxidation of equal amounts of fatty acids and glucose.

Ketogenesis
Carbohydrate depletion slows the entry of acetyl-CoA into the TCA cycle secondary to depleted TCA intermediates and enzyme activity. Increased lipolysis and reduced systemic carbohydrate availability during starvation diverts excess acetyl-CoA.

**Figure 2-20.** Fat mobilization in adipose tissue. Triglyceride lipase activation by hormonal stimulation of adipose cells occurs through the cyclic adenosine monophosphate (cAMP) pathway. Triglycerides are serially hydrolyzed with resultant free fatty acid (FFA) release at every step. The FFAs diffuse readily into the capillary bed for transport. Tissues with glycerokinase can use glycerol for fuel by forming glycerol-3-phosphate. Glycerol-3-phosphate can esterify with FFAs to form triglycerides or can be used as a precursor for renal and hepatic gluconeogenesis. Skeletal muscle and adipose cells have little glycerokinase and thus do not use glycerol for fuel.
Figure 2-21. Free fatty acids (FFAs) in the cells form fatty acylcoenzyme A (CoA) with CoA. Fatty acyl-CoA cannot enter the inner mitochondrial membrane and requires carnitine as a carrier protein (carnitine shuttle). Once inside the mitochondria, carnitine dissociates, and fatty acyl-CoA is reformed. The carnitine molecule is transported back into the cytosol for reuse. The fatty acyl-CoA undergoes β-oxidation to form acetyl-CoA for entry into the tricarboxylic acid cycle. “R” represents a part of the acyl group of acyl-CoA.

toward hepatic ketogenesis. A number of extrahepatic tissues, but not the liver itself, are capable of using ketones for fuel. Ketosis represents a state in which hepatic ketone production exceeds extrahepatic ketone utilization.

The rate of ketogenesis appears to be inversely related to the severity of injury. Major trauma, severe shock, and sepsis attenuate ketogenesis by increasing insulin levels and by causing rapid tissue oxidation of free fatty acids. Minor injuries and infections are associated with modest elevations in plasma free fatty acid concentrations and ketogenesis. However, in minor stress states ketogenesis does not exceed that in nonstressed starvation.

Carbohydrate Metabolism

Ingested and enteral carbohydrates are primarily digested in the small intestine, where pancreatic and intestinal enzymes reduce the complex carbohydrates to dimeric units. Disaccharidases (e.g., sucrase, lactase, and maltase) within intestinal brush borders dismantle the complex carbohydrates into simple hexose units, which are transported into the intestinal mucosa. Glucose and galactose are primarily absorbed by energy-dependent active transport coupled to the sodium pump. Fructose absorption, however, occurs by concentration-dependent facilitated diffusion. Neither fructose and galactose within the circulation nor exogenous mannitol (for neurologic injury) evokes an insulin response. Intravenous administration of low-dose fructose in fasting humans has been associated with nitrogen conservation, but the clinical utility of fructose administration in human injury remains to be demonstrated.

Discussion of carbohydrate metabolism primarily refers to the utilization of glucose. The oxidation of 1 g of carbohydrate yields 4 kcal, but sugar solutions such as those found in intravenous fluids or parenteral nutrition provide only 3.4 kcal/g of dextrose. In starvation, glucose production occurs at the expense of protein stores (i.e., skeletal muscle). Hence, the primary goal for maintenance glucose administration in surgical patients is to minimize muscle wasting. The exogenous administration of small amounts of glucose (approximately 50 g/d) facilitates fat entry into the TCA cycle and reduces ketosis. Unlike in starvation in healthy subjects, in septic and trauma patients provision of exogenous glucose never has been shown to fully suppress amino acid degradation for gluconeogenesis. This suggests that during periods of stress, other hormonal and proinflammatory mediators have a profound influence on the rate of protein degradation and that some degree of muscle wasting is inevitable. The administration of insulin, however, has been shown to reverse protein catabolism during severe stress by stimulating protein synthesis in skeletal muscles and by inhibiting hepatocyte protein degradation. Insulin also stimulates the incorporation of elemental precursors into nucleic acids in association with RNA synthesis in muscle cells.

In cells, glucose is phosphorylated to form glucose-6-phosphate. Glucose-6-phosphate can be polymerized during glycogenesis or catabolized in glycogenolysis. Glucose catabolism occurs by cleavage to pyruvate or lactate (pyruvic acid pathway) or by decarboxylation to pentoses (pentose shunt) (Fig. 2-22).

Excess glucose from overfeeding, as reflected by RQs >1.0, can result in conditions such as glucosuria, thermogenesis, and conversion to fat (lipogenesis). Excessive glucose administration results in elevated carbon dioxide production, which may be deleterious in patients with suboptimal pulmonary function, as well as hyperglycemia, which may contribute to infectious risk and immune suppression.

Injury and severe infections acutely induce a state of peripheral glucose intolerance, despite ample insulin production at levels several-fold above baseline. This may occur in part due to reduced skeletal muscle pyruvate dehydrogenase activity after injury, which diminishes the conversion of pyruvate to acetyl-CoA and subsequent entry into the TCA cycle. The three-carbon structures (e.g., pyruvate and lactate) that consequently accumulate are shunted to the liver as substrate for gluconeogenesis. Furthermore, regional tissue catheterization and isotope dilution studies have shown an increase in net splanchnic glucose production by 50% to 60% in septic patients and a 50% to 100% increase in burn patients. The increase in plasma glucose levels is proportional to the severity of injury, and this net hepatic gluconeogenic response is believed to be under the influence of glucagon. Unlike in the nonstressed subject, in the hypermetabolic, critically ill patient the hepatic gluconeogenic response to injury or sepsis cannot be suppressed by exogenous or excess glucose administration but rather persists. Hepatic gluconeogenesis, arising primarily from alanine and glutamine catabolism, provides a ready fuel source for tissues such as those of the nervous system, wounds, and erythrocytes, which do not require insulin for glucose transport. The elevated glucose concentrations also provide a necessary energy source for leukocytes in inflamed tissues and in sites of microbial invasions.
The shunting of glucose away from nonessential organs such as skeletal muscle and adipose tissues is mediated by catecholamines. Experiments with infusing catecholamines and glucagon in animals have demonstrated elevated plasma glucose levels as a result of increased hepatic gluconeogenesis and peripheral insulin resistance. Interestingly, although glucocorticoid infusion alone does not increase glucose levels, it does prolong and augment the hyperglycemic effects of catecholamines and glucagon when glucocorticoid is administered concurrently with the latter.

Glycogen stores within skeletal muscles can be mobilized by epinephrine activation of β-adrenergic receptors, GTP-binding proteins (G-proteins), which subsequently activates the second messenger, cAMP. The cAMP activates phosphorylase kinase, which in turn leads to conversion of glycogen to glucose-1-phosphate. Phosphorylase kinase also can be activated by the second messenger, calcium, through the breakdown of phosphatidylinositol phosphate, which is the case in vasopressin-mediated hepatic glycogenolysis.

Protein and Amino Acid Metabolism

The average protein intake in healthy young adults ranges from 80 to 120 g/d, and every 6 g of protein yields approximately 1 g of nitrogen. The degradation of 1 g of protein yields approximately 4 kcal of energy, similar to the yield in carbohydrate metabolism. After injury, the initial systemic proteolysis, mediated primarily by glucocorticoids, increases urinary nitrogen excretion to levels in excess of 30 g/d, which roughly corresponds to a loss in lean body mass of 1.5% per day. An injured individual who does not receive nutrition for 10 days can theoretically lose 15% lean body mass. Therefore, amino acids cannot be considered a long-term fuel reserve, and indeed excessive protein depletion (i.e., 25% to 30% of lean body weight) is not compatible with sustaining life.

Protein catabolism after injury provides substrates for gluconeogenesis and for the synthesis of acute phase proteins. Radiolabeled amino acid incorporation studies and protein analyses confirm that skeletal muscles are preferentially depleted acutely after injury, whereas visceral tissues (e.g., the liver and kidney) remain relatively preserved. The accelerated urea excretion after injury also is associated with the excretion of intracellular elements such as sulfur, phosphorus, potassium, magnesium, and creatinine. Conversely, the rapid utilization of elements such as potassium and magnesium during recovery from major injury may indicate a period of tissue healing.

The net changes in protein catabolism and synthesis correspond to the severity and duration of injury (Fig. 2-23). Elective operations and minor injuries result in lower protein synthesis and moderate protein breakdown. Severe trauma, burns, and sepsis are associated with increased protein catabolism. The rise in urinary nitrogen and negative nitrogen balance can be detected early after injury and peak by 7 days. This state of protein catabolism may persist for as long as 3 to 7 weeks. The patient’s prior physical status and age appear to influence the degree of proteolysis after injury or sepsis. Activation of the ubiquitin-proteosome system in muscle cells is one of the major pathways for protein degradation during acute injury. This response is accentuated by tissue hypoxia, acidosis, insulin resistance, and elevated glucocorticoid levels.

Figure 2-22. Simplified schema of glucose catabolism through the pentose monophosphate pathway or by breakdown into pyruvate. Glucose-6-phosphate becomes an important “crossroad” for glucose metabolism.

7 Nutrition in the Surgical Patient

The goal of nutritional support in the surgical patient is to prevent or reverse the catabolic effects of disease or injury. Although several important biologic parameters have been used to measure the efficacy of nutritional regimens, the ultimate validation for nutritional support in surgical patients should be improvement in clinical outcome and restoration of function.

Estimation of Energy Requirements

All patients admitted to the hospital should have their nutritional status assessed. Overall nutritional assessment is undertaken to
determine the severity of nutrient deficiencies or excess and to aid in predicting nutritional requirements. Pertinent information is obtained by determining the presence of weight loss, chronic illnesses, or dietary habits that influence the quantity and quality of food intake. Social habits predisposing to malnutrition and the use of medications that may influence food intake or urination should also be investigated. Physical examination seeks to assess loss of muscle and adipose tissues, organ dysfunction, and subtle changes in skin, hair, or neuromuscular function reflecting frank or impending nutritional deficiency. Anthropometric data (i.e., weight change, skinfold thickness, and arm circumference muscle area) and biochemical determinations (i.e., creatinine excretion, albumin level, prealbumin level, total lymphocyte count, and transferrin level) may be used to substantiate the patient’s history and physical findings. This information, in conjunction with nutritional risk assessment scoring, can identify patients who may benefit from early nutritional support.

For critically ill and injured patients, validated scoring systems such as the Nutritional Risk Screening (NRS)\textsuperscript{245} or the Nutrition Risk in the Critically Ill (NUTRIC)\textsuperscript{246} score should be employed to make this determination and should be performed in conjunction with assessment of GI tract function and risk of aspiration. Appreciation for the stresses and natural history of the disease process, in combination with nutritional assessment, remains the basis for identifying patients in acute or anticipated need of nutritional support. Currently, specialized enteral nutrition can be avoided in patients who are deemed to be a low nutritional risk with low disease severity for up to one week. However, their nutritional status should be reassessed regularly.\textsuperscript{247}

A fundamental goal of nutritional support is to meet the energy requirements for essential metabolic processes and tissue repair. Failure to provide adequate nonprotein energy sources will lead to consumption of lean tissue stores. The requirement for energy may be measured by indirect calorimetry, which is the gold standard in hospitalized patients and is recommended for the critically ill.\textsuperscript{243} However, the use of indirect calorimetry, particularly in the critically ill patient, may not be available or feasible in this setting. Moreover, it may lead to an overestimation of caloric requirements, which has been associated with increased risk of infectious complications.\textsuperscript{248}

In the absence of indirect calorimetry, resting energy expenditure may also be estimated using a published predictive equation. Adjusted for the type of surgical stress, such equations are suitable for estimating energy requirements in the majority of hospitalized patients. Alternately, a simple weight-based equation of 25 to 30 kcal/kg (using dry or usual body weight) per day is appropriate with a low risk of overfeeding and is consistent with current recommendations from ASPEN (American Society of Parenteral and Enteral Nutrition) in the ICU setting.\textsuperscript{247} After trauma or sepsis, energy substrate demands are increased during the recovery phase and may necessitate greater nonprotein calories beyond calculated energy expenditure (Table 2-9). These additional nonprotein calories provided after injury are usually 1.2 to 2.0 times greater than calculated resting energy expenditure, depending on the type of injury. It is seldom appropriate to exceed this level of nonprotein energy intake during the height of the catabolic phase. Currently, standard enteral nutrition delivers 49% to 53% of calories as carbohydrate and 29% to 30% of calories as fat, which is consistent with current recommendations. For parenteral nutrition, dextrose-containing stock solutions are prepared and available in different concentrations. The percentage of calories that is contributed by dextrose should be determined on a per-patient basis according to the severity of injury/illness and the estimated caloric needs. Lipid emulsions can be included in the total mixture or be administered separately in 10% or 20% solutions with 1.1 kcal/ml and 2 kcal/ml, respectively.

The second objective of nutritional support is to meet the substrate requirements for protein synthesis. Protein nutritional support is especially important for maintaining immune function and lean body mass and is more closely linked to positive outcomes than total caloric intake. Although the mean protein requirement of healthy individuals is defined as 0.8 g/kg per day by the Food and Nutrition Board of the U.S. National Research Council, current recommendations for protein dosing exceed this

![Figure 2-23. The effect of injury severity on nitrogen wasting. (Reproduced with permission from Long CL, Schaffel N, Geiger JW, et al: Metabolic response to injury and illness: estimation of energy and protein needs from indirect calorimetry and nitrogen balance, JPEN J Parenter Enteral Nutr. 1979 Nov-Dec;3(6):452-456.)](image-url)
amount (1.2–2 gm protein/kg per day), especially for the critically ill and injured. Higher protein intake seems to support improved nitrogen balance and high-protein nutritional support is currently recommended for patients with body mass index >30.247

**Vitamins and Minerals**

The requirements for vitamins and essential trace minerals usually can be met easily in the average patient with an uncomplicated postoperative course. Therefore, vitamins usually are not given in the absence of preoperative deficiencies. Patients maintained on elemental diets or parenteral hyperalimentation require complete vitamin and mineral supplementation. Commercial enteral diets contain varying amounts of essential minerals and vitamins. It is necessary to ensure that adequate replacement is available in the diet or by supplementation. Numerous commercial vitamin preparations are available for intravenous or intramuscular use, although most do not contain vitamin K and some do not contain vitamin B12 or folic acid. Supplemental trace minerals may be given intravenously via commercial preparations. Essential fatty acid supplementation also may be necessary, especially in patients with depletion of adipose stores.

**Overfeeding**

Overfeeding usually results from overestimation of caloric needs, as occurs when actual body weight is used to calculate the BEE in patient populations such as the critically ill with significant fluid overload and the obese. Indirect calorimetry can be used to quantify energy requirements but frequently overestimates BEE by 10% to 15% in stressed patients, particularly if they are receiving ventilatory support. In these instances, estimated dry weight should be obtained from preinjury records or family members. Adjusted lean body weight also can be calculated. Overfeeding may contribute to clinical deterioration via increased oxygen consumption, increased carbon dioxide production and prolonged need for ventilatory support, fatty liver, suppression of leukocyte function, hyperglycemia, and increased risk of infection.

**ENTERAL NUTRITION**

**Rationale for Enteral Nutrition**

Enteral nutrition (EN) is preferred over parenteral nutrition (PN) based on the lower cost of enteral feeding and the associated risks of the intravenous route, including vascular access complications.240 Of further consideration are the consequences of gastrointestinal tract disuse, which include diminished soluble IgA production and cytokine production as well as bacterial overgrowth and altered mucosal barrier function and immune defenses. In support of this idea, recent meta-analysis demonstrated a significant reduction in infectious complications in critically ill or injured patients receiving EN when compared to PN as well as ICU length of stay.250 However, no increase in overall survival was noted. While EN is recommended as the first choice for nutritional support in patients who can tolerate it, a recent large trial from Europe comparing early isocaloric EN vs. PN in adult critically ill patients with shock did not reduce mortality or the risk of secondary infections but was associated with a greater risk of digestive complications including intestinal ischemia.251

The benefits of enteral feeding in patients undergoing elective surgery appear to be linked to their preoperative nutritional status. Historical studies comparing postoperative enteral and parenteral nutrition in patients undergoing gastrointestinal surgery have demonstrated reduced infectious complications and acute phase protein production in those fed by the enteral route. Yet prospectively randomized studies of patients with adequate nutritional status (albumin ≥4 g/dL) undergoing gastrointestinal surgery demonstrate no differences in outcome and complications between those administered enteral nutrition and those given maintenance intravenous fluids alone in the initial days after surgery.252

**Early vs. Late Feeding**

Current recommendations support early enteral nutrition (within 48 hours) in critically ill patients, but with a caveat.253 Early “full nutrition” is likely to be harmful and is associated with a higher infection rate. The aim therefore is a caloric target below the actual energy expenditure, with the goal of providing >80% of estimated total energy goals gradually by 3 to 4 days. Early EN may be protective of the enteral epithelial barrier function and help to maintain the diversity of the microbiome. While early caloric limitation seems to benefit the critically ill patient when compared to overfeeding, the restriction likely creates a significant shortfall in protein provision considering the low protein-to-calorie ratio of most enteral products.254

In this regard, it is important to distinguish “permissive underfeeding” in which the total calories provided average 1500 kcal/d with 40 gm/d of protein from hypocaloric nutrition which has the same total calories with 140 gm/d protein. Hypocaloric nutrition is currently recommended for critically ill obese patients, but some investigators argue that this nutritional strategy may also benefit nonobese patients especially during

---

### Table 2-9

<table>
<thead>
<tr>
<th>CONDITION</th>
<th>kcal/kg PER DAY</th>
<th>ADJUSTMENT ABOVE BEE</th>
<th>GRAMS OF PROTEIN/kg PER DAY</th>
<th>NONPROTEIN CALORIES: NITROGEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal/moderate malnutrition</td>
<td>25–30</td>
<td>1.1</td>
<td>1.0</td>
<td>150:1</td>
</tr>
<tr>
<td>Mild stress</td>
<td>25–30</td>
<td>1.2</td>
<td>1.2</td>
<td>150:1</td>
</tr>
<tr>
<td>Moderate stress</td>
<td>30</td>
<td>1.4</td>
<td>1.5</td>
<td>120:1</td>
</tr>
<tr>
<td>Severe stress</td>
<td>30–35</td>
<td>1.6</td>
<td>2.0</td>
<td>90–120:1</td>
</tr>
<tr>
<td>Burns</td>
<td>35–40</td>
<td>2.0</td>
<td>2.5</td>
<td>90–100:1</td>
</tr>
</tbody>
</table>

Caloric adjustments above basal energy expenditure (BEE) in hypermetabolic conditions
the early acute period of critical illness. This recommendation excludes those patients whose pre-ICU weight loss indicate that they are malnourished.

As the patient enters the recovery period of their illness, total protein and caloric requirements are likely to significantly increase. Based on our understanding of starvation, increased extrinsic delivery of both calories and protein are likely to be required during this period.

For patients undergoing elective surgery, healthy patients without malnutrition who are undergoing uncomplicated surgery can tolerate 10 days of partial starvation (i.e., maintenance intravenous fluids only) before any clinically significant protein catabolism occurs. Earlier intervention is likely indicated for patients in whom preoperative protein-calorie malnutrition has been identified. Other clinical scenarios for which the benefits of enteral nutritional support have been substantiated include permanent neurologic impairment, oropharyngeal dysfunction, short-bowel syndrome, and bone marrow transplantation.

Initiation of enteral nutrition should occur as soon as feasible after adequate resuscitation, most readily determined by adequate urine output. The presence of bowel sounds and the passage of flatus or stool are not absolute prerequisites for initiation of enteral nutrition, but in the setting of gastroparesis feedings should be administered distal to the pylorus. Gastric residuals of 200 mL or more in a 4- to 6-hour period or abdominal distention requires cessation of feeding and adjustment of the infusion rate. Concomitant gastric decompression with distal small-bowel feedings may be appropriate in certain patients such as closed-head injury patients with gastroparesis. There is no evidence to support withholding enteral feedings for patients after bowel resection or for those with low-output enterocutaneous fistulas of <500 mL/d. In fact, a recent systematic review of studies of early enteral feeding (within 24 hours of gastrointestinal surgery) showed no effect on anastomotic leak and a reduction in mortality. Early enteral feeding is also associated in reduced incidence of fistula formation in patients with open abdomen. Enteral feeding should also be offered to patients with short-bowel syndrome or clinical malabsorption, but necessary calories, essential minerals, and vitamins should be supplemented using parenteral modalities.

**Intermittent vs. Continuous Enteral Feeding**

Enteral nutrition can be administered either continuously or intermittently; however, the standard choice for critically injured adults is continuous enteral feeding (CEF) due to the lower complication rates. Data also suggest that CEF may promote protein anabolism by inhibiting protein breakdown.

**Enteral Formulas**

For most critically ill patients, the choice of enteral formula will be determined by a number of factors and will include a clinical judgment as to the “best fit” for the patient’s needs. In general, feeding formulas to consider are GI tolerance-promoting, anti-inflammatory, immune-modulating, organ supportive, and standard enteral nutrition. In addition, guidelines from professional nutrition societies identify certain populations of patients who can benefit from formulations with specific pharmaconutrients. For many others, each physician must use his or her own clinical judgment about what formula will best meet the patient’s needs.

The functional status of the gastrointestinal tract determines the type of enteral solutions to be used. Patients with an intact gastrointestinal tract will tolerate complex solutions, but patients who have not been fed via the gastrointestinal tract for prolonged periods are less likely to tolerate complex carbohydrates. In those patients who are having difficulty tolerating standard enteral formulas, peptide- and medium-chain triglyceride-based formulas with probiotics can lessen GI tolerance problems. Additionally, in patients with demonstrated malabsorption issues, such as with inflammatory bowel diseases or short bowel syndrome, current guidelines endorse the provision of hydrolyzed protein formulas to improve absorption. Guidelines have not yet been made with regard to the fiber content of enteral formulas. However, recent evidence indicates that supplementation of enteral formulas with soluble dietary fiber may be beneficial for improving stool consistency in patients suffering from diarrhea.

Factors that influence the choice of enteral formula also include the extent of organ dysfunction (e.g., renal, pulmonary, hepatic, or gastrointestinal), the nutrients needed to restore optimal function and healing, and the cost of specific products. There are still no conclusive data to recommend one category of product over another, and nutritional support committees typically develop the most cost-efficient enteral formulary for the most commonly encountered disease categories within the institution.

As discussed extensively in the first sections of this chapter, surgery and trauma result in a significant “sterile” inflammatory response that impacts for the innate and adaptive immune systems. The provision of immune-modulating nutrients, termed “immunonutrition,” is one mechanism by which the immune response can be supported and an attempt made to lower infectious risk. At present, the best studied of immune-nutrients are glutamine, arginine, and omega-3 PUFAs.

**Immunonutrients.** As discussed extensively in the first sections of this chapter, surgery and trauma result in a significant “sterile” inflammatory response that impacts both the innate and adaptive immune systems. The provision of immune-modulating nutrients, termed “immunonutrition,” is one mechanism by which the immune response can be supported and an attempt made to lower infectious risk. Studies have shown that a variety of nutrients, including amino acids (glutamine and arginine); lipids (omega-3 PUFAs); and micronutrients (e.g., vitamin C and selenium) can provide support to the immune system. While current evidence does not support their use universally, benefit may exist for individual patients. At present, the best studied of immune-nutrients are glutamine, arginine, and omega-3 PUFAs.

**Glutamine** is the most abundant amino acid in the human body, comprising nearly two thirds of the free intracellular amino acid pool. Considered a nonessential amino acid, glutamine is a necessary substrate for nucleotide synthesis in most dividing cells and hence provides a major fuel source for enterocytes. It also serves as an important fuel source for immuneocytes. During stress states, peripheral glutamine stores are rapidly depleted, and the amino acid is preferentially shunted as a fuel source toward the visceral organs and tumors, respectively. These situations create, at least experimentally, a glutamine-depleted environment with potential immune consequences, thus generating interest in both enteral and parenteral glutamine supplementation. However, recently reported data from two large randomized controlled clinical trials in which critically ill patients received glutamine supplementation demonstrated...
Therefore, glutamine supplementation in the critically ill patient is not currently recommended.247

Arginine, also a nonessential amino acid in healthy subjects, first attracted attention for its immunoenhancing properties, wound-healing benefits, and association with improved survival in animal models of sepsis and injury.261 However, arginine can be metabolized to nitric oxide, via nitric oxide synthase (NOS). If NOS is upregulated, with arginine as available substrate, NO production can also increase, which can have a negative impact on the critically ill patient. As with glutamine, the benefits of experimental arginine supplementation during stress states are diverse. In clinical studies involving critically ill and injured patients and patients who have undergone surgery for certain malignancies, enteral administration of arginine has led to net nitrogen retention and protein synthesis, whereas isonitrogenous diets have not. Some of these studies also provide in vitro evidence of enhanced immunocyte function. The clinical utility of arginine supplementation in improving overall patient outcome remains an area of investigation.262

As previously discussed, omega-3 polyunsaturated fatty acids (PUFAs, canola oil, or fish oil) displaces omega-6 fatty acids in cell membranes, which theoretically reduces the pro-inflammatory response from prostaglandin production. Hence, there has been significant interest in reducing the ratio of omega-6 to omega-3. The data regarding supplementation of enteral feedings with fish oil as a source for omega-3 PUFAs has been mixed, however, with no demonstrated improvement in respiratory complications in severe trauma patients and possible benefits in patients with mild sepsis.263

Standard Polymeric Formulas. Most polymeric formulas provide a caloric density from 1 to 2 kcal/mL, and approximately 1500 to 1800 mL are required to meet daily requirements. These compositions provide baseline carbohydrates, protein, electrolytes, water, fat, and fat-soluble vitamins (some do not have vitamin K). These contain no fiber bulk and therefore leave minimum residue. These solutions usually are considered to be the standard or first-line formulas for stable patients with an intact gastrointestinal tract. Normal digestive function is required for this formula.

Fiber-Containing Formulas. Isotonic formulas with fiber contain soluble and insoluble fiber, which is most often soy based. Physiologically, fiber-based solutions delay intestinal transit time and may reduce the incidence of diarrhea compared with nonfiber solutions. It is most beneficial in this regard in patients who have a high number of loose stools.264 Fiber stimulates pancreatic lipase activity and is degraded by gut bacteria into short-chain fatty acids (SCFAs), an important fuel for colonicocytes. Recent data have also demonstrated the expression of SCFA receptors on leukocytes, suggesting that fiber fermentation by the colonic microbiome may indirectly regulate immune cell function. Another potential plus of fiber-containing formulas is the inclusion of prebiotic fibers with the goal of positively impacting bacterial targets in the gut as well as gut barrier function. While there has been limited research in this area to determine the possible impact on clinical outcomes, addition of these fermentable soluble fiber additives is something that should be considered in the ICU patient as a measure that can aid in the maintenance or restoration of a healthy balance of commensal gut bacteria.

Immune-Enhancing Formulas. Immune-enhancing formulas are fortified with special nutrients that are purported to enhance various aspects of immune or solid organ function as previously discussed. Such additives include glutamine, arginine, omega-3 fatty acids, and nucleotides.264 Although several trials have proposed that one or more of these additives reduce surgical complications and improve outcome, these results have not been uniformly corroborated by other trials. The Canadian Clinical Practice Guidelines currently do not recommend the addition of arginine supplements for critically ill patients due to the potential for harm when used in septic patients.265 Omega-3 PUFAs results from the EDEN-Omega study demonstrated that twice-daily enteral supplementation of n-3 fatty acids, γ-linolenic acid, and antioxidants did not improve the primary end point of ventilator-free days or other clinical outcomes in patients with acute lung injury and may be harmful.266 Glutamine supplementation should be strictly guided by the individual patient condition for the reasons discussed previously.

Calorie-Dense Formulas. The primary distinction of calorie-dense formulas is a greater caloric value for the same volume. Most commercial products of this variety provide 1.5 to 2 kcal/mL and therefore are suitable for patients requiring fluid restriction or those unable to tolerate large-volume infusions. As expected, these solutions have higher osmolality than standard formulas and are suitable for intragastric feedings.

High-Protein Formulas/Bariatric Formulas. High-protein formulas are available in isotonic and nonisotonic mixtures and are proposed for critically ill or trauma patients with high protein requirements. These formulas have nonprotein-calorie to nitrogen ratios between 80:1 and 120:1. While some observational studies show improved outcomes with higher protein intakes in critically ill patients, there is limited data from randomized trials that prevents making strong conclusions about the dose of protein in critically ill patients.

As discussed previously, there has been support for high-protein, hypocaloric feeding in obese patients. As such, enteral formulas termed “bariatric formulas” have been developed. As an example, one product has 1 kcal/mL of formula, with 37% of the calories coming from protein. As the evidence for high-protein, hypocaloric feeding is low grade, it is unclear whether clinical outcomes with respect to survival and infectious complications is improved, and more data is required for definitive recommendation.267

Elemental Formulas. Elemental formulas contain predigested nutrients and provide proteins in the form of small peptides. Complex carbohydrates are limited, and fat content, in the form of MCTs and LCTs, is minimal. The primary advantage of such a formula is ease of absorption, but the inherent scarcity of fat, associated vitamins, and trace elements limits its long-term use as a primary source of nutrients. Due to its high osmolality, dilution, or slow infusion rates usually are necessary, particularly in critically ill patients. These formulas have been used frequently in patients with malabsorption, gut impairment, and pancreatitis, but their cost is significantly higher than that of standard formulas. To date, there has been no evidence of their benefit in routine use.

Renal-Failure Formulas. The primary benefits of renal formulas are the lower fluid volume and concentrations of potassium, phosphorus, and magnesium needed to meet daily calorie requirements. This type of formulation almost exclusively
contains essential amino acids and has a high nonprotein-calorie to nitrogen ratio; however, it does not contain trace elements or vitamins. Current guidelines suggest that patients with chronic kidney disease (CKD) who require enteral feeding should be placed on “standard enteral formulations.” Moreover, standard recommendations for both protein and calories support are also recommended.

**Hepatic-Failure Formulas.** Close to 50% of the proteins in hepatic-failure formulas are branched-chain amino acids (e.g., leucine, isoleucine, and valine). The goal of such a formula is to reduce aromatic amino acid levels and increase the levels of branched-chain amino acids, which can potentially reverse encephalopathy in patients with hepatic failure. The use of these formulas is controversial, however, because no clear benefits have been proven by clinical trials. Protein restriction should be avoided in patients with end-stage liver disease because such patients have significant protein energy malnutrition that predisposes them to additional morbidity and mortality. Similar to patients with CKD, standard formulations are recommended initially unless the patient develops hepatic encephalopathy that is refractory to standard treatment. With regard to protein supplementation, data indicate that providing 1.5 gm protein/kg per day improves clinical outcomes in these patients.

**Access for Enteral Nutritional Support**

The available techniques and repertoire for enteral access have provided multiple options for feeding the gut. Presently used methods and preferred indications are summarized in Table 2-10.

**Nasoenteric Tubes.** Nasogastric feeding should be reserved for those with intact mentation and protective laryngeal reflexes to minimize risks of aspiration. Even in intubated patients, nasogastric feedings often can be recovered from tracheal suction. Nasojejunal feedings are associated with fewer pulmonary complications including risk of pneumonia, but access past the pylorus requires greater effort to accomplish. Therefore, routine use of small bowel feedings is preferred in units where small bowel access is readily feasible. Where there may be difficulties obtaining access, small bowel feedings may be considered a priority for those patients at high risk for intolerance to enteral nutrition (e.g., high gastric residuals).

Blind insertion of nasogastric feeding tubes is fraught with misplacement, and air instillation with auscultation is inaccurate for ascertaining proper positioning. Radiographic confirmation is usually required to verify the position of the nasogastric feeding tube.

Several methods have been recommended for the passage of nasoenteric feeding tubes into the small bowel, including use of prokinetic agents, right lateral decubitus positioning, gastric insufflation, tube angulation, and application of clockwise torque. However, the successful placement of feeding tubes by these methods is highly variable and operator dependent. Furthermore, it is time consuming, and success rates for intubation past the duodenum into the jejunum by these methods are <20%. Fluoroscopy-guided intubation past the pylorus has a >90% success rate, and more than half of these intubations result in jejunal placement. Similarly, endoscopy-guided placement past the pylorus has high success rates, but attempts to advance the tube beyond the second portion of the duodenum using a standard gastroduodenoscope is unlikely to be successful.

**Table 2-10**

<table>
<thead>
<tr>
<th>Options for enteral feeding access</th>
<th>ACCESS OPTION</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nasogastric tube</td>
<td>Short-term use only; aspiration risks; nasopharyngeal trauma; frequent dislodgment</td>
<td></td>
</tr>
<tr>
<td>Nasoduodenal/ nasojejunal tube</td>
<td>Short-term use; lower aspiration risks in jejunum; placement challenges (radiographic assistance often necessary)</td>
<td></td>
</tr>
<tr>
<td>Percutaneous endoscopic gastrostomy (PEG)</td>
<td>Endoscopy skills required; may be used for gastric decompression or bolus feeds; aspiration risks; can last 12–24 mo; slightly higher complication rates with placement and site leaks</td>
<td></td>
</tr>
<tr>
<td>Surgical gastrostomy</td>
<td>Requires general anesthesia and small laparotomy; procedure may allow placement of extended duodenal/ jejunal feeding ports; laparoscopic placement possible</td>
<td></td>
</tr>
<tr>
<td>Fluoroscopic gastrostomy</td>
<td>Blind placement using needle and T-prongs to anchor to stomach; can thread smaller catheter through gastrostomy into duodenum/jejunum under fluoroscopy</td>
<td></td>
</tr>
<tr>
<td>PEG-jejunal tube</td>
<td>Jejunal placement with regular endoscope is operator dependent; jejunal tube often dislodges retrograde; two-stage procedure with PEG placement, followed by fluoroscopic conversion with jejunal feeding tube through PEG</td>
<td></td>
</tr>
<tr>
<td>Direct percutaneous endoscopic jejunal gastrostomy (DPEJ)</td>
<td>Direct endoscopic tube placement with enteroscope; placement challenges; greater injury risks</td>
<td></td>
</tr>
<tr>
<td>Surgical jejunostomy</td>
<td>Commonly carried out during laparotomy; general anesthesia; laparoscopic placement usually requires assistant to thread catheter; laparoscopy offers direct visualization of catheter placement</td>
<td></td>
</tr>
<tr>
<td>Fluoroscopic jejunostomy</td>
<td>Difficult approach with injury risks; not commonly done</td>
<td></td>
</tr>
</tbody>
</table>

Small-bowel feeding is more reliable for delivering nutrition than nasogastric feeding. Furthermore, the risks of aspiration pneumonia can be reduced by 25% with small-bowel feeding compared with nasogastric feeding. The disadvantages of the use of nasoenteric feeding tubes are clogging, kinking, and inadvertent displacement or removal of the tube as well as nasopharyngeal complications. If nasoenteric feeding will be required for longer than 30 days, access should be converted to a percutaneous one.
Percutaneous Endoscopic Gastrostomy. The most common indications for percutaneous endoscopic gastrostomy (PEG) include impaired swallowing mechanisms, oropharyngeal or esophageal obstruction, and major facial trauma. It is frequently used for debilitated patients requiring caloric supplementation, hydration, or frequent medication dosing. It is also appropriate for patients requiring passive gastric decompression. Relative contraindications for PEG placement include ascites, coagulopathy, gastric varices, gastric neoplasm, and lack of a suitable abdominal site. Most tubes are 18F to 28F in size and may be used for 12 to 24 months.

Identification of the PEG site requires endoscopic transillumination of the anterior stomach against the abdominal wall. A 14-gauge angiocatheter is passed through the abdominal wall into the fully insufflated stomach. A guidewire is threaded through the angiocatheter, grasped by snare or forceps, and pulled out through the mouth. The tapered end of the PEG tube is secured to the guidewire and is pulled into position out of the abdominal wall. The PEG tube is secured without tension against the abdominal wall, and many have reported using the tube within hours of placement. It has been the practice of some to connect the PEG tube to a drainage bag for passive decompression for 24 hours before use, allowing more time for the stomach to seal against the peritoneum.

If endoscopy is not available or technical obstacles preclude PEG placement, the interventional radiologist can attempt the procedure percutaneously under fluoroscopic guidance by first insufflating the stomach against the abdominal wall with a nasogastric tube. If this also is unsuccessful, surgical gastrostomy tube placement can be considered, particularly with minimally invasive methods. When surgery is contemplated, it may be wise to consider directly accessing the small bowel for nutrition delivery.

Although PEG tubes enhance nutritional delivery, facilitate nursing care, and are superior to nasogastric tubes, serious complications occur in approximately 3% of patients. These complications include wound infection, necrotizing fasciitis, peritonitis, aspiration, leaks, dislodgment, bowel perforation, enteric fistulas, bleeding, and aspiration pneumonia. For patients with significant gastroparesis or gastric outlet obstruction, feedings through PEG tubes are hazardous. In such cases, the PEG tube can be used for decompression and allow access for converting the PEG tube to a transpyloric feeding tube.

Percutaneous Endoscopic Gastrostomy-Jejunostomy and Direct Percutaneous Endoscopic Jejunostomy. Although gastric bolus feedings are more physiologic, patients who cannot tolerate gastric feedings or who have significant aspiration risks should be fed directly past the pylorus. In the percutaneous endoscopic gastrostomy-jejunostomy (PEG-J) method, a 9F to 12F tube is passed through an existing PEG tube, past the pylorus, and into the duodenum. This can be achieved by endoscopic or fluoroscopic guidance. With weighted catheter tips and guidewires, the tube can be further advanced past the ligament of Treitz. However, the incidence of long-term PEG-J tube malfunction has been reported to be >50% as a result of retrograde tube migration into the stomach, kinking, or clogging.

Direct percutaneous endoscopic jejunostomy (DPEJ) tube placement uses the same techniques as PEG tube placement but requires an enteroscope or colonoscope to reach the jejunum. DPEJ tube malfunctions are probably less frequent than PEG-J tube malfunctions, and kinking or clogging is usually averted by placement of larger-caliber catheters. The success rate of DPEJ tube placement is variable because of the complexity of endoscopic skills required to locate a suitable jejunal site. In such cases where endoscopic means are not feasible, surgical jejunostomy tube placement is more appropriate, especially when minimally invasive techniques are available.

Surgical Gastrostomy and Jejunostomy. For a patient undergoing complex abdominal or trauma surgery, thought should be given during surgery to the possible routes for subsequent nutritional support because laparotomy affords direct access to the stomach or small bowel. The only absolute contraindication to feeding jejunostomy is distal intestinal obstruction. Relative contraindications include severe edema of the intestinal wall, radiation enteritis, inflammatory bowel disease, ascites, severe immunodeficiency, and bowel ischemia. Needle-catheter jejunostomies also can be done with a minimal learning curve. The biggest drawback usually is possible clogging and knotting of the 6F catheter.

Abdominal distention and cramps are common adverse effects of early enteral nutrition. Some have also reported impaired respiratory mechanics as a result of intolerance to enteral feedings. These are mostly correctable by temporarily discontinuing feedings and resuming at a lower infusion rate.

Pneumatosis intestinalis and small-bowel necrosis are infrequent but significant problems in patients receiving jejunal tube feedings. Several contributing factors have been proposed, including the hyperosmolarity of enteral solutions, bacterial overgrowth, fermentation, and accumulation of metabolic breakdown products. The common pathophysiology is believed to be bowel distention and consequent reduction in bowel wall perfusion. Risk factors for these complications include cardiogenic and circulatory shock, vasopressor use, diabetes mellitus, and chronic obstructive pulmonary disease. Therefore, enteral feedings in the critically ill patient should be delayed until adequate resuscitation has been achieved. As alternatives, diluting standard enteral formula, delaying the progression to goal infusion rates, or using monomeric solutions with low osmolality requiring less digestion by the gastrointestinal tract all have been successfully used.

PARENTERAL NUTRITION

Parenteral nutrition is the continuous infusion of a hyperosmolar solution containing carbohydrates, proteins, fat, and other necessary nutrients through an indwelling catheter inserted into the superior vena cava. To obtain the maximum benefit, the calorie to protein ratio must be adequate (at least 100 to 150 kcal/g nitrogen), and both carbohydrates and proteins must be infused simultaneously. When the sources of calories and nitrogen are given at different times, there is a significant decrease in nitrogen utilization. These nutrients can be given in quantities considerably greater than the basic caloric and nitrogen requirements, and this method has proved to be highly successful in achieving growth and development, positive nitrogen balance, and weight gain in a variety of clinical situations. Clinical trials and meta-analysis of studies of parenteral feeding in the perioperative period have suggested that preoperative nutritional support may benefit some surgical patients, particularly those with extensive malnutrition.

Historically, short-term use of parenteral nutrition (PN) in critically ill patients (i.e., duration of <7 days) when enteral
nutrition (EN) may have been instituted was associated with higher rates of infectious complications. It appears, however, that the increased mortality associated with PN may have been associated with excessive caloric delivery. More recent data have shown no mortality difference between EN and PN when caloric delivery was reduced and matched. A recent meta-analysis confirmed this result and noted no increase in infectious complications. That being said, the risk/benefit for PN in the ICU is much smaller and in a patient with low nutritional risk provides little benefit over the first week of hospitalization in the ICU.

**Rationale for Parenteral Nutrition**

The principal indications for parenteral nutrition are malnutrition, sepsis, or surgical or traumatic injury in seriously ill patients for whom use of the gastrointestinal tract for feedings is not possible. Parenteral nutrition should not be used based solely on the medical diagnosis or disease state. Rather, PN use is recommended for those critically ill or injured patients who are at high nutritional risk, when EN is not possible. Alternatively, PN can also be used to supplement EN after 1 week of use if use of EN is unable to meet >60% of energy and protein requirements.

The safe and successful use of parenteral nutrition requires proper selection of patients with specific nutritional needs, experience with the technique, and an awareness of the associated complications. In patients with significant malnutrition, parenteral nutrition can rapidly improve nitrogen balance, which may enhance immune function. Routine postoperative use of parenteral nutrition is not shown to have clinical benefit and may be associated with a significant increase in complication rate. As with enteral nutrition, the fundamental goals are to provide sufficient calories and nitrogen substrate to promote tissue repair and to maintain the integrity or growth of lean tissue mass.

**Total Parenteral Nutrition**

Total parenteral nutrition (TPN), also referred to as central parenteral nutrition, requires access to a large-diameter vein to deliver the entire nutritional requirements of the individual. Dextrose content of the solution is high (15%–25%), and all other macronutrients and micronutrients are deliverable by this route.

**Peripheral Parenteral Nutrition**

The lower osmolarity of the solution used for peripheral parenteral nutrition (PPN), secondary to reduced levels of dextrose (5% to 10%) and protein (3%), allows its administration via peripheral veins. Some nutrients cannot be supplemented because they cannot be concentrated into small volumes. Therefore, PPN is not appropriate for repleting patients with severe malnutrition. It can be considered if central routes are not available or if supplemental nutritional support is required. Typically, PPN is used for short periods (<2 weeks). Beyond this time, TPN should be instituted.

**Initiation of Parenteral Nutrition**

The basic solution for parenteral nutrition contains a final concentration of 15% to 25% dextrose and 3% to 5% crystalline amino acids. The solutions usually are prepared in sterile conditions in the pharmacy from commercially available kits containing the component solutions and transfer apparatus. Preparation in the pharmacy under laminar flow hoods reduces the incidence of bacterial contamination of the solution. Proper preparation with suitable quality control is absolutely essential to avoid septic complications.

The proper provision of electrolytes and amino acids must take into account routes of fluid and electrolyte loss, renal function, metabolic rate, cardiac function, and the underlying disease state.

Intravenous vitamin preparations also should be added to parenteral formulas. Vitamin deficiencies are rare occurrences if such preparations are used. In addition, because vitamin K is not part of any commercially prepared vitamin solution, it should be supplemented on a weekly basis. During prolonged parenteral nutrition with fat-free solutions, essential fatty acid deficiency may become clinically apparent and manifests as dry, scaly dermatitis and loss of hair. The syndrome may be prevented by periodic infusion of a fat emulsion at a rate equivalent to 10% to 15% of total calories. Essential trace minerals may be required after prolonged TPN and may be supplied by direct addition of commercial preparations. The most frequent presentation of trace mineral deficiencies is the eczematoid rash developing both diffusely and at intertriginous areas in zinc-deficient patients. Other rare trace mineral deficiencies include a microcytic anemia associated with copper deficiency, and glucose intolerance presumably related to chromium deficiency. The latter complications are seldom seen except in patients receiving parenteral nutrition for extended periods. The daily administration of commercially available trace mineral supplements will obviate most such problems.

Depending on fluid and nitrogen tolerance, parenteral nutrition solutions generally can be increased over 2 to 3 days toward the desired infusion rate. Current recommendations suggest that hypocaloric nutrition (high protein with lower caloric dosing) be considered in the critically ill or injured over the first week in the ICU. The suggested target dose is <20 kcal/kg per day or <80% of estimated caloric needs with adequate protein (>1.2 g/kg per day). This strategy is suggested to minimize risk of both hyperglycemia and insulin resistance, which may reduce infectious complications. Insulin may be supplemented as necessary to ensure glucose tolerance, with a targeted blood glucose range of 140 or 150 to 180 mg/dL for the general ICU population. Administration of additional intravenous fluids and electrolytes may occasionally be necessary in patients with persistently high fluid losses.

The patient should be carefully monitored for development of electrolyte, volume, acid-base, and septic complications. Vital signs and urinary output should be measured regularly, and the patient should be weighed regularly. Frequent adjustments of the volume and composition of the solutions are necessary during the course of therapy. Samples for measurement of electrolytes are drawn daily until levels are stable and every 2 or 3 days thereafter. Blood counts, blood urea nitrogen level, levels of liver function indicators, and phosphate and magnesium levels are determined at least weekly.

The urine or capillary blood glucose level is checked every 6 hours, and serum glucose concentration is checked at least once daily during the first few days of the infusion and at frequent intervals thereafter. Relative glucose intolerance, which often manifests as glycosuria, may occur after initiation of parenteral nutrition. If blood glucose levels remain elevated or glycosuria persists, the dextrose concentration may be decreased, the infusion rate slowed, or regular insulin added to each bottle. The rise in blood glucose concentration observed after initiating parenteral nutrition may be temporary, as the normal pancreas...
increases its output of insulin in response to the continuous carbohydrate infusion. In patients with diabetes mellitus, additional insulin may be required.

Potassium is essential to achieve positive nitrogen balance and replace depleted intracellular stores. In addition, a significant shift of potassium ion from the extracellular to the intracellular space may take place because of the large glucose infusion, with resultant hypokalemia, metabolic alkalosis, and poor glucose utilization. In some cases, as much as 240 mEq of potassium ion daily may be required. Hypokalemia may cause glycemic increases, which would be treated with potassium, not insulin. Thus, before giving insulin, the serum potassium level must be checked to avoid exacerbating the hypokalemia.

Patients with insulin-dependent diabetes mellitus may exhibit wide fluctuations in blood glucose levels while receiving parenteral nutrition. This may require protocol-driven intravenous insulin therapy. In addition, partial replacement of dextrose calories with lipid emulsions may alleviate these problems in selected patients.

Lipid emulsions derived from soybean or safflower oils are widely used as an adjunctive nutrient to prevent the development of essential fatty acid deficiency, although recent data support reducing the overall omega-6 PUFA load in favor of omega-3 PUFA or MCT. Current recommendations are to limit intravenous fat emulsion infusion over the first week of hospitalization to a maximum of 100 g per week delivered in two divided doses. This is based on standard emulsions that are soy-based. As data is acquired for omega-3 PUFA-based emulsions, including fish-oil or olive-oil based emulsions, these recommendations may alter.

The delivery of parenteral nutrition requires central intravenous access. Temporary or short-term access can be achieved with a 16-gauge percutaneous catheter inserted into a subclavian or internal jugular vein and threaded into the superior vena cava. More permanent access with the intention of providing long-term or home parenteral nutrition can be achieved by placement of a catheter with a subcutaneous port for access by tunneling a catheter with a substantial subcutaneous length or threading a long catheter through the basilic or cephalic vein into the superior vena cava.

Complications of Parenteral Nutrition

Technical Complications. One of the more common and serious complications associated with long-term parenteral feeding is sepsis secondary to contamination of the central venous catheter. Contamination of solutions should also be considered, but it is rare when proper pharmacy protocols have been followed. Central line-associated blood stream infections (CLABSI) occur as a consequence of hematogenous seeding of the catheter with bacteria. One of the earliest signs of systemic sepsis from CVA-BSI may be the sudden development of glucose intolerance (with or without temperature increase) in a patient who previously has been maintained on parenteral alimentation without difficulty. When this occurs, or if high fever (> 38.5°C [101.3°F]) develops without obvious cause, a diligent search for a potential septic focus is indicated. Other causes of fever should also be investigated. If fever persists, the infusion catheter should be removed and submitted for culture. If the catheter is the cause of the fever, removal of the infectious source is usually followed by rapid defervescence. Some centers are now replacing catheters considered at low risk for infection over a guidewire. However, if blood cultures are positive and the catheter tip is also positive, then the catheter should be removed and placed in a new site. Should evidence of infection persist over 24 to 48 hours without a definable source, the catheter should be replaced into the opposite subclavian vein or into one of the internal jugular veins, and the infusion should be restarted.

The use of multilumen catheters may be associated with a slightly increased risk of infection. This is most likely associated with greater catheter manipulation and intensive use. The rate of catheter infection is highest for those placed in the femoral vein, lower for those in the jugular vein, and lowest for those in the subclavian vein. When catheters are indwelling for <3 days, infection risks are negligible. If indwelling time is 3 to 7 days, the infection risk is 3% to 5%. Indwelling times of >7 days are associated with a catheter infection risk of 5% to 10%. Strict adherence to barrier precautions also reduces the rate of infection as can the implementation of procedure checklists to ensure compliance with evidence-based guidelines shown to reduce infectious risk.

Other complications related to catheter placement include the development of pneumothorax, hemothorax, hydrothorax, subclavian artery injury, thoracic duct injury, cardiac arrhythmia, air embolism, catheter embolism, and cardiac perforation with tamponade. All of these complications may be avoided by strict adherence to proper techniques. Further, the use of ultrasonographic guidance during CV line placement has been demonstrated to significantly decrease the failure rate, complication rate, and number of attempts required for successful access.

Metabolic Complications. Hyperglycemia may develop with normal rates of infusion in patients with impaired glucose tolerance or in any patient if the hypertonic solutions are administered too rapidly. This is a particularly common complication in patients with latent diabetes and in patients subjected to severe surgical stress or trauma. Treatment of the condition consists of volume replacement with correction of electrolyte abnormalities and the administration of insulin. This complication can be avoided with careful attention to daily fluid balance and frequent monitoring of blood glucose levels and serum electrolytes.

Increasing experience has emphasized the importance of not overfeeding the parenterally nourished patient. This is particularly true for the depleted patient in whom excess calories infusion may result in carbon dioxide retention and respiratory insufficiency. In addition, excess feeding also has been related to the development of hepatic steatosis or marked glycogen deposition in selected patients. Cholestasis and formation of gallstones are common in patients receiving long-term parenteral nutrition. Mild but transient abnormalities of serum transaminase, alkaline phosphatase, and bilirubin levels occur in many parenterally nourished patients. Failure of the liver enzymes to plateau or return to normal over 7 to 14 days should suggest another etiology.

Intestinal Atrophy. Lack of intestinal stimulation is associated with intestinal mucosal atrophy, diminished villous height, bacterial overgrowth, reduced lymphoid tissue size, reduced immunoglobulin A production, and impaired gut immunity. The full clinical implications of these changes are not well realized, although bacterial translocation has been demonstrated in animal models. The most efficacious method to prevent these changes is to provide at least some nutrients enterally. In patients requiring TPN, it may be feasible to infuse small amounts of feedings via the gastrointestinal tract.
REFERENCES
Entries highlighted in bright blue are key references.


Chapter 2: Systemic Response to Injury and Metabolic Support


78

PART I
BASIC CONSIDERATIONS

106. Yang Y, Bazhin AV, Werner J, Karakhanova S. Reactive
108. C. N, Cunningham-Bussell A. Beyond oxidative stress: an
immunologist’s guide to reactive oxygen species. Nat Rev
109. Tschopp J, Schroder K. NLRP3 inflammasome activation: the
convergence of multiple signalling pathways on ROS production?
ars.2012.5149
111. Belikov AV, Schraven B, Simeoni L. T cells and reactive
s12929-015-0194-3
112. Schroder M, Kaufman RJ. The mammalian unfolded protein
annurev.biochem.73.011303.074134
113.	Grootjans J, Kaser A, Kaufman RJ, Blumberg RS. The
unfolded protein response in immunity and inflammation. Nat
114. Jeschke MG, Boehning D. Endoplasmic reticulum stress
and insulin resistance post-trauma: similarities to type
10.1111/j.1582-4934.2011.01405.x
doi: 10.1371/journal.pone.0130635
116. Smith JA. Regulation of cytokine production by the unfolded
protein response; implications for infection and autoimmunity.
is associated with insulin resistance, endoplasmic reticulum
2012;255(2):370-378. doi: 10.1097/SLA.0b013e31823e76e7
118. Thiessen SE, Van den Berghe G, Vanhorebeek I. Mitochondrial
and endoplasmic reticulum dysfunction and related defense
mechanisms in critical illness-induced multiple organ failure.
Biochim Biophys Acta. 2017;1863(10 pt B):2534-2545. doi:
10.1016/j.bbadis.2017.02.015
response stimulates FGF21 expression: systemic enhancer
cellsig.2017.08.009
120. Thiessen SE, Vanhorebeek I, Derese I, Gunst J, Van den
Berghe G. FGF21 response to critical illness: effect of blood
glucose control and relation with cellular stress and survival.
J Clin Endocrinol Metab. 2015;100(10):E1319-E1327. doi:
10.1210/jc.2015-2700
10.1038/icb.2012.82
122. Levine B, Mizushima N, Virgin HW. Autophagy in immunity
and inflammation. Nature. 2011;469(7330):323-335. doi:
10.1038/nature09782
123. Yen YT, Yang HR, Lo HC, et al. Enhancing autophagy with
10.1016/j.surg.2012.11.021
125. Kaczmarek A, Vandenabeele P, Krysko DV. Necroptosis:
the release of damage-associated molecular patterns and its

10.1016/j.immuni.2013.02.003
10.1016/j.immuni.2011.09.020
128.	Man SM, Karki R, Kanneganti TD. Molecular mechanisms and functions of pyroptosis, inflammatory caspases
2017;277(1):61-75. doi: 10.1111/imr.12534
doi: 10.1097/MD.0000000000009859
130. Waters JP, Pober JS, Bradley JR. Tumour necrosis factor
10.1002/path.4187
131. Khalil AA, Hall JC, Aziz FA, Price P. Tumour necrosis factor:
1010-1016. doi: 10.1111/j.1445-2197.2006.03921.x
132. Dekker AB, Krijnen P, Schipper IB. Predictive value of cytokines for developing complications after polytrauma. World
v5.i3.187
133. Di Paolo NC, Shayakhmetov DM. Interleukin 1alpha and the
10.1038/ni.3503
134. Dinarello CA. Interleukin-1 in the pathogenesis and treatment
doi: 10.1182/blood-2010-07-273417
135. Sitia R, Rubartelli A. The unconventional secretion of IL1beta: handling a dangerous weapon to optimize inflammatory responses. [Published online ahead of print April 4, 2018.]
136. Stylianou E, Saklatvala J. Interleukin-1. Int J Biochem Cell
137. Sun M, Brady RD, Wright DK, et al. Treatment with an
138. Dinarello CA, Fantuzzi G. Interleukin-18 and host defense
doi: 10.1086/374751
139. Kaplanski G. Interleukin-18: Biological properties and role
doi: 10.1111/imr.12616
140. Novick D, Kim S, Kaplanski G, Dinarello CA. Interleukin-18,
more than a Th1 cytokine. Semin Immunol. 2013;25(6):
439-448. doi: 10.1016/j.smim.2013.10.014
rccm.201201-0003OC
142. Kinoshita M, Miyazaki H, Ono S, Seki S. Immunoenhancing therapy with interleukin-18 against bacterial infection in
immunocompromised hosts after severe surgical stress. J Leukoc
143. Molofsky AB, Savage AK, Locksley RM. Interleukin-33
in tissue homeostasis, injury, and inflammation. Immunity.
2015;42(6):1005-1019. doi: 10.1016/j.immuni.2015.06.006
144. Bessa J, Meyer CA, de Vera Mudry MC, et al. Altered subcellular localization of IL-33 leads to non-resolving lethal
jaut.2014.02.012


79

SYSTEMIC RESPONSE TO INJURY AND METABOLIC SUPPORT

indicates lung injury and on-going inflammatory changes after
severe trauma in a porcine long-term model. Prostaglandins
plefa.2017.09.014
165. Zhang MJ, Spite M. Resolvins: anti-inflammatory and proresolving mediators derived from omega-3 polyunsaturated
annurev-nutr-071811-150726
166. Calder PC, Adolph M, Deutz NE, et al. Lipids in the intensive
care unit: recommendations from the ESPEN Expert Group.
167. Calder PC. n-3 fatty acids, inflammation, and immunity—
relevance to postsurgical and critically ill patients. Lipids.
Nat Immunol. 2010;11(9):785-797. doi: 10.1038/ni.1923
169. Huber-Lang M, Gebhard F, Schmidt CQ, Palmer A, Denk S,
Wiegner R. Complement therapeutic strategies in trauma,
hemorrhagic shock and systemic inflammation—closing
10.1016/j.smim.2016.04.005
complement, coagulation, and kallikrein/kinin cross-talk act
as a linchpin in the events leading to thromboinflammation.
172. Faerber L, Drechsler S, Ladenburger S, Gschaidmeier H,
Fischer W. The neuronal 5-HT3 receptor network after 20
years of research—evolving concepts in management of pain
10.1016/j.ejphar.2007.01.028
173. Duerschmied D, Suidan GL, Demers M, et al. Platelet serotonin promotes the recruitment of neutrophils to sites of acute
inflammation in mice. Blood. 2013;121(6):1008-1015. doi:
10.1182/blood-2012-06-437392
tips.2005.07.002
175. O’Mahony L, Akdis M, Akdis CA. Regulation of the immune
10.1016/j.jaci.2011.06.051
2016;113(35):9852-9857. doi: 10.1073/pnas.1610253113
annurev.immunol.16.1.293
patterns of muscle loss in chronic systemic inflammation.
180. Duncan SA, Baganizi DR, Sahu R, Singh SR, Dennis VA.
SOCS proteins as regulators of inflammatory responses
induced by bacterial infections: a review. Front Microbiol.
molecules: the growing players in macrophage polarization

CHAPTER 2

145. Liao W, Lin JX, Leonard WJ. Interleukin-2 at the crossroads of
effector responses, tolerance, and immunotherapy. Immunity.
2007;8(12):1142-1148. doi: 10.1038/sj.embor.7401099
147. Jawa RS, Anillo S, Huntoon K, Baumann H, Kulaylat M.
10.1177/0885066610395679
148. Rose-John S. IL-6 trans-signaling via the soluble IL-6 receptor: importance for the proinflammatory activities of IL-6. Int
the first 24 h following trauma to predict immunological complications and mortality in trauma patients: a meta-analysis.
[Published online ahead of print November 14, 2017.] Eur J
151. Hutchins AP, Diez D, Miranda-Saavedra D. The IL-10/
10.1038/ni.2366
153. Weijer S, Florquin S, van der Poll T. Endogenous interleukin-12 improves the early antimicrobial host response to
murine Escherichia coli peritonitis. Shock. 2005;23(1):
54-58.
divergence of blunt trauma survival and mortality.
CCM.0000000000001951
the immune response to trauma: a multiparametric systems immunology approach. Crit Care Med. 2017;45(9):
1523-1530. doi: 10.1097/CCM.0000000000002577
156. Korn T, Bettelli E, Oukka M, Kuchroo VK. IL-17 and Th17
annurev.immunol.021908.132710
158. Rauch I, Muller M, Decker T. The regulation of inflammation by interferons and their STATs. JAK-STAT. 2013;2(1):
e23820-1-13.
159. Kernbauer E, Maier V, Rauch I, Muller M, Decker T. Route
of infection determines the impact of type I interferons on
2013;8(6):e65007. doi: 10.1371/journal.pone.0065007
10.1189/jlb.0213066
lymph diversion abrogates 5-lipoxygenase activation in the
kidney following trauma and hemorrhagic shock. J Trauma
TA.0000000000000231


81

SYSTEMIC RESPONSE TO INJURY AND METABOLIC SUPPORT

238. Uehara M, Plank LD, Hill GL. Components of energy expenditure in patients with severe sepsis and major trauma: a basis
2003;27(3):168-175.
242. Dahn MS, Mitchell RA, Lange MP, Smith S, Jacobs LA.
Hepatic metabolic response to injury and sepsis. Surgery.
243. Vidal-Puig A, O’Rahilly S. Metabolism. Controlling the
glucose factory. Nature. 2001;413(6852):125-126. doi:
10.1038/35093198
244. Volpi E, Sheffield-Moore M, Rasmussen BB, Wolfe RR. Basal
muscle amino acid kinetics and protein synthesis in healthy
EWG. Nutritional risk screening (NRS 2002): a new method
246.	Heyland DK, Dhaliwal R, Jiang X, Day AG. Identifying
critically ill patients who benefit the most from nutrition
therapy: the development and initial validation of a novel
10.1186/cc10546
247.	Taylor BE, McClave SA, Martindale RG, et al. Guidelines for the provision and assessment of nutrition support
therapy in the adult critically ill patient: Society of Critical
Care Medicine (SCCM) and American Society for Parenteral and Enteral Nutrition (A.S.P.E.N.). Crit Care Med.
(TICACOS): a prospective, randomized, controlled pilot study
of nutritional support in critically ill patients. Intensive Care
249. Heslin MJ, Brennan MF. Advances in perioperative nutrition:
parenteral nutrition in critically ill patients: an updated
s13054-016-1298-1
251.	Reignier J, Boisrame-Helms J, Brisard L, et al. Enteral versus parenteral early nutrition in ventilated adults with shock:
discussion 77-80.
253. Reintam Blaser A, Berger MM. Early or late feeding after
nu9121278
254. Oshima T, Deutz NE, Doig G, Wischmeyer PE, Pichard C.
Protein-energy nutrition in the ICU is the power couple: a
doi: 10.1016/j.clnu.2015.10.016
255. Hoffer LJ. High-protein hypocaloric nutrition for non-obese
doi: 10.1002/ncp.10091
continuous enteral feeding to prevent catabolism in acutely ill
adult and pediatric patients. Curr Opin Clin Nutr Metab Care.

CHAPTER 2

interleukin-10 and morbidity in human blunt trauma.
CCM.0000000000000248
220. Namas RA, Mi Q, Namas R, et al. Insights into the role of
chemokines, damage-associated molecular patterns, and
lymphocyte-derived mediators from computational models
of trauma-induced inflammation. Antioxid Redox Signal.
2015;23(17):1370-1387. doi: 10.1089/ars.2015.6398
221. Fortin CF, McDonald PP, Fulop T, Lesur O. Sepsis, leukocytes, and nitric oxide (NO): an intricate affair. Shock.
2010;33(4):344-352. doi: 10.1097/SHK.0b013e3181c0f068
10.1016/j.it.2015.01.003
223.	Darwiche SS, Pfeifer R, Menzel C, et al. Inducible nitric
SHK.0b013e31826c5afe
557-565. doi: 10.1038/sj.ki.5002340
iNOS inhibitor versus norepinephrine in the treatment of
SHK.0b013e3181d75967
doi: 10.1089/neu.2015.4340
230. Yeager ME, Belchenko DD, Nguyen CM, Colvin KL, Ivy DD,
Stenmark KR. Endothelin-1, the unfolded protein response,
and persistent inflammation: role of pulmonary artery smooth
doi: 10.1165/rcmb.2010-0506OC
233. Rondina MT, Weyrich AS, Zimmerman GA. Platelets
CIRCRESAHA.113.300512
234. Zimmerman GA, McIntyre TM, Prescott SM, Stafforini DM.
The platelet-activating factor signaling system and its regulators in syndromes of inflammation and thrombosis. Crit Care
235. Varpula M, Pulkki K, Karlsson S, Ruokonen E, Pettila V. Predictive value of N-terminal pro-brain natriuretic peptide in
1277-1283. doi: 10.1097/01.CCM.0000261893.72811.0F
236. Qian A, Zhang M, Zhao G. Dynamic detection of N-terminal
pro-B-type natriuretic peptide helps to predict the outcome
2015;41(1):57-64. doi: 10.1007/s00068-014-0406-7
237. Mitch WE, Price SR. Mechanisms activating proteolysis to
cause muscle atrophy in catabolic conditions. J Ren Nutr.




INTRODUCTION
Fluid and electrolyte management is paramount to the care of the surgical patient. Changes in both fluid volume and electrolyte composition occur preoperatively, intraoperatively, and postoperatively, as well as in response to trauma and sepsis. The sections that follow review the normal anatomy of body fluids, electrolyte composition and concentration abnormalities and treatments, common metabolic derangements, and alternative resuscitative fluids. These concepts are then discussed in relationship to management of specific surgical patients and their commonly encountered fluid and electrolyte abnormalities.

BODY FLUIDS

Total Body Water
Water constitutes approximately 50% to 60% of total body weight. The relationship between total body weight and total body water (TBW) is relatively constant for an individual and is primarily a reflection of body fat. Lean tissues such as muscle and solid organs have higher water content than fat and bone. As a result, young, lean males have a higher proportion of body weight as water than elderly or obese individuals. In an average young adult male, TBW accounts for 60% of total body weight, whereas in an average young adult female, it is 50%. The lower percentage of TBW in females correlates with a higher percentage of adipose tissue and lower percentage of muscle mass in most. Estimates of percentage of TBW should be adjusted downward approximately 10% to 20% for obese individuals and upward by 10% for malnourished individuals. The highest percentage of TBW is found in newborns, with approximately 80% of their total body weight comprised of water. This decreases to approximately 65% by 1 year of age and thereafter remains fairly constant.

Concentration Changes / 86
Composition Changes: Etiology and Diagnosis / 88
Acid-Base Balance / 91
Metabolic Derangements / 91
Fluid and Electrolyte Therapy / 93
Parenteral Solutions / 93
Alternative Resuscitative Fluids / 94
Correction of Life-Threatening Electrolyte Abnormalities / 94
Preoperative Fluid Therapy / 96
Intraoperative Fluid Therapy / 97

TBW is divided into three functional fluid compartments: plasma, extravascular interstitial fluid, and intracellular fluid (Fig. 3-1). The extracellular fluids (ECF), plasma, and interstitial fluid together compose about one-third of the TBW, and the intracellular compartment composes the remaining two thirds. The extracellular water composes 20% of the total body weight and is divided between plasma (5% of body weight) and interstitial fluid (15% of body weight). Intracellular water makes up approximately 40% of an individual’s total body weight, with the largest proportion in the skeletal muscle mass.

Composition of Fluid Compartments
The normal chemical composition of the body fluid compartments is shown in Fig. 3-2. The ECF compartment is balanced between sodium, the principal cation, and chloride and bicarbonate, the principal anions. The intracellular fluid compartment is composed primarily of the cations potassium and magnesium, the anions phosphate and sulfate, and proteins. The concentration gradient between compartments is maintained by adenosine triphosphate–driven sodium-potassium pumps located within the cell membranes. The composition of the plasma and interstitial fluid differs only slightly in ionic composition. The slightly higher protein content (organic anions) in plasma results in a higher plasma cation composition relative to the interstitial fluid, as explained by the Gibbs-Donnan equilibrium equation. Proteins add to the osmolality of the plasma and contribute to the balance of forces that determine fluid balance across the capillary endothelium. Although the movement of ions and proteins between the various fluid compartments is restricted, water is freely diffusible. Water is distributed evenly throughout all fluid compartments of the body so that a given volume of water increases the volume of any one compartment.
Key Points

1. Proper management of fluid and electrolytes facilitates crucial homeostasis that allows cardiovascular perfusion, organ system function, and cellular mechanisms to respond to surgical illness.

2. Knowledge of the compartmentalization of body fluids forms the basis for understanding pathologic shifts in these fluid spaces in disease states. Although difficult to quantify, a deficiency in the functional extracellular fluid compartment often requires resuscitation with isotonic fluids in surgical and trauma patients.

3. Alterations in the concentration of serum sodium have profound effects on cellular function due to water shifts between the intracellular and extracellular spaces.

4. Different rates of compensation between respiratory and metabolic components of acid-base homeostasis require frequent laboratory reassessment during therapy.

5. Although active investigation continues, alternative resuscitation fluids have limited clinical utility, other than the correction of specific electrolyte abnormalities.

6. Enhanced recovery after surgery (ERAS) protocols have markedly changed perioperative fluid management and are being used more frequently. ERAS minimizes perioperative fluid administration and focuses on early enteral intake to reduce morbidity associated with IV fluid administration.

7. Most acute surgical illnesses are accompanied by some degree of volume loss or redistribution. Consequently, isotonic fluid administration is the most common initial intravenous fluid strategy, while attention is being given to alterations in concentration and composition.

8. Some surgical patients with neurologic illness, malnutrition, acute renal failure, or cancer require special attention to well-defined, disease-specific abnormalities in fluid and electrolyte status.

relatively little. Sodium, however, is confined to the ECF compartment, and because of its osmotic and electrical properties, it remains associated with water. Therefore, sodium-containing fluids are distributed throughout the ECF and add to the volume of both the intravascular and interstitial spaces. Although the administration of sodium-containing fluids expands the intravascular volume, it also expands the interstitial space by approximately three times as much as the plasma.

Osmotic Pressure

The physiologic activity of electrolytes in solution depends on the number of particles per unit volume (millimoles per liter, or mmol/L), the number of electric charges per unit volume (milliequivalents per liter, or mEq/L), and the number of osmotically active ions per unit volume (milliosmoles per liter, or mOsm/L). The concentration of electrolytes usually is expressed in terms of the chemical combining activity, or equivalents. An equivalent of an ion is its atomic weight expressed in grams divided by the valence:

\[
\text{Equivalent} = \frac{\text{atomic weight (g)}}{\text{valence}}
\]

For univalent ions such as sodium, 1 mEq is same as 1 mmol. For divalent ions such as magnesium, 1 mmol equals 2 mEq. The number of milliequivalents of cations must be balanced by the same number of milliequivalents of anions. However, the expression of molar equivalents alone does not allow a physiologic comparison of solutes in a solution.

The movement of water across a cell membrane depends primarily on osmosis. To achieve osmotic equilibrium, water moves across a semipermeable membrane to equalize the concentration on both sides. This movement is determined by the concentration of the solutes on each side of the membrane. Osmotic pressure is measured in units of osmoles (osm) or milliosmoles (mOsm) that refer to the actual number of osmotically active particles. For example, 1 mmol of sodium chloride contributes to 2 mOsm (one from sodium and one from chloride). The principal determinants of osmolality are the concentrations of sodium, glucose, and urea (blood urea nitrogen, or BUN):

\[
\text{Calculated serum osmolality} = 2 \times \text{sodium} + \left( \frac{\text{glucose}}{18} \right) + \left( \frac{\text{BUN}}{2.8} \right)
\]

The osmolality of the intracellular and extracellular fluids is maintained between 290 and 310 mOsm in each compartment. Because cell membranes are permeable to water, any...
change in osmotic pressure in one compartment is accompanied by a redistribution of water until the effective osmotic pressure between compartments is equal. For practical clinical purposes, most significant gains and losses of body fluid are directly from the extracellular compartment.

**BODY FLUID CHANGES**

**Normal Exchange of Fluid and Electrolytes**

The healthy person consumes an average of 2000 mL of water per day, approximately 75% from oral intake and the rest extracted from solid foods. Daily water losses include 800 to 1200 mL in urine, 250 mL in stool, and 600 mL in insensible losses. Insensible losses of water occur through both the skin (75%) and lungs (25%) and can be increased by such factors as fever, hypermetabolism, and hyperventilation. Sensible water losses such as sweating or pathologic loss of gastrointestinal (GI) fluids vary widely, but these include the loss of electrolytes as well as water (Table 3-1). To clear the products of metabolism, the kidneys must excrete a minimum of 500 to 800 mL of urine per day, regardless of the amount of oral intake.

The typical individual consumes 3 to 5 g of dietary salt per day, with the balance maintained by the kidneys. With hypo- or hypovolemia, sodium excretion can be reduced to as little as 1 mEq/d or maximized to as much as 5000 mEq/d to achieve balance except in people with salt-wasting kidneys. Sweat is hypotonic, and sweating usually results in only a small sodium loss. GI losses are isotonic to slightly hypotonic and contribute little to net gain or loss of free water when measured and appropriately replaced by isotonic salt solutions.

**Classification of Body Fluid Changes**

Disorders in fluid balance may be classified into three general categories: disturbances in (a) volume, (b) concentration, and (c) composition. Although each of these may occur simultaneously, each is a separate entity with unique mechanisms demanding individual correction. Isotonic gain or loss of salt solution results in extracellular volume changes, with little impact on intracellular fluid volume. If free water is added or lost from the ECF, water will pass between the ECF and intracellular fluid until solute concentration or osmolarity is equalized between the compartments. Unlike with sodium, the concentration of most other ions in the ECF can be altered without significant change in the total number of osmotically active particles, producing only a compositional change.

**Disturbances in Fluid Balance**

Extracellular volume deficit is the most common fluid disorder in surgical patients and can be either acute or chronic. Acute volume deficit is associated with cardiovascular and central nervous system signs, whereas chronic deficits display tissue signs, such as a decrease in skin turgor and sunken eyes, in addition to cardiovascular and central nervous system signs (Table 3-2). Laboratory examination may reveal an elevated BUN level if the deficit is severe enough to reduce glomerular filtration and hemoconcentration. Urine osmolality usually will be higher than serum osmolality, and urine sodium will be low, typically <20 mEq/L. Serum sodium concentration does not necessarily reflect volume status and therefore may be high, normal, or low when a volume deficit is present. The most common cause of volume deficit in surgical patients is a loss of GI fluids (Table 3-3) from nasogastric suction, vomiting, diarrhea, or...
enterocutaneous fistula. In addition, sequestration secondary to soft tissue injuries, burns, and intra-abdominal processes such as peritonitis, obstruction, or prolonged surgery can also lead to massive volume deficits.

Extracellular volume excess may be iatrogenic or secondary to renal dysfunction, congestive heart failure, or cirrhosis. Both plasma and interstitial volumes usually are increased. Symptoms are primarily pulmonary and cardiovascular (see Table 3-2). In fit patients, edema and hyperdynamic circulation are common and well tolerated. However, the elderly and patients with cardiac disease may quickly develop congestive heart failure and pulmonary edema in response to only a moderate volume excess.

### Table 3-2

<table>
<thead>
<tr>
<th>SYSTEM</th>
<th>VOLUME DEFICIT</th>
<th>VOLUME EXCESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generalized</td>
<td>Weight loss</td>
<td>Weight gain</td>
</tr>
<tr>
<td>Cardiac</td>
<td>Tachycardia</td>
<td>Increased cardiac output</td>
</tr>
<tr>
<td>Orthostasis/ hypotension</td>
<td>Increased central venous pressure</td>
<td></td>
</tr>
<tr>
<td>Collapsed neck veins</td>
<td>Distended neck veins</td>
<td></td>
</tr>
<tr>
<td>Lungs and skin</td>
<td>600</td>
<td>1500</td>
</tr>
</tbody>
</table>

### Volume Control

Volume changes are sensed by both osmoreceptors and baroreceptors. Osmoreceptors are specialized sensors that detect even small changes in fluid osmolality and drive changes in thirst and diuresis through the kidneys. For example, when plasma osmolality is increased, thirst is stimulated and water consumption increases, although the exact cell mechanism is not known. Additionally, the hypothalamus is stimulated to secrete vasopressin, which increases water reabsorption in the kidneys. Together, these two mechanisms return the plasma osmolality to normal. Baroreceptors also modulate volume in response to changes in pressure and circulating volume through specialized pressure sensors located in the aortic arch and carotid sinuses. Baroreceptor responses are both neural, through sympathetic and parasympathetic pathways, and hormonal, through substances including renin-angiotensin, aldosterone, atrial natriuretic peptide, and renal prostaglandins. The net result of alterations in renal sodium excretion and free water reabsorption is restoration of volume to the normal state.

### Concentration Changes

Changes in serum sodium concentration are inversely proportional to TBW. Therefore, abnormalities in TBW are reflected by abnormalities in serum sodium levels.

**Hyponatremia.** A low serum sodium level occurs when there is an excess of extracellular water relative to sodium. Extracellular volume can be high, normal, or low (Fig. 3-3). In most cases of hyponatremia, sodium concentration is decreased as a consequence of either sodium depletion or dilution. Dilutional hyponatremia frequently results from excess extracellular water and therefore is associated with a high extracellular volume status. Excessive oral water intake or iatrogenic intravenous (IV) excess free water administration can cause hyponatremia. Postoperative patients are particularly prone to increased secretion of antidiuretic hormone (ADH), which increases reabsorption of water.
of free water from the kidneys with subsequent volume expansion and hyponatremia. This is usually self-limiting in that both hyponatremia and volume expansion decrease ADH secretion. Additionally, a number of drugs can cause water retention and subsequent hyponatremia, such as antipsychotics and tricyclic antidepressants as well as angiotensin-converting enzyme inhibitors. The elderly are particularly susceptible to drug-induced hyponatremia. Physical signs of volume overload usually are absent, and laboratory evaluation reveals hemodilution. Depletional causes of hyponatremia are associated with either a decreased intake or increased loss of sodium-containing fluids. A concomitant ECF volume deficit is common. Causes include decreased sodium intake, such as consumption of a low-sodium diet or use of enteral feeds, which are typically low in sodium;

<table>
<thead>
<tr>
<th>TYPE OF SECRETION</th>
<th>VOLUME (mL/24 h)</th>
<th>NA (mEq/L)</th>
<th>K (mEq/L)</th>
<th>CL (mEq/L)</th>
<th>HCO₃⁻ (mEq/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stomach</td>
<td>1000–2000</td>
<td>60–90</td>
<td>10–30</td>
<td>100–130</td>
<td>0</td>
</tr>
<tr>
<td>Small intestine</td>
<td>2000–3000</td>
<td>120–140</td>
<td>5–10</td>
<td>90–120</td>
<td>30–40</td>
</tr>
<tr>
<td>Colon</td>
<td>—</td>
<td>60</td>
<td>30</td>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td>Pancreas</td>
<td>600–800</td>
<td>135–145</td>
<td>5–10</td>
<td>70–90</td>
<td>95–115</td>
</tr>
<tr>
<td>Bile</td>
<td>300–800</td>
<td>135–145</td>
<td>5–10</td>
<td>90–110</td>
<td>30–40</td>
</tr>
</tbody>
</table>

Figure 3-3. Evaluation of sodium abnormalities. ADH = antidiuretic hormone; SIADH = syndrome of inappropriate secretion of antidiuretic hormone.
GI losses from vomiting, prolonged nasogastric suctioning, or diarrhea; and renal losses due to diuretic use or primary renal disease.

Hyponatremia also can be seen with an excess of solute relative to free water, such as with untreated hyperglycemia or mannitol administration. When hyponatremia in the presence of hyperglycemia is being evaluated, the corrected sodium concentration should be calculated as follows:

For every 100-mg/dL increment in plasma glucose above normal, the plasma sodium should decrease by 1.6 mEq/L.

Lastly, extreme elevations in plasma lipids and proteins can cause pseudohyponatremia because there is no true decrease in extracellular sodium relative to water.

Signs and symptoms of hyponatremia (Table 3-4) are dependent on the degree of hyponatremia and the rapidity with which it occurred. Clinical manifestations primarily have a central nervous system origin and are related to cellular water intoxication and associated increases in intracranial pressure. Oliguric renal failure also can be a rapid complication in the setting of severe hyponatremia.

A systematic review of the etiology of hyponatremia should reveal its cause in a given instance. Hyperosmolar causes, including hyperglycemia or mannitol infusion and pseudohyponatremia, should be easily excluded. Next, depletional versus dilutional causes of hyponatremia are evaluated. In the absence of renal disease, depletion is associated with low urine sodium levels (<20 mEq/L), whereas renal sodium wasting shows high urine sodium levels (>20 mEq/L). Dilutional causes of hyponatremia usually are associated with hypervolemic circulation. A normal volume status in the setting of hyponatremia should prompt an evaluation for a syndrome of inappropriate secretion of ADH.

**Hyponatremia.** Hyponatremia results from either a loss of free water or a gain of sodium in excess of water. Like hyponatremia, it can be associated with an increased, normal, or decreased extracellular volume (see Fig. 3-3). Hypervolemic hyponatremia usually is caused either by iatrogenic administration of sodium-containing fluids, including excess sodium bicarbonate, or mineralocorticoid as seen in hyperaldosteronism, Cushing’s syndrome, and congenital adrenal hyperplasia. Urine sodium concentration is typically >20 mEq/L, and urine osmolality is >300 mOsm/L. Normovolemic hyponatremia can result from renal causes, including diabetes insipidus, diuretic use, and renal disease, or from nonrenal water loss from the GI tract or skin, although the same conditions can result in hypovolemic hyponatremia. When hypovolemia is present, the urine sodium concentration is <20 mEq/L and urine osmolality is <300 to 400 mOsm/L. Nonrenal water loss can occur secondary to relatively isotonic GI fluid losses such as that caused by diarrhea, to hypotonic skin fluid losses such as loss due to fever, or to losses via tracheotomies during hyperventilation. Additionally, thyrotoxicosis can cause water loss, as can the use of hypertonic glucose solutions for peritoneal dialysis. With nonrenal water loss, the urine sodium concentration is <15 mEq/L, and the urine osmolality is >400 mOsm/L.

Symptomatic hyponatremia usually occurs only in patients with impaired thirst or restricted access to fluid because thirst will result in increased water intake. Symptoms are rare until the serum sodium concentration exceeds 160 mEq/L but, once present, are associated with significant morbidity and mortality. Because symptoms are related to hyperosmolality, central nervous system effects predominate (see Table 3-4). Water shifts from the intracellular to the extracellular space in response to a hyperosmolar extracellular space, which results in cellular dehydration. This can put traction on the cerebral vessels and lead to subarachnoid hemorrhage. Central nervous system symptoms can range from restlessness and irritability to seizures, coma, and death. The classic signs of hypovolemic hyponatremia (tachycardia, orthostasis, and hypotension) may be present, as well as the unique findings of dry, sticky mucous membranes.

### Table 3-4

<table>
<thead>
<tr>
<th>BODY SYSTEM</th>
<th>HYponatremia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central nervous system</td>
<td>Headache, confusion, hyperactive or hypoactive deep tendon reflexes, seizures, coma, increased intracranial pressure</td>
</tr>
<tr>
<td>Musculoskeletal</td>
<td>Weakness, fatigue, muscle cramps/twitching</td>
</tr>
<tr>
<td>GI</td>
<td>Anorexia, nausea, vomiting, watery diarrhea</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>Hypertension and bradycardia if intracranial pressure increases significantly</td>
</tr>
<tr>
<td>Tissue</td>
<td>Lacrimation, salivation</td>
</tr>
<tr>
<td>Renal</td>
<td>Oliguria</td>
</tr>
</tbody>
</table>

### Composition Changes: Etiology and Diagnosis

#### Potassium Abnormalities.** The average dietary intake of potassium is approximately 50 to 100 mEq/d, which in the absence of hypokalemia is excreted primarily in the urine. Extracellular potassium is maintained within a narrow range, principally by renal excretion of potassium, which can range from 10 to 700 mEq/d. Although only 2% of the total body potassium (4.5 mEq/L × 14 L = 63 mEq) is located within the extracellular compartment, this small amount is critical to cardiac and neuromuscular function; thus, even minor changes can have major effects on cardiac activity. The intracellular and extracellular distribution of potassium is influenced by a number of factors, including surgical stress, injury, acidosis, and tissue catabolism.
Hyperkalemia

Hyperkalemia is defined as a serum potassium concentration above the normal range of 3.5 to 5.0 mEq/L. It is caused by excessive potassium intake, increased release of potassium from cells, or impaired potassium excretion by the kidneys (Table 3-5). Increased intake can be either from oral or IV supplementation, or from red cell lysis after transfusion. Hemolysis, rhabdomyolysis, and crush injuries can disrupt cell membranes and release intracellular potassium into the ECF. Acidosis and a rapid rise in extracellular osmolality from hyperglycemia or IV mannitol can raise serum potassium levels by causing a shift of potassium ions to the extracellular compartment. Because 98% of total body potassium is in the intracellular fluid compartment, even small shifts of intracellular potassium out of the intracellular fluid compartment can lead to a significant rise in extracellular potassium. A number of medications can contribute to hyperkalemia, particularly in the presence of renal insufficiency, including potassium-sparing diuretics, angiotensin-converting enzyme inhibitors, and non-steroidal anti-inflammatory drugs (NSAIDs). Spironolactone and angiotensin-converting enzyme inhibitors interfere with aldosterone activity, inhibiting the normal renal mechanism of potassium excretion. Acute and chronic renal insufficiency also impairs potassium excretion.

Symptoms of hyperkalemia are primarily GI, neuromuscular, and cardiovascular (Table 3-6). GI symptoms include nausea, vomiting, intestinal colic, and diarrhea. Neuromuscular symptoms range from weakness to ascending paralysis to respiratory failure. Early cardiovascular signs may be apparent from electrocardiogram (ECG) changes and eventually lead to hemodynamic symptoms of arrhythmia and cardiac arrest. ECG changes that may be seen with hyperkalemia include high peaked T waves (early), widened QRS complex, flattened P wave, prolonged PR interval (first-degree block), sine wave formation, and ventricular fibrillation.

Hypokalemia

Hypokalemia is much more common than hyperkalemia in the surgical patient. It may be caused by inadequate potassium intake; excessive renal potassium excretion; potassium loss in pathologic GI secretions, such as with diarrhea, fistulas, vomiting, or high nasogastric output; or intracellular shifts from metabolic alkalosis or insulin therapy (see Table 3-5).
The change in potassium associated with alkalosis can be calculated by the following formula:

\[
\text{Potassium decreases by 0.3 mEq/L for every 0.1 increase in pH above normal.}
\]

Additionally, drugs such as amphotericin, aminoglycosides, cisplatin, and ifosfamide that induce magnesium depletion cause renal potassium wastage. In cases in which potassium deficiency is due to magnesium depletion, potassium repletion is difficult unless hypomagnesemia is first corrected.

The symptoms of hypokalemia (see Table 3-6), like those of hyperkalemia, are primarily related to failure of normal contractility of GI smooth muscle, skeletal muscle, and cardiac muscle. Findings may include ileus, constipation, weakness, fatigue, diminished tendon reflexes, paralysis, and cardiac arrest. In the setting of ECF depletion, symptoms may be masked initially and then worsened by further dilution during volume repletion. ECG changes suggestive of hypokalemia include U waves, T-wave flattening, ST-segment changes, and arrhythmias (with digitalis therapy).

**Calcium Abnormalities.** The vast majority of the body’s calcium is contained within the bone matrix, with <1% found in the ECF. Serum calcium is distributed among three forms: protein bound (40%), complexed to phosphate and other anions (10%), and ionized (50%). It is the ionized fraction that is responsible for neuromuscular stability and can be measured directly. When total serum calcium levels are measured, the albumin concentration must be taken into consideration:

\[
\text{Adjust total serum calcium down by 0.8 mg/dL for every 1 g/dL decrease in albumin.}
\]

Unlike changes in albumin, changes in pH will affect the ionized calcium concentration. Acidosis decreases protein binding, thereby increasing the ionized fraction of calcium.

Daily calcium intake is 1 to 3 g/d. Most of this is excreted via the bowel, with urinary excretion relatively low. Total body calcium balance is under complex hormonal control, but disturbances in metabolism are relatively long term and less important in the acute surgical setting. However, attention to the critical role of ionized calcium in neuromuscular function often is required.

**Hypercalcemia** Hypercalcemia is defined as a serum calcium level above the normal range of 8.5 to 10.5 mEq/L or an increase in the ionized calcium level above 4.2 to 4.8 mg/dL. Primary hyperparathyroidism in the outpatient setting and malignancy in hospitalized patients, from either bony metastasis or secretion of parathyroid hormone–related protein, account for most cases of symptomatic hypercalcemia. Symptoms of hypercalcemia (see Table 3-6), which vary with the degree of severity, include neurologic impairment, musculoskeletal weakness and pain, renal dysfunction, GI symptoms of nausea, vomiting, and abdominal pain. Cardiac symptoms can be manifested as hypertension, cardiac arrhythmias, and a worsening of digitalis toxicity. ECG changes in hypercalcemia include shortened QT interval, prolonged PR and QRS intervals, increased QRS voltage, T-wave flattening and widening, and atrioventricular block (which can progress to complete heart block and cardiac arrest).

**Hypocalcemia** Hypocalcemia is defined as a serum calcium level below 8.5 mEq/L or a decrease in the ionized calcium level below 4.2 mg/dL. The causes of hypocalcemia include pancreatitis, massive soft tissue infections such as necrotizing fasciitis, renal failure, pancreatic and small bowel fistulas, hypoparathyroidism, toxic shock syndrome, abnormalities in magnesium levels, and tumor lysis syndrome. In addition, transient hypocalcemia commonly occurs after removal of a parathyroid adenoma due to atrophy of the remaining glands and avid bone remineralization, and sometimes requires high-dose calcium supplementation. Additionally, malignancies associated with increased osteoblastic activity, such as breast and prostate cancer, can lead to hypocalcemia from increased bone formation. Calcium precipitation with organic anions is also a cause of hypocalcemia and may occur during hyperphosphatemia from tumor lysis syndrome or rhabdomyolysis. Pancreatitis may sequester calcium via chelation with free fatty acids. Massive blood transfusion with citrate binding is another mechanism. Hypocalcemia rarely results solely from decreased intake because bone reabsorption can maintain normal levels for prolonged periods.

Asymptomatic hypocalcemia may occur when hypophosphatemia results in a normal ionized calcium level. Conversely, symptoms can develop with a normal serum calcium level during alkalosis, which decreases ionized calcium. In general, neuromuscular and cardiac symptoms do not occur until the ionized fraction falls below 2.5 mg/dL (see Table 3-6). Clinical findings may include paresthesias of the face and extremities, muscle cramps, carpopedal spasm, stridor, tetany, and seizures. Patients will demonstrate hyperreflexia and may exhibit positive Chvostek’s sign (spasm resulting from tapping over the facial nerve) and Trousseau’s sign (spasm resulting from pressure applied to the nerves and vessels of the upper extremity with a blood pressure cuff). Hypocalcemia may lead to decreased cardiac contractility and heart failure. ECG changes of hypocalcemia include prolonged QT interval, T-wave inversion, heart block, and ventricular fibrillation.

**Phosphorus Abnormalities.** Phosphorus is the primary intracellular divalent anion and is abundant in metabolically active cells. Phosphorus is involved in energy production during glycolysis and is found in high-energy phosphate products such as adenosine triphosphate. Serum phosphate levels are tightly controlled by renal excretion.

**Hyperphosphatemia** Hyperphosphatemia can be due to decreased urinary excretion, increased intake, or endogenous mobilization of phosphorus. Most cases of hyperphosphatemia are seen in patients with impaired renal function. Hypoparathyroidism or hyperthyroidism also can decrease urinary excretion of phosphorus and thus lead to hyperphosphatemia. Increased release of endogenous phosphorus can be seen in association with any clinical condition that results in cell destruction, including rhabdomyolysis, tumor lysis syndrome, hemolysis, sepsis, severe hyperthermia, and malignant hyperthermia. Excessive phosphate administration from IV hyperalimentation solutions or phosphorus-containing laxatives may also lead to elevated phosphate levels. Most cases of hyperphosphatemia are asymptomatic, but significant prolonged hyperphosphatemia can lead to metastatic deposition of soft tissue calcium-phosphorus complexes.

**Hypophosphatemia** Hypophosphatemia can be due to a decrease in phosphorus intake, an intracellular shift of phosphorus, or an increase in phosphorus excretion. Decreased GI uptake due to malabsorption or administration of phosphate binders and decreased dietary intake from malnutrition are causes of chronic hypophosphatemia. Most acute cases are due to an intracellular
shift of phosphorus in association with respiratory alkalosis, insulin therapy, refeeding syndrome, and hungry bone syndrome. Clinical manifestations of hypophosphatemia usually are absent until levels fall significantly. In general, symptoms are related to adverse effects on the oxygen availability of tissue and to a decrease in high-energy phosphates, and can be manifested as cardiac dysfunction or muscle weakness.

**Magnesium Abnormalities.** Magnesium is the fourth most common mineral in the body and, like potassium, is found primarily in the intracellular compartments. Approximately one-half of the total body content of 2000 mEq is incorporated in bone and is slowly exchangeable. Of the fraction found in the extracellular space, one-third is bound to serum albumin. Therefore, the plasma level of magnesium may be a poor indicator of total body stores in the presence of hypoalbuminemia. Magnesium should be replaced until levels are in the upper limit of normal. The normal dietary intake is approximately 20 mEq/d and is excreted in both the feces and urine. The kidneys have a remarkable ability to conserve magnesium, with renal excretion <1 mEq/d during magnesium deficiency.

**Hypermagnesemia** Hypermagnesemia is rare but can be seen with severe renal insufficiency and parallel changes in potassium excretion. Magnesium-containing antacids and laxatives can produce toxic levels in patients with renal failure. Excess intake in conjunction with total parenteral nutrition (TPN), or, rarely, massive trauma, thermal injury, and severe acidosis, may be associated with symptomatic hypermagnesemia. Clinical examination (see Table 3-6) may find nausea and vomiting; neuromuscular dysfunction with weakness, lethargy, and hyporeflexia; and impaired cardiac conduction leading to hypotension and arrest. ECG changes are similar to those seen with hyperkalemia and include increased PR interval, widened QRS complex, and elevated T waves.

**Hypomagnesemia** Magnesium depletion is a common problem in hospitalized patients, particularly in the critically ill. The kidney is primarily responsible for magnesium homeostasis through regulation by calcium/magnesium receptors on the renal tubular cells that respond to serum magnesium concentrations. Hypomagnesemia may result from alterations of intake, renal excretion, and pathologic losses. Poor intake may occur in cases of starvation, alcoholism, prolonged IV fluid therapy, and TPN with inadequate supplementation of magnesium. Losses are seen in cases of increased renal excretion from alcohol abuse, diuretic use, administration of amphotericin B, and primary aldosteronism, as well as GI losses from diarrhea, malabsorption, and acute pancreatitis. The magnesium ion is essential for proper function of many enzyme systems. Depletion is characterized by neuromuscular and central nervous system hyperactivity. Symptoms are similar to those of calcium deficiency, including hyperactive reflexes, muscle tremors, tetany, and positive Chvostek’s and Trousseau’s signs (see Table 3-6). Severe deficiencies can lead to delirium and seizures. A number of ECG changes also can occur and include prolonged QT and PR intervals, ST-segment depression, flattening or inversion of P waves, torsades de pointes, and arrhythmias. Hypomagnesemia is important not only because of its direct effects on the nervous system but also because it can produce hypocalcemia and lead to persistent hypokalemia. When hypokalemia or hypocalcemia coexists with hypomagnesemia, magnesium should be aggressively replaced to assist in restoring potassium or calcium homeostasis.

**Acid-Base Balance**

**Acid-Base Homeostasis.** The pH of body fluids is maintained within a narrow range despite the ability of the kidneys to generate large amounts of HCO$_3^-$ and the normal large acid load produced as a by-product of metabolism. This endogenous acid load is efficiently neutralized by buffer systems and ultimately excreted by the lungs and kidneys.

Important buffers include intracellular proteins and phosphates and the extracellular bicarbonate–carbonic acid system. Compensation for acid-base derangements can be by respiratory mechanisms (for metabolic derangements) or metabolic mechanisms (for respiratory derangements). Changes in ventilation in response to metabolic abnormalities are mediated by hydrogen-sensitive chemoreceptors found in the carotid body and brain stem. Acidosis stimulates the chemoreceptors to increase ventilation, whereas alkalosis decreases the activity of the chemoreceptors and thus decreases ventilation. The kidneys provide compensation for respiratory abnormalities by either increasing or decreasing bicarbonate reabsorption in response to respiratory acidosis or alkalosis, respectively. Unlike the prompt change in ventilation that occurs with metabolic abnormalities, the compensatory response in the kidneys to respiratory abnormalities is delayed. Significant compensation may not begin for 6 hours and then may continue for several days. Because of this delayed compensatory response, respiratory acid-base derangements before renal compensation are classified as acute, whereas those persisting after renal compensation are categorized as chronic. The predicted compensatory changes in response to metabolic or respiratory derangements are listed in Table 3-7. If the predicted change in pH is exceeded, then a mixed acid-base abnormality may be present (Table 3-8).

**Metabolic Derangements**

**Metabolic Acidosis.** Metabolic acidosis results from an increased intake of acids, an increased generation of acids, or an increased loss of bicarbonate (Table 3-9). The body responds by several mechanisms, including producing buffers (extracellular bicarbonate and intracellular buffers from bone and muscle), increasing ventilation (Kussmaul’s respirations), and increasing renal reabsorption and generation of bicarbonate. The kidney also will increase secretion of hydrogen and thus increase urinary excretion of NH$_4^+$ (H$^+$ + NH$_3$ $\rightarrow$ NH$_4^+$). Evaluation of a patient with a low serum bicarbonate level and metabolic

<table>
<thead>
<tr>
<th>Table 3-7 Predicted changes in acid-base disorders</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DISORDER</strong></td>
</tr>
<tr>
<td><strong>Metabolic</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Respiratory</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

P$\text{CO}_2$ = partial pressure of carbon dioxide.
acidosis includes determination of the anion gap (AG), an index of unmeasured anions.

\[ \text{AG} = (\text{Na}) - (\text{Cl} + \text{HCO}_3^-) \]

The normal AG is <12 mmol/L and is due primarily to the albumin effect, so that the estimated AG must be adjusted for albumin (hypoaalbuminemia reduces the AG) \(^{19}\).

Corrected AG = actual AG – (2.5[4.5 – albumin])

Metabolic acidosis with an increased AG occurs either from ingestion of exogenous acid such as from ethylene glycol, salicylates, or methanol, or from increased endogenous acid production of the following:

- β-Hydroxybutyrate and acetoacetate in ketoacidosis
- Lactate in lactic acidosis
- Organic acids in renal insufficiency

A common cause of severe metabolic acidosis in surgical patients is lactic acidosis. In circulatory shock, lactate is produced in the presence of hypoxia from inadequate tissue perfusion. The treatment is to restore perfusion with volume resuscitation rather than to attempt to correct the abnormality with exogenous bicarbonate. With adequate perfusion, the lactic acid is rapidly metabolized by the liver, and the pH level returns to normal. In clinical studies of lactic acidosis and ketoacidosis, the administration of bicarbonate has not reduced morbidity or mortality or improved cellular function. \(^{20}\) Administered bicarbonate can combine with the excess hydrogen ions to form carbonic acid; this is then converted to CO\(_2\) and water, which thus raises the partial pressure of CO\(_2\) (Pco\(_2\)). This hypercarbia could compound ventilation abnormalities in patients with underlying acute respiratory distress syndrome. This CO\(_2\) can diffuse into cells, but bicarbonate remains extracellular, which thus worsens intracellular acidosis. Clinically, lactate levels may not be useful in directing resuscitation, although lactate levels may be higher in nonsurvivors of serious injury. \(^{21}\)

Metabolic acidosis with a normal AG results from exogenous acid administration (HCl or NH\(_4\)^+), from loss of bicarbonate due to GI disorders such as diarrhea and fistulas or ureterosigmoidostomy, or from renal losses. In these settings, the bicarbonate loss is accompanied by a gain of chloride; thus, the AG remains unchanged. To determine whether the loss of bicarbonate has a renal cause, the urinary (NH\(_4\)^+) can be measured. A low urinary (NH\(_4\)^+) in the face of hyperchloremic acidosis would indicate that the kidney is the site of loss, and evaluation for renal tubular acidosis should be undertaken. Proximal renal tubular acidosis results from decreased tubular reabsorption of HCO\(_3^-\), whereas distal renal tubular acidosis results from decreased acid excretion. The carbonic anhydrase inhibitor acetazolamide also causes bicarbonate loss from the kidneys.

### Table 3-8

#### Respiratory and metabolic components of acid-base disorders

<table>
<thead>
<tr>
<th>TYPE OF ACID-BASE DISORDER</th>
<th>ACUTE UNCOMPENSATED</th>
<th>CHRONIC (PARTIALLY COMPENSATED)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pH</td>
<td>PCO(_2) (RESPIRATORY COMPONENT)</td>
</tr>
<tr>
<td>Respiratory acidosis</td>
<td>↓↓</td>
<td>↑↑</td>
</tr>
<tr>
<td>Respiratory alkalosis</td>
<td>↑↑</td>
<td>↓↓</td>
</tr>
<tr>
<td>Metabolic acidosis</td>
<td>↓↓</td>
<td>N</td>
</tr>
<tr>
<td>Metabolic alkalosis</td>
<td>↑↑</td>
<td>N</td>
</tr>
</tbody>
</table>

\(^a\)Measured as standard bicarbonate, whole blood buffer base, CO\(_3\) content, or CO\(_2\) combining power. The base excess value is positive when the standard bicarbonate is above normal and negative when the standard bicarbonate is below normal.

N = normal; PCO\(_2\) = partial pressure of carbon dioxide.

### Table 3-9

#### Etiology of metabolic acidosis

**Increased Anion Gap Metabolic Acidosis**
- Exogenous acid ingestion
  - Ethylene glycol
  - Salicylate
  - Methanol
- Endogenous acid production
  - Ketoacidosis
  - Lactic acidosis
  - Renal insufficiency

**Normal Anion Gap**
- Acid administration (HCl)
- Loss of bicarbonate
- GI losses (diarrhea, fistulas)
- Ureterosigmoidostomy
- Renal tubular acidosis
- Carbonic anhydrase inhibitor
and intestinal secretions, vomiting with an obstructed pylorus results only in the loss of gastric fluid, which is high in chloride and hydrogen, and therefore results in a hypochloremic alkalosis. Initially, the urinary bicarbonate level is high in compensation for the alkalosis. Hydrogen ion reabsorption also ensues, with an accompanied potassium ion excretion. In response to the associated volume deficit, aldosterone-mediated sodium reabsorption increases potassium excretion. The resulting hypocalemia leads to the excretion of hydrogen ions in the face of alkalosis, a paradoxical aciduria. Treatment includes replacement of the volume deficit with isotonic saline and then potassium replacement once adequate urine output is achieved.

Respiratory Derangements. Under normal circumstances, blood PCO₂ is tightly maintained by alveolar ventilation, controlled by the respiratory centers in the pons and medulla oblongata.

Respiratory Acidosis Respiratory acidosis is associated with the retention of CO₂ secondary to decreased alveolar ventilation. The principal causes are listed in Table 3-11. Because compensation is primarily a renal mechanism, it is a delayed response. Treatment of acute respiratory acidosis is directed at the underlying cause. Measures to ensure adequate ventilation are also initiated. This may entail patient-initiated volume expansion using noninvasive bilevel positive airway pressure or may require endotracheal intubation to increase minute ventilation. In the chronic form of respiratory acidosis, the partial pressure of arterial CO₂ remains elevated, and the bicarbonate concentration rises slowly as renal compensation occurs.

Respiratory Alkalosis In the surgical patient, most cases of respiratory alkalosis are acute and secondary to alveolar hyperventilation. Causes include pain, anxiety, and neurologic disorders, including central nervous system injury and assisted ventilation. Drugs such as salicylates, fever, gram-negative bacteremia, thyrotoxicosis, and hypoxemia are other possibilities. Acute hypocapnia can cause an uptake of potassium and phosphate into cells and increased binding of calcium to albumin, leading to symptomatic hypokalemia, hypophosphatemia, and hypocalcemia with subsequent arrhythmias, paresthesias, muscle cramps, and seizures. Treatment should be directed at the underlying cause, but direct treatment of the hyperventilation using controlled ventilation may also be required.

### FLUID AND ELECTROLYTE THERAPY

#### Parenteral Solutions

A number of commercially available electrolyte solutions are available for parenteral administration. The most commonly used solutions are listed in Table 3-12. The type of fluid administered depends on the patient’s volume status and the type of concentration or compositional abnormality present. Plasma-Lyte, lactated Ringer’s solution, and normal saline are considered isotonic and are useful in replacing GI losses and correcting extracellular volume deficits. Lactated Ringer’s is slightly hypotonic in that it contains 130 mEq of lactate. Lactate is used rather than bicarbonate because it is more stable in IV fluids during storage. It is converted into bicarbonate by the liver after infusion, even in the face of hemorrhagic shock.

Sodium chloride is mildly hypertonic, containing 154 mEq of sodium that is balanced by 154 mEq of chloride. The high chloride concentration imposes a significant chloride load on the kidneys and may lead to a hyperchloremic metabolic acidos.

Sodium chloride is an ideal solution, however, for correcting volume deficits associated with hypernatremia, hypochloremia, and metabolic alkalosis.

Plasma-Lyte is the crystalloid preparation that most closely resembles the electrolyte composition of human plasma. In addition to the favorable, isotonic electrolyte composition, Plasma-Lyte contains a number of additional buffers that create a favorable profile for addressing acidosis. These characteristics have resulted in Plasma-Lyte emerging as one of the most popular isotonic fluids for use in surgery. It should be noted that Plasma-Lyte contains small quantities of potassium; although the likelihood of inducing hyperkalemia is very low, care should be taken when using Plasma-Lyte in patients at risk for renal impairment.

The less concentrated sodium solutions, such as 0.45% sodium chloride, are useful for replacement of ongoing GI losses as well as for maintenance fluid therapy in the postoperative period. This solution provides sufficient free water for insensible losses and enough sodium to aid the kidneys in adjustment of serum sodium levels. The addition of 5% dextrose...
BASIC CONSIDERATIONS

PART I

(50 g of dextrose per liter) supplies 200 kcal/L, and dextrose is always added to solutions containing <0.45% sodium chloride to maintain osmolality and thus prevent the lysis of red blood cells that may occur with rapid infusion of hypotonic fluids. The addition of potassium is useful once adequate renal function and urine output are established.

Alternative Resuscitative Fluids

A number of alternative solutions for volume expansion and resuscitation are available (Table 3-13). Hypertonic saline solutions (3.5% and 5%) are used for correction of severe sodium deficits and are discussed elsewhere in this chapter. Hypertonic saline (7.5%) has been used as a treatment modality in patients with closed head injuries. It has been shown to increase cerebral perfusion and decrease intracranial pressure, thus decreasing brain edema. However, there have also been concerns about increased bleeding because hypertonic saline is an arteriolar vasodilator. A recent meta-analysis of hypertonic saline in severe traumatic brain injury revealed that, in a total of 11 eligible studies, there was no mortality benefit associated with hypertonic saline compared to other solutions. Colloids also are used in surgical patients, and their effectiveness as volume expanders compared with isotonic crystalloids has long been debated. Due to their molecular weight, they are confined to the intravascular space, and their infusion results in more efficient transient plasma volume expansion. However, under conditions of severe hemorrhagic shock, capillary membrane permeability increases; this permeability permits colloids to enter the interstitial space, which can worsen edema and impair tissue oxygenation. Four major types of colloids are available—albumin, dextrans, hetastarch, and gelatins—and are described by their molecular weight and size in Table 3-13. In a large randomized trial of patients admitted to an intensive care unit with hypovolemic shock, administration of colloid showed no improvement in mortality at 30 days as compared to crystalloid resuscitation. Interestingly, there was a suggestion of improvement of 90-day mortality and more days alive without mechanical ventilation in patients receiving colloid; however, these were not trial primary endpoints and were considered exploratory. Furthermore, a Cochrane Database systematic review on the topic concluded that there is no available evidence from randomized trials to support colloid use over crystalloid to reduce the risk of death following trauma, burns, or surgery. Colloids are markedly more expensive than crystalloids, and certain colloids, such as hydroxyethyl starch, have been associated with increased morbidity including the need for renal replacement therapy. Taken together, the use of colloid for resuscitation of critically ill and surgical patients has limited application.

Correction of Life-Threatening Electrolyte Abnormalities

Sodium

**Hypernatremia** Treatment of hypernatremia usually consists of treatment of the associated water deficit. In hypovolemic patients, volume should be restored with normal saline before the concentration abnormality is addressed. Once adequate volume has been achieved, the water deficit is replaced using a hypertonic fluid such as 5% dextrose, 5% dextrose in one-quarter normal saline, or enterally administered water. The formula used to estimate the amount of water required to correct hypernatremia is as follows:

$$\text{Water deficit (L)} = \frac{\text{serum sodium} - 140}{140} \times \text{TBW}$$

Estimate TBW as 50% of lean body mass in men and 40% in women.
The rate of fluid administration should be titrated to achieve a decrease in serum sodium concentration of no more than 1 mEq/h and 12 mEq/d for the treatment of acute symptomatic hypernatremia. Even slower correction should be undertaken for chronic hypernatremia (0.7 mEq/h) because overly rapid correction can lead to cerebral edema and herniation. The type of fluid used depends on the severity and ease of correction. Oral or enteral replacement is acceptable in most cases, or IV replacement with half- or quarter-normal saline can be used. Caution also should be exercised when using 5% dextrose in water to avoid overly rapid correction. Frequent neurologic evaluation as well as frequent evaluation of serum sodium levels also should be performed. Hypernatremia is less common than hyponatremia, but has a worse prognosis, and is an independent predictor of mortality in critical illness.29

Hyponatremia Most cases of hyponatremia can be treated by free water restriction and, if severe, the administration of sodium. In patients with normal renal function, symptomatic hyponatremia usually does not occur until the serum sodium level is ≤120 mEq/L. If neurologic symptoms are present, 3% normal saline should be used to increase the sodium by no more than 1 mEq/L per hour until the serum sodium level reaches 130 mEq/L or neurologic symptoms are improved. Correction of asymptomatic hyponatremia should increase the sodium level by no more than 0.5 mEq/L per hour to a maximum increase of 12 mEq/L per day, and even more slowly in chronic hyponatremia. The rapid correction of hyponatremia can lead to pontine myelinolysis,30 with seizures, weakness, paresis, akinetic movements, and unresponsiveness, and may result in permanent brain damage and death. Serial magnetic resonance imaging may be necessary to confirm the diagnosis.31

Potassium

Hyperkalemia Treatment options for symptomatic hyperkalemia are listed in Table 3-14. The goals of therapy include reducing the total body potassium, shifting potassium from the extracellular to the intracellular space, and protecting the cells from the effects of increased potassium. For all patients, exogenous sources of potassium should be removed, including potassium supplementation in IV fluids and enteral and parenteral solutions. Potassium can be removed from the body using a cation-exchange resin such as Kayexalate that binds potassium in exchange for sodium. It can be administered either orally, in alert patients, or rectally. Immediate measures also should include attempts to shift potassium intracellularly with glucose and bicarbonate infusion. Nebulized albuterol (10 to 20 mg) may also be used. Use of glucose alone will cause a rise in insulin secretion, but in the acutely ill, this response may be blunted, and therefore both glucose and insulin may be necessary. Circulatory overload and hypernatremia may result from the administration of Kayexalate and bicarbonate, so care should be exercised when administering these agents to patients with fragile cardiac function. When ECG changes are present, calcium chloride or calcium gluconate (5 to 10 mL of 10% solution) should be administered immediately to counteract the myocardial effects of hyperkalemia. Calcium infusion should be used cautiously in patients receiving digitalis because digitalis toxicity may be precipitated. All of the aforementioned measures are temporary, lasting from 1 to approximately 4 hours. Dialysis should be considered in severe hyperkalemia when conservative measures fail.

Hypokalemia Treatment for hypokalemia consists of potassium repletion, the rate of which is determined by the symptoms (Table 3-15). Oral repletion is adequate for mild, asymptomatic hypokalemia. If IV repletion is required, usually no more than 10 mEq/h is advisable in an unmonitored setting. This amount can be increased to 40 mEq/h when accompanied by continuous ECG monitoring, and even more in the case of imminent cardiac arrest from a malignant arrhythmia-associated hypokalemia. Caution should be exercised when oliguria or impaired renal function is coexistent.

Calcium

Hypercalcemia Treatment is required when hypercalcemia is symptomatic, which usually occurs when the serum level exceeds 12 mg/dL. The critical level for serum calcium is 15 mg/dL, when symptoms noted earlier may rapidly progress to death. The initial treatment is aimed at repleting the associated volume deficit and then inducing a brisk diuresis with normal saline. Treatment of hypercalcemia associated with malignancies is discussed later in this chapter.

Hypocalcemia Asymptomatic hypocalcemia can be treated with oral or IV calcium (see Table 3-15). Acute symptomatic hypocalcemia should be treated with IV 10% calcium gluconate to achieve a serum concentration of 7 to 9 mg/dL. Associated deficits in magnesium, potassium, and pH must also be corrected. Hypocalcemia will be refractory to treatment if coexisting hypomagnesemia is not corrected first. Routine calcium supplementation is no longer recommended in association with massive blood transfusions.32

Phosphorus

Hyperphosphatemia Phosphate binders such as sucralfate or aluminum-containing antacids can be used to lower serum phosphorus levels. Calcium acetate tablets also are useful when hypocalcemia is simultaneously present. Dialysis usually is reserved for patients with renal failure.

Hypophosphatemia Depending on the level of depletion and tolerance to oral supplementation, a number of enteral and parenteral repletion strategies are effective for the treatment of hypophosphatemia (see Table 3-15).

Magnesium

Hypermagnesemia Treatment for hypermagnesemia consists of measures to eliminate exogenous sources of magnesium,
Table 3-15
Electrolyte replacement therapy protocol

**Potassium**
Serum potassium level <4.0 mEq/L:
- Asymptomatic, tolerating enteral nutrition: KCl 40 mEq per enteral access × 1 dose
- Asymptomatic, not tolerating enteral nutrition: KCl 20 mEq IV q2h × 2 doses
- Symptomatic: KCl 20 mEq IV q1h × 4 doses
Recheck potassium level 2 h after end of infusion; if <3.5 mEq/L and asymptomatic, replace as per above protocol

**Magnesium**
Magnesium level 1.0–1.8 mEq/L:
- Magnesium sulfate 0.5 mEq/kg in normal saline 250 mL infused IV over 24 h × 3 d
Recheck magnesium level in 3 d
Magnesium level <1.0 mEq/L:
- Magnesium sulfate 1 mEq/kg in normal saline 250 mL infused IV over 24 h × 1 d, then 0.5 mEq/kg in normal saline 250 mL infused IV over 24 h × 2 d
Recheck magnesium level in 3 d
If patient has gastric access and needs a bowel regimen:
- Milk of magnesia 15 mL (approximately 49 mEq magnesium) q24h per gastric tube; hold for diarrhea

**Calcium**
Ionized calcium level <4.0 mg/dL:
- With gastric access and tolerating enteral nutrition: Calcium carbonate suspension 1250 mg/5 mL q6h per gastric access; recheck ionized calcium level in 3 d
- Without gastric access or not tolerating enteral nutrition: Calcium gluconate 2 g IV over 1 h × 1 dose; recheck ionized calcium level in 3 d

**Phosphate**
Phosphate level 1.0–2.5 mg/dL:
- Tolerating enteral nutrition: Neutra-Phos 2 packets q6h per gastric tube or feeding tube
- No enteral nutrition: KPHO₄ or NaPO₄ 0.15 mmol/kg IV over 6 h × 1 dose
Recheck phosphate level in 3 d
Phosphate level <1.0 mg/dL:
- Tolerating enteral nutrition: KPHO₄ or NaPO₄ 0.25 mmol/kg over 6 h × 1 dose
Recheck phosphate level 4 h after end of infusion; if <2.5 mg/dL, begin Neutra-Phos 2 packets q6h
- Not tolering enteral nutrition: KPHO₄ or NaPO₄ 0.25 mmol/kg (LBW) over 6 h × 1 dose; recheck phosphate level 4 h after end of infusion; if <2.5 mg/dL, then KPHO₄ or NaPO₄ 0.15 mmol/kg (LBW) IV over 6 h × 1 dose

3 mmol KPHO₄ = 3 mmol Phos and 4.4 mEq K⁺ = 1 mL
3 mmol NaPO₄ = 3 mmol Phos and 4 mEq Na⁺ = 1 mL
Neutra-Phos 1 packet = 8 mmol Phos, 7 mEq K⁺, 7 mEq Na⁺
Use patient’s lean body weight (LBW) in kilograms for all calculations.
Disregard protocol if patient has renal failure, is on dialysis, or has a creatinine clearance <30 mL/min.

Correct concurrent volume deficits, and correct acidosis if present. To manage acute symptoms, calcium chloride (5 to 10 mL) should be administered to immediately antagonize the cardiovascular effects. If elevated levels or symptoms persist, hemodialysis may be necessary.

**Hypomagnesemia**
Correction of magnesium depletion can be oral if asymptomatic and mild. Otherwise, IV repletion is indicated and depends on severity (see Table 3-15) and clinical symptoms. For those with severe deficits (<1.0 mEq/L) or those who are symptomatic, 1 to 2 g of magnesium sulfate may be administered IV over 15 minutes. Under ECG monitoring, it may be given over 2 minutes if necessary to correct torsades de pointes (irregular ventricular rhythm). Caution should be taken when giving large amounts of magnesium because magnesium toxicity may develop. The simultaneous administration of calcium gluconate will counteract the adverse side effects of a rapidly rising magnesium level and correct hypocalcemia, which is frequently associated with hypomagnesemia.

**Preoperative Fluid Therapy**
The administration of maintenance fluids should be all that is required in an otherwise healthy individual who may be under orders to receive nothing by mouth for some period before the time of surgery. This does not, however, include replenishment of a preexisting deficit or ongoing fluid losses. The following is a frequently used formula for calculating the volume of maintenance fluids in the absence of preexisting abnormalities:

<table>
<thead>
<tr>
<th>Weight Range</th>
<th>Fluid Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>For the first 0–10 kg</td>
<td>Give 100 mL/kg per day</td>
</tr>
<tr>
<td>For the next 10–20 kg</td>
<td>Give an additional 50 mL/kg per day</td>
</tr>
<tr>
<td>For weight &gt;20 kg</td>
<td>Give an additional 20 mL/kg per day</td>
</tr>
</tbody>
</table>

For example, a 60-kg female would receive a total of 2300 mL of fluid daily: 1000 mL for the first 10 kg of body
weight (10 kg × 100 mL/kg per day), 500 mL for the next 20 kg (10 kg × 50 mL/kg per day), and 800 mL for the last 40 kg (40 kg × 20 mL/kg per day).

An alternative approach is to replace the calculated daily water losses in urine, stool, and insensible loss with a hypotonic saline solution rather than water alone, which allows the kidney some sodium excess to adjust for concentration. Although there should be no “routine” maintenance fluid orders, both of these methods would yield an appropriate choice of 5% dextrose in 0.45% sodium chloride at 100 mL/h as initial therapy, with potassium added for patients with normal renal function. However, many surgical patients have volume and/or electrolyte abnormalities associated with their surgical disease. Preoperative evaluation of a patient’s volume status and preexisting electrolyte abnormalities is an important part of overall preoperative assessment and care. Volume deficits should be considered in patients who have obvious GI losses, such as through emesis or diarrhea, as well as in patients with poor oral intake secondary to their disease. Less obvious are those fluid losses known as third-space or nonfunctional ECF losses that occur with GI obstruction, peritoneal or bowel inflammation, ascites, crush injuries, burns, and severe soft tissue infections such as necrotizing fasciitis. The diagnosis of an acute volume deficit is primarily clinical (see Table 3-2), although the physical signs may vary with the duration of the deficit. Cardiovascular signs of tachycardia and orthostasis predominate with acute volume loss, usually accompanied by oliguria and hemoconcentration. Acute volume deficits should be corrected as much as possible before the time of operation.

Once a volume deficit is diagnosed, prompt fluid replacement should be instituted, usually with an isotonic crystalloid, depending on the measured serum electrolyte values. Patients with cardiovascular signs of volume deficit should receive a bolus of 1 to 2 L of isotonic fluid followed by a continuous infusion. Close monitoring during this period is imperative. Resuscitation should be guided by the reversal of the signs of volume deficit, such as restoration of acceptable values for vital signs, maintenance of adequate urine output (0.5 to 1 mL/kg per hour in an adult), and correction of base deficit. Patients whose volume deficit is not corrected after this initial volume challenge and those with impaired renal function and the elderly should be considered for more intensive monitoring in an intensive care unit setting. In these patients, early invasive monitoring of central venous pressure or cardiac output may be necessary.

If symptomatic electrolyte abnormalities accompany volume deficit, the abnormality should be corrected to the point that the acute symptom is relieved before surgical intervention. For correction of severe hypernatremia associated with a volume deficit, an unsafe rapid fall in extracellular osmolarity from 5% dextrose infusion is avoided by slowly correcting the hypernatremia with 0.45% saline or even lactated Ringer’s solution rather than 5% dextrose alone. This will safely and slowly correct the hypernatremia while also correcting the associated volume deficit.

Intraoperative Fluid Therapy

With the induction of anesthesia, compensatory mechanisms are lost, and hypotension will develop if volume deficits are not appropriately corrected before the time of surgery. Hemodynamic instability during anesthesia is best avoided by correcting known fluid losses, replacing ongoing losses, and providing adequate maintenance fluid therapy preoperatively. In addition to measured blood loss, major open abdominal surgeries are associated with continued extracellular losses in the form of bowel wall edema, peritoneal fluid, and the wound edema during surgery. Large soft tissue wounds, complex fractures with associated soft tissue injury, and burns are all associated with additional third-space losses that must be considered in the operating room. These represent distributional shifts, in that the functional volume of ECF is reduced but fluid is not externally lost from the body. These functional losses have been referred to as parasitic losses, sequestration, or third-space edema because the lost volume no longer participates in the normal functions of the ECF.

Until the 1960s saline solutions were withheld during surgery. Administered saline was retained and was felt to be an inappropriate challenge to a physiologic response of intraoperative salt intolerance. Basic and clinical research began to change this concept, eventually leading to the current concept that saline administration is necessary to restore the obligate ECF losses noted earlier. Although no accurate formula can predict intraoperative fluid needs, replacement of ECF during surgery often requires 500 to 1000 mL/h of a balanced salt solution to support homeostasis. The addition of albumin or other colloid-containing solutions to intraoperative fluid therapy is not necessary. Manipulation of colloidal oncotic forces by albumin infusion during major vascular surgery showed no advantage in supporting cardiac function or avoiding the accumulation of extravascular lung water.

Postoperative Fluid Therapy

Postoperative fluid therapy should be based on the patient’s current estimated volume status and projected ongoing fluid losses. Any deficits from either preoperative or intraoperative losses should be corrected, and ongoing requirements should be included along with maintenance fluids. Third-space losses, although difficult to measure, should be included in fluid replacement strategies. In the initial postoperative period, an isotonic solution should be administered. The adequacy of resuscitation should be guided by the restoration of acceptable values for vital signs and urine output and, in more complicated cases, by the correction of base deficit or lactate. Adjuncts to assessing volume status in the postoperative patient include such tools as a straight leg raise, point-of-care ultrasound, and assessment of respiratory variation via use of an arterial catheter in a mechanically ventilated patient. After the initial 24 to 48 hours, fluids can be changed to 5% dextrose in 0.45% saline in patients unable to tolerate enteral nutrition. If normal renal function and adequate urine output are present, potassium may be added to the IV fluids. Daily fluid orders should begin with assessment of the patient’s volume status and assessment of electrolyte abnormalities. There is rarely a need to check electrolyte levels in the first few days of an uncomplicated postoperative course. However, postoperative diuresis may require attention to replacement of urinary potassium loss. All measured losses, including losses through vomiting, nasogastric suctioning, drains, and urine output, as well as insensible losses, are replaced with the appropriate parenteral solutions as previously reviewed.

Fluid Management in Enhanced Recovery After Surgery (ERAS) Pathways

As pioneered by the Danish surgeon Henrik Kehlet, ERAS pathways have been designed to guide the perioperative management
of various types of surgical procedures. ERAS consists of a multimodal strategy to maximize and maintain preoperative organ function, and implementation of ERAS protocols has resulted in a decrease in length of stay, improved patient satisfaction, cost savings, and a reduction in complications. A full discussion of ERAS is outside the scope of this chapter; however, it is important to note that perioperative fluid management is a major tenet of ERAS protocols. The 2011 European Society of Anaesthesiology guidelines were among the first formal recommendations to alter standard recommendations for preoperative enteral intake. These recommendations include allowance of clear liquids up to 2 hours prior to surgery. Many ERAS protocols include the use of carbohydrate and electrolyte-rich fluids to enhance hydration and metabolic response to surgery. In addition to preoperative enteral hydration, a major focus of ERAS protocols is the restriction of intra- and postoperative sodium and intravenous fluids. Fluid overload has been associated with prolonged ileus and coagulation abnormalities. Goal-directed fluid therapy has been shown to reduce postoperative morbidity and length of stay independent of the other multimodal components of ERAS, making minimizing fluids a major target of intervention. Postoperatively, early enteral intake is advised, with prompt discontinuation of intravenous fluids. These strategies targeting euvolemia have been shown to be safe and improve outcomes, making ERAS a rapidly evolving strategy that will continue to influence the perioperative fluid and electrolyte management of surgical patients.

Special Considerations for the Postoperative Patient

Volume excess is a common disorder in the postoperative period. The administration of isotonic fluids in excess of actual needs may result in excess volume expansion. This may be due to the overestimation of third-space losses or to ongoing GI losses that are difficult to measure accurately. The earliest sign of volume overload is weight gain. The average postoperative patient who is not receiving nutritional support should lose approximately 0.25 to 0.5 lb/d (0.11 to 0.23 kg/d) from catabolism. Additional signs of volume excess may also be present as listed in Table 3-2. Peripheral edema may not necessarily be associated with intravascular volume overload because overexpansion of total ECF may exist in association with a deficit in the circulating plasma volume.

Volume deficits also can be encountered in surgical patients if preoperative losses were not completely corrected, intraoperative losses were underestimated, or postoperative losses were greater than appreciated. The clinical manifestations are described in Table 3-2 and include tachycardia, orthostasis, and oliguria. Hemoconcentration also may be present. Treatment will depend on the amount and composition of fluid lost. In most cases of volume depletion, replacement with an isotonic fluid will be sufficient while alterations in concentration and composition are being evaluated.

ELECTROLYTE ABNORMALITIES IN SPECIFIC SURGICAL PATIENTS

Neurologic Patients

Syndrome of Inappropriate Secretion of Antidiuretic Hormone. The syndrome of inappropriate secretion of antidiuretic hormone (SIADH) can occur after head injury or surgery to the central nervous system, but it also is seen in association with administration of drugs such as morphine, nonsteroidals, and oxytocin, and in a number of pulmonary and endocrine diseases, including hypothyroidism and glucocorticoid deficiency. Additionally, it can be seen in association with a number of malignancies, most often small cell cancer of the lung but also pancreatic carcinoma, thymoma, and Hodgkin’s disease.

SIADH should be considered in patients who are euvoletic and hyponatremic with elevated urine sodium levels and urine osmolality. ADH secretion is considered inappropriate when it is not in response to osmotic or volume-related conditions. Correction of the underlying problem should be attempted when possible. In most cases, restriction of free water will improve the hyponatremia. The goal is to achieve net water balance while avoiding volume depletion that may compromise renal function. Furosemide also can be used to induce free water loss. If hyponatremia persists after fluid restriction, the addition of isotonic or hypertonic fluids may be effective. The administration of isotonic saline may sometimes worsen the problem if the urinary sodium concentration is higher than the infused sodium concentration. The use of loop diuretics may be helpful in this situation by preventing further urine concentration. In chronic SIADH, when long-term fluid restriction is difficult to maintain or is ineffective, demeclocycline and lithium can be used to induce free water loss.

Diabetes Insipidus. Diabetes insipidus (DI) is a disorder of ADH stimulation and is manifested by dilute urine in the case of hypernatremia. Central DI results from a defect in ADH secretion, and nephrogenic DI results from a defect in end-organ responsiveness to ADH. Central DI is frequently seen in association with pituitary surgery, closed head injury, and anoxic encephalopathy. Nephrogenic DI occurs in association with hypokalemia, administration of radiocontrast dye, and use of certain drugs such as aminoglycosides and amphotericin B. In patients tolerating oral intake, volume status usually is normal because thirst stimulates increased intake. However, volume depletion can occur rapidly in patients who are incapable of oral intake. The diagnosis can be confirmed by documenting a paradoxical increase in urine osmolality in response to a period of water deprivation. In mild cases, free water replacement may be adequate therapy. In more severe cases, vasopressin can be added. The usual dosage of vasopressin is 5 U subcutaneously every 6 to 8 hours. However, serum electrolytes and osmolality should be monitored to avoid excess vasopressin administration with resulting iatrogenic SIADH.

Cerebral Salt Wasting. Cerebral salt wasting is a diagnosis of exclusion that occurs in patients with a cerebral lesion and renal wasting of sodium and chloride with no other identifiable cause. Natriuresis in a patient with a contracted extracellular volume should prompt the possible diagnosis of cerebral salt wasting. Hyponatremia is frequently observed but is nonspecific and occurs as a secondary event, which differentiates it from SIADH.

Malnourished Patients: Refeeding Syndrome

Refeeding syndrome is a potentially lethal condition that can occur with rapid and excessive feeding of patients with severe underlying malnutrition due to starvation, alcoholism, delayed nutritional support, anorexia nervosa, or massive weight loss in obese patients. With refeeding, a shift in metabolism from fat to carbohydrate substrate stimulates insulin release, which
results in the cellular uptake of electrolytes, particularly phosphate, magnesium, potassium, and calcium. However, severe hyperglycemia may result from blunted basal insulin secretion. The refeeding syndrome can be associated with enteral or parenteral refeeding, and symptoms from electrolyte abnormalities include cardiac arrhythmias, confusion, respiratory failure, and even death. To prevent the development of refeeding syndrome, underlying electrolyte and volume deficits should be corrected. Additionally, thiamine should be administered before the initiation of feeding. Caloric repletion should be instituted slowly and should gradually increase over the first week. Vital signs, fluid balance, and electrolytes should be closely monitored and any deficits corrected as they evolve.

**Acute Renal Failure Patients**

A number of fluid and electrolyte abnormalities are specific to patients with acute renal failure. With the onset of renal failure, an accurate assessment of volume status must be made. If prerenal azotemia is present, prompt correction of the underlying volume deficit is mandatory. Once acute tubular necrosis is established, measures should be taken to restrict daily fluid intake to match urine output and insensible and GI losses. Oliguric renal failure requires close monitoring of serum potassium levels. Measures to correct hyperkalemia as reviewed in Table 3-14 should be instituted early, including consideration of early hemodialysis. Hyponatremia is common in established renal failure as a result of the breakdown of proteins, carbohydrates, and fats, as well as the administration of free water. Dialysis may be required for severe hyponatremia. Hypocalcemia, hypomagnesemia, and hyperphosphatemia also are associated with acute renal failure. Hypocalcemia should be verified by measuring ionized calcium, because many patients also are hypoalbuminemic. Phosphate binders can be used to control hyperphosphatemia, but dialysis may be required in more severe cases. Metabolic acidosis is commonly seen with renal failure, as the kidneys lose their ability to clear acid by-products. Bicarbonate can be useful, but dialysis often is needed. Although dialysis may be either intermittent or continuous, renal recovery may be improved by continuous renal replacement.

**Cancer Patients**

Fluid and electrolyte abnormalities are common in patients with cancer. The causes may be common to all patient populations or may be specific to cancer patients and their treatment. Hypotension is frequently hypovolemic due to renal loss of sodium caused by diuretics or salt-wasting nephropathy as seen with some chemotherapeutic agents such as cisplatin. Cerebral salt wasting also can occur in patients with intracerebral lesions. Normovolemic hyponatremia may occur in association with SIADH from cervical cancer, lymphoma, and leukemia, or from certain chemotherapeutic agents. Hypernatremia in cancer patients most often is due to poor oral intake or GI volume losses, which are common side effects of chemotherapy. Central DI also can lead to hypernatremia in patients with central nervous system lesions.

Hypokalemia can develop from GI losses associated with diarrhea caused by radiation enteritis or chemotherapy, or from tumors such as villous adenomas of the colon. Tumor lysis syndrome can precipitate severe hyperkalemia from massive tumor cell destruction.

Hypocalcemia can be seen after removal of a thyroid or parathyroid tumor or after a central neck dissection, which can damage the parathyroid glands. Hungry bone syndrome produces acute and profound hypocalcemia after parathyroid surgery for secondary or tertiary hyperparathyroidism because calcium is rapidly taken up by bones. Prostate and breast cancer can result in increased osteoblastic activity, which decreases serum calcium by increasing bone formation. Acute hypocalcemia also can occur with hyperphosphatemia because phosphorus complexes with calcium. Hypomagnesemia is a side effect of ifosfamide and cisplatin therapy. Hyperphosphatemia can be seen in hyperparathyroidism, due to decreased phosphorus reabsorption, and in oncogenic osteomalacia, which increases the urinary excretion of phosphorus. Other causes of hyperphosphatemia in cancer patients include renal tubular dysfunction from multiple myeloma, Bence Jones proteins, and certain chemotherapeutic agents. Acute hyperphosphatemia can occur as rapidly proliferating malignant cells take up phosphorus in acute leukemia. Tumor lysis syndrome or the use of bisphosphonates to treat hypercalcemia also can result in hyperphosphatemia.

Malignancy is the most common cause of hypercalcemia in hospitalized patients and is due to increased bone resorption or decreased renal excretion. Bone destruction occurs from bony metastasis as seen in breast or renal cell cancer but also can occur in multiple myeloma. With Hodgkin’s and non-Hodgkin’s lymphoma, hypercalcemia results from increased calcitriol formation, which increases both absorption of calcium from the GI tract and mobilization from bone. Humoral hypercalcemia of malignancy is a common cause of hypercalcemia in cancer patients. As in primary hyperparathyroidism, a parathyroid-related protein is secreted that binds to parathyroid receptors, stimulating calcium resorption from bone and decreasing renal excretion of calcium. The treatment of hypercalcemia of malignancy should begin with saline volume expansion, which will decrease renal reabsorption of calcium as the associated volume deficit is corrected. Once an adequate volume status has been achieved, a loop diuretic may be added. Unfortunately, these measures are only temporary, and additional treatment is often necessary. A variety of drugs are available with varying times of onset, durations of action, and side effects. The bisphosphonates etidronate and pamidronate inhibit bone resorption and osteoclastic activity. They have a slow onset of action, but effects can last for 2 weeks. Calcitonin also is effective, inhibiting bone resorption and increasing renal excretion of calcium. It acts quickly, within 2 to 4 hours, but its use is limited by the development of tachyphylaxis. Corticosteroids may decrease tachyphylaxis in response to calcitonin and can be used alone to treat hypercalcemia. Gallium nitrates are potent inhibitors of bone resorption. They display a long duration of action but can cause nephrotoxicity. Mitomycin is an antibiotic that blocks osteoclastic activity, but it can be associated with liver, renal, and hematologic abnormalities, which limits its use to the treatment of Paget’s disease of bone. For patients with severe, refractory hypercalcemia who are unable to tolerate volume expansion due to pulmonary edema or congestive heart failure, dialysis is an option.

Tumor lysis syndrome results when the release of intracellular metabolites overwhelms the kidneys’ excretory capacity. This rapid release of uric acid, potassium, and phosphorus can result in marked hyperuricemia, hyperkalemia, hyperphosphatemia, hypocalcemia, and acute renal failure. It is typically seen with poorly differentiated lymphomas and leukemias but also can occur with a number of solid tumor malignancies. Tumor lysis syndrome most commonly develops during treatment.
with chemotherapy or radiotherapy. Once it develops, volume expansion should be undertaken and any associated electrolyte abnormalities corrected. In this setting, hypocalcaemia should not be treated unless it is symptomatic to avoid metastatic calcifications. Dialysis may be required for management of impaired renal function or correction of electrolyte abnormalities.

REFERENCES

Entries highlighted in bright blue are key references.


This meta-analysis demonstrated no mortality benefit or effect on intracranial pressure control utilizing hypertonic saline and was a key summary of eleven major trials in the field.


This multi-center randomized trial found no difference in mortality at 28 days in a mixed ICU population of patients with hypovolemia when comparing colloid to crystalloid resuscitation. The authors observed a reduction in 90-day mortality in the group receiving colloids; however, they cautioned that this observation was exploratory and required further analysis.


This Cochrane review of 70 trials including mortality data found no benefit to colloid resuscitation for reduction of risk of mortality, while the use of hydroxystethyl starch (HES) was found to have a possible association with increased mortality. The recommendation of the review was that colloid use in the ICU could not be supported due to increased cost and no mortality benefit.


In a randomized trial designed to address a primary outcome of 90-day mortality, the use of HES revealed no mortality benefit over saline with a statistically significant increase in the rate of renal replacement therapy after HES use.


This early paper by Shires and colleagues produced important early observations of fluid shifts and resuscitation strategies following surgery and shaped some of the early questions in the field of resuscitation science after surgery.


Index guidelines published by the European Society of Anaesthesiology were among the first to support the use of ERAS principles and provided a template for the design of the rapidly evolving practice of ERAS worldwide.


Meta-analysis of randomized controlled trials demonstrated a marked reduction in morbidity, hospital length of stay (LOS),
ICU LOS, and time to passage of feces utilizing goal-directed fluid therapy (GDFT) for intraoperative fluid management. Of note, in those studies where a multimodal ERAS pathway was utilized, the benefits of GDFT were reduced, although reduction in major morbidities remained. These data suggest that the multimodal benefits of ERAS may extend beyond just fluid management.


This page intentionally left blank
Hemostasis, Surgical Bleeding, and Transfusion

Ronald Chang, John B. Holcomb, Evan Leibner, Matthew Pommerening, and Rosemary A. Kozar

**BIOLOGY OF HEMOSTASIS**

Hemostasis is a complex process whose function is to limit blood loss from an injured vessel. Four major physiologic events participate in the hemostatic process: vascular constriction, platelet plug formation, fibrin formation, and fibrinolysis. Although each tends to be activated in order, the four processes are interrelated so that there is a continuum and multiple reinforcements. The process is shown schematically in Fig. 4-1.

**Vascular Constriction**

Vascular constriction is the initial response to vessel injury. It is more pronounced in vessels with medial smooth muscles and is dependent on local contraction of smooth muscle. Vasoconstriction is subsequently linked to platelet plug formation. Thromboxane A₂ (TXA₂) is produced locally at the site of injury via the release of arachidonic acid from platelet membranes and is a potent constrictor of smooth muscle. Similarly, endothelin synthesized by injured endothelium and serotonin (5-hydroxytryptamine [5-HT]) released during platelet aggregation are potent vasoconstrictors. Lastly, bradykinin and fibrinopeptides, which are involved in the coagulation schema, are also capable of contracting vascular smooth muscle.

The extent of vasoconstriction varies with the degree of vessel injury. A small artery with a lateral incision may remain open due to physical forces, whereas a similarly sized vessel that is completely transected may contract to the extent that bleeding ceases spontaneously.

**Platelet Function**

Platelets are anucleate fragments of megakaryocytes. The normal circulating number of platelets ranges between 150,000 and 400,000/μL. Up to 30% of circulating platelets may be sequestered in the spleen. If not consumed in a clotting reaction, platelets are normally removed by the spleen and have an average life span of 7 to 10 days.

Platelets play an integral role in hemostasis by forming a hemostatic plug and by contributing to thrombin formation (Fig. 4-2). Platelets do not normally adhere to each other or to the vessel wall but can form a plug that aids in cessation of bleeding when vascular disruption occurs. Injury to the intimal layer in the vascular wall exposes subendothelial collagen to which platelets adhere. This process requires von Willebrand factor (vWF), a protein in the subendothelium that is lacking in patients with von Willebrand’s disease. vWF binds to glycoprotein (GP) I/IX/V on the platelet membrane. Following adhesion, platelets initiate a release reaction that recruits other platelets from the circulating blood to seal the disrupted vessel. Up to this point, this process is known as primary hemostasis. Platelet aggregation is reversible and is not associated with secretion. Additionally, heparin does not interfere with this reaction, and thus, hemostasis can occur in the heparinized patient. Adenosine diphosphate (ADP) and serotonin are the principal mediators in platelet aggregation.

Arachidonic acid released from the platelet membranes is converted by cyclooxygenase to prostaglandin G₂ (PGG₂) and then to prostaglandin H₂ (PGH₂), which, in turn, is converted to TXA₂. TXA₂ has potent vasoconstriction and platelet aggregation effects. Arachidonic acid may also be shuttled to adjacent endothelial cells and converted to prostacyclin (PGL₂), which is a vasodilator and acts to inhibit platelet aggregation. Platelet cyclooxygenase is irreversibly inhibited by aspirin and reversibly blocked by nonsteroidal anti-inflammatory agents but is not affected by cyclooxygenase-2 (COX-2) inhibitors.

In the second wave of platelet aggregation, a release reaction occurs in which several substances including ADP, Ca²⁺, serotonin, TXA₂, and α-granule proteins are discharged.
The life span of platelets ranges from 7 to 10 days. Drugs that interfere with platelet function include aspirin, clopidogrel, prasugrel, dipyridamole, and the glycoprotein IIb/IIIa (GP IIb/IIIa) inhibitors. Approximately 5 to 7 days should pass from the time the drug is stopped until an elective procedure is performed.

Laboratory evidence of trauma-induced coagulopathy is found in up to one-third of severely injured patients at admission. It is distinct from disseminated intravascular coagulopathy and iatrogenic causes of coagulopathy such as hemodilution. Several non–mutually exclusive mechanisms have been proposed. However, the relationship between laboratory coagulation abnormalities and clinically evident coagulopathic bleeding is unclear.

Direct oral anticoagulants have no readily available method for monitoring anticoagulation. A new monoclonal antibody has been approved to reverse coagulopathy due to dabigatran, and agents are currently in clinical trials for the reversal of direct factor Xa oral anticoagulants.

When determining the need for bridging of therapeutic anticoagulation in the preoperative and postoperative setting, the patient’s risk of bleeding should be carefully considered against the risk of thromboembolism and used to guide the need for reversal of anticoagulation therapy preoperatively and the timing of its reinstatement postoperatively.

Damage control resuscitation has three basic components: permissive hypotension, minimizing crystalloid-based resuscitation, and the administration of balanced ratios of blood products.

The need for massive transfusion should be anticipated, and guidelines should be in place to provide early and balanced amounts of red blood cells, plasma, and platelets.

Fibrinogen is a required cofactor for this process, acting as a bridge for the GP IIb/IIIa receptor on the activated platelets. The release reaction results in compaction of the platelets into a plug, a process that is no longer reversible. Thrombospondin, another protein secreted by the α-granule, stabilizes fibrinogen binding to the activated platelet surface and strengthens the platelet-platelet interactions. Platelet factor 4 (PF4) and α-thromboglobulin are also secreted during the release reaction. PF4 is a potent heparin antagonist. The second wave of platelet aggregation is inhibited by aspirin and nonsteroidal anti-inflammatory drugs, by cyclic adenosine monophosphate (cAMP), and by nitric oxide. As a consequence of the release reaction, alterations occur in the phospholipids of the platelet membrane that allow calcium and clotting factors to bind to the platelet surface, forming enzymatically active complexes. The altered lipoprotein surface (sometimes referred to as platelet factor 3) catalyzes reactions that are involved in the conversion of prothrombin (factor II) to thrombin (factor IIa) by activated factor X (Xa) in the presence of factor V and calcium, and it is involved in the reaction by which activated factor IX (IXa), factor VIII, and calcium activated factor X. Platelets may also play a role in the initial activation of factors XI and XII.

Coagulation

Hemostasis involves a complex interplay and combination of interactions between platelets, the endothelium, and multiple circulating or membrane-bound coagulation factors. While overly simplistic and not reflective of the depth or complexity of these interactions, the coagulation cascade has traditionally been depicted as two possible pathways converging into a single
common pathway (Fig. 4-3). While this pathway reflects the basic process and sequences that lead to the formation of a clot, the numerous feedback loops, endothelial interplay, and platelet functions are not included. The intrinsic pathway begins with the activation of factor XII that subsequently activates factors XI, IX, and VIII. In this pathway, each of the primary factors is “intrinsic” to the circulating plasma, whereby no surface is required to initiate the process. In the extrinsic pathway, tissue factor (TF) is released or exposed on the surface of the endothelium, binding to circulating factor VII, facilitating its activation to VIIa. Each of these pathways continues on to a common sequence that begins with the activation of factor X to Xa (in the presence of VIIIa). Subsequently, Xa (with the help of factor Va) converts factor II (prothrombin) to thrombin and then factor I (fibrinogen) to fibrin. Clot formation occurs after fibrin monomers are cross-linked to polymers with the assistance of factor XIII.

One convenient feature of depicting the coagulation cascade with two merging arms is that commonly used laboratory tests segregate abnormalities of clotting to one of the two arms. An elevated activated partial thromboplastin time (aPTT) is associated with abnormal function of the intrinsic arm of the cascade (II, IX, X, XI, XII), while the prothrombin time (PT) is associated with the extrinsic arm (II, VII, X). Vitamin K deficiency or warfarin use affects factors II, VII, IX, and X.

Expanding from the basic concept of Fig. 4-3, the cell-based model of hemostasis, divided into the initiation, amplification, and propagation phases, provides a more complete picture of clot formation. During initiation, the primary pathway for coagulation is initiated by TF exposure following subendothelial injury. TF binds to VIIa, and this complex catalyzes the activation of factor X to Xa and IX to IXa, which in turn activates factor V to Va. This “prothrombinase” complex generates small amounts of thrombin from prothrombin in a calcium-dependent process. During amplification, platelets adhere to extracellular matrix components at the site of injury and become activated upon exposure to thrombin and other stimuli. Finally, during the propagation phase, “tenase” (factor VIIIa/IXa) and prothrombinase (factor Va/Xa) complexes are assembled on the surfaces of activated platelets. This results in large-scale generation of thrombin (“thrombin burst”) and fibrin.

In building on the redundancy inherent in the coagulation system, factor VIIIa combines with IXa to form the intrinsic factor complex. Factor IXa is responsible for the bulk of the conversion of factor X to Xa. This complex (VIIIa-IXa) is 50 times more effective at catalyzing factor X activation than is the extrinsic (TF-VIIa) complex and five to six orders of magnitude more effective than factor IXa alone.

Once formed, thrombin leaves the membrane surface and converts fibrinogen by two cleavage steps into fibrin and two small peptides termed fibrinopeptides A and B. Removal of fibrinopeptide A permits end-to-end polymerization of the fibrin molecules, whereas cleavage of fibrinopeptide B allows side-to-side polymerization of the fibrin clot. This latter step is
facilitated by thrombin-activatable fibrinolysis inhibitor (TAFI), which acts to stabilize the resultant clot.

In seeking to balance profound bleeding with overwhelming clot burden, several related processes exist to prevent propagation of the clot beyond the site of injury. First, feedback inhibition on the coagulation cascade deactivates the enzyme complexes leading to thrombin formation. Thrombomodulin (TM) presented by the endothelium serves as a “thrombin sink” by forming a complex with thrombin, rendering it no longer available to cleave fibrinogen. This then activates protein C (APC) and reduces further thrombin generation by inhibiting factors V and VIII. Second, tissue plasminogen activator (tPA) is released from the endothelium following injury, cleaving plasminogen to initiate fibrinolysis. APC then consumes plasminogen activator inhibitor-1 (PAI-1), leading to increased tPA activity and fibrinolysis. Building on the anticoagulant response to inhibit thrombin formation, tissue factor pathway inhibitor (TFPI) is released, blocking the TF-VIIa complex and reducing the production of factors Xa and IXa. Antithrombin III (AT-III) then neutralizes all of the procoagulant serine proteases and also inhibits the TF-VIIa complex. The most potent mechanism of thrombin inhibition involves the APC system. APC forms a complex with its cofactor, protein S, on a phospholipid surface. This complex then cleaves factors Va and VIIIa so that they are no longer able to participate in the formation of TF-VIIa or prothrombinase complexes. This is of interest clinically in the form of a genetic mutation, called factor V Leiden. In this setting, factor V is resistant to cleavage by APC, thereby remaining active as a procoagulant. Patients with factor V Leiden are predisposed to venous thromboembolic events.

Degradation of fibrin clot is accomplished by plasmin, a serine protease derived from the proenzyme plasminogen. Plasmin formation occurs as a result of one of several plasminogen activators. tPA is made by the endothelium and other cells of the vascular wall and is the main circulating form of this family of enzymes. tPA is selective for fibrin-bound plasminogen so that endogenous fibrinolytic activity occurs predominately at the site of clot formation. The other major plasminogen activator, urokinase plasminogen activator (uPA), is also produced by endothelial cells as well as by urothelium, is not selective for fibrin-bound plasminogen. Of note, the thrombin-TM complex activates TAFI, leading to a mixed effect on clot stability. In addition to inhibiting fibrinolysis directly, removal of the terminal lysine on the fibrin molecule by TAFI renders the clot more susceptible to lysis by plasmin.

**Fibrinolysis**

Fibrin clot breakdown (lysis) allows restoration of blood flow during the healing process following injury and begins at the same time clot formation is initiated. Fibrin polymers are degraded by plasmin, a serine protease derived from the proenzyme plasminogen. Plasminogen is converted to plasmin by one of several plasminogen activators, including tPA. Plasmin then degrades the fibrin mesh at various places, leading to the production of circulating fragments, termed fibrin degradation products (FDPs), cleared by other proteases or by the kidney and liver (Fig. 4-4). Fibrinolysis is directed by circulating kinases, tissue activators, and kallikrein present in vascular endothelium. tPA is synthesized by endothelial cells and released by the cells on thrombin stimulation. Bradykinin, a potent endothelial-dependent vasodilator, is cleaved from high molecular weight kininogen by kallikrein and enhances the release of tPA. Both tPA and plasminogen bind to fibrin as it forms, and this trimeric complex cleaves fibrin very efficiently. After plasmin is generated, however, it cleaves fibrin somewhat less efficiently.

As with clot formation, fibrinolysis is also kept in check through several robust mechanisms. tPA activates plasminogen more efficiently when it is bound to fibrin, so that plasmin is formed selectively on the clot. Plasmin is inhibited by α2-antiplasmin, a protein that is cross-linked to fibrin by factor XIII, which helps to ensure that clot lysis does not occur too quickly. Any circulating plasmin is also inhibited by α2-antiplasmin and circulating tPA or urokinase. Clot lysis yields FDPs including E-nodules and D-dimers. These smaller fragments interfere with normal platelet aggregation, and the larger fragments may be incorporated into the clot in lieu of normal fibrin monomers. This may result in an unstable clot as seen in cases of severe coagulopathy such as hyperfibrinolysis associated with trauma-induced coagulopathy or disseminated intravascular coagulopathy. The presence of D-dimers in the circulation may serve as a marker of thrombosis or other conditions in which a significant activation of the fibrinolytic system is present. Another inhibitor of the fibrinolytic system is TAFI, which removes lysine residues from fibrin that are essential for binding plasminogen.

**CONGENITAL FACTOR DEFICIENCIES**

**Coagulation Factor Deficiencies**

Inherited deficiencies of all of the coagulation factors are seen. However, the three most frequent are factor VIII deficiency (hemophilia A or von Willebrand’s disease), factor IX deficiency (hemophilia B or Christmas disease), and factor XI deficiency. Hemophilia A and hemophilia B are inherited as sex-linked recessive disorders with males being affected almost exclusively. The clinical severity of hemophilia A and hemophilia B depends on the measurable level of factor VIII or factor IX in the patient’s plasma. Plasma factor levels less than 1% of normal are considered severe disease, factor levels between 1% and 5% moderately severe disease, and levels between 5% and 30% mild disease. Patients with severe hemophilia have spontaneous bleeds, frequently into joints, leading to crippling arthropathies. Intracranial bleeding, intramuscular hematomas, retroperitoneal hematomas, and gastrointestinal, genitourinary, and retropharyngeal bleeding are added clinical sequelae seen with severe disease. Patients with moderately severe hemophilia have less spontaneous bleeding but are likely to bleed severely...
HEMOSTASIS, SURGICAL BLEEDING, AND TRANSFUSION

CHAPTER 4

after trauma or surgery. Mild hemophiliacs do not bleed spontaneously and have only minor bleeding after major trauma or surgery. Since platelet function is normal in hemophiliacs, patients may not bleed immediately after an injury or minor surgery as they have a normal response with platelet activation and formation of a platelet plug. At times, the diagnosis of hemophilia is not made in these patients until after their first minor procedure (e.g., tooth extraction or tonsillectomy).

Patients with hemophilia A or B are treated with factor VIII or factor IX concentrate, respectively. Recombinant factor VIII is strongly recommended for patients not treated previously and is generally recommended for patients who are both human immunodeficiency virus (HIV) and hepatitis C virus (HCV) seronegative. For factor IX replacement, the preferred products are recombinant or high-purity factor IX. In general, activity levels should be restored to 30% to 40% for mild hemorrhage, 50% for severe bleeding, and 80% to 100% for life-threatening bleeding. Up to 20% of hemophiliacs with factor VIII deficiency develop inhibitors that can neutralize FVIII. For patients with low titers, inhibitors can be overcome with higher doses of factor VIII. For patients with high inhibitor titers, alternate treatments should be used and may include porcine factor VIII, prothrombin complex concentrates, activated prothrombin complex concentrates, or recombinant factor VIIIa. For patients undergoing elective surgical procedures, a multidisciplinary approach with preoperative planning and replacement is recommended.2

von Willebrand’s Disease. von Willebrand’s disease (vWD), the most common congenital bleeding disorder, is characterized by a quantitative or qualitative defect in vWF, a large glycoprotein responsible for carrying factor VIII and platelet adhesion. The latter is important for normal platelet adhesion to exposed subendothelium and for aggregation under high shear conditions. Patients with vWD have bleeding that is characteristic of platelet disorders such as easy bruising and mucosal bleeding. Menorrhagia is common in women. vWD is classified into three types. Type I is a partial quantitative deficiency; type II is a qualitative defect; type III is total deficiency. For bleeding, type I patients usually respond well to desmopressin (DDAVP). Type II patients may respond, depending on the particular defect. Type III patients are usually unresponsive. These patients may require vWF concentrates.3

Factor XI Deficiency. Factor XI deficiency, an autosomal recessive inherited condition sometimes referred to as hemophilia C, is more prevalent in the Ashkenazi Jewish population but found in all races. Spontaneous bleeding is rare, but bleeding may occur after surgery, trauma, or invasive procedures. Treatment of patients with factor XI deficiency who present with bleeding or in whom surgery is planned and who are known to have bled previously is with fresh frozen plasma (FFP). Each milliliter of plasma contains 1 unit of factor XI activity, so the volume needed depends on the patient’s baseline level, the desired level, and the plasma volume. Antifibrinolitics may be useful in patients with menorrhagia. Factor VIIa is recommended for patients with anti-factor XI antibodies, although thrombosis has been reported.4 There has been renewed interest in factor XI inhibitors as antithrombotic agents because patients with factor XI deficiency generally have only minimal bleeding risk unless a severe deficiency is present and seem to be protected from thrombosis.5

Deficiency of Factors II (Prothrombin), V, and X. Inherited deficiencies of factors II, V, and X are rare. These deficiencies are inherited as autosomal recessive. Significant bleeding in homozygotes with less than 1% of normal activity is encountered. Bleeding with any of these deficiencies is treated with FFP. Similar to factor XI, FFP contains one unit of activity of each per milliliter. However, factor V activity is decreased because of its inherent instability. The half-life of prothrombin (factor II) is long (approximately 72 hours), and only about 25% of a normal level is needed for hemostasis. Prothrombin complex concentrates can be used to treat deficiencies of prothrombin or factor X. Daily infusions of FFP are used to treat bleeding in factor V deficiency, with a goal of 20% to 25% activity. Factor V deficiency may be co-inherited with factor VIII deficiency. Treatment of bleeding in individuals with the combined deficiency requires factor VIII concentrate and FFP. Some patients with factor V deficiency are also lacking the factor V normally present in platelets and may need platelet transfusions as well as FFP.

Factor VII Deficiency. Inherited factor VII deficiency is a rare autosomal recessive disorder. Clinical bleeding can vary widely and does not always correlate with the level of FVII coagulant activity in plasma. Bleeding is uncommon unless the level is less than 3%. The most common bleeding manifestations involve easy bruising and mucosal bleeding, particularly epistaxis or oral mucosal bleeding. Postoperative bleeding is also common, reported in 30% of surgical procedures.6 Treatment is with FFP or recombinant factor VIIa. The half-life of recombinant factor VIIa is only approximately 2 hours, but excellent hemostasis can be achieved with frequent infusions. The half-life of factor VII in FFP is up to 4 hours.

Factor XIII Deficiency. Congenital factor XIII (FXIII) deficiency, originally recognized by Duckert in 1960, is a rare autosomal recessive disease usually associated with a severe bleeding diathesis.7 The male-to-female ratio is 1:1. Although acquired FXIII deficiency has been described in association with hepatic failure, inflammatory bowel disease, and myeloid leukemia, the only significant association with bleeding in children is the inherited deficiency.8 Bleeding is typically delayed because clots form normally but are susceptible to fibrinolysis. Umbilical stump bleeding is characteristic, and there is a high risk of intracranial bleeding. Spontaneous abortion is usual in women with factor XIII deficiency unless they receive replacement therapy. Replacement can be accomplished with FFP, cryoprecipitate, or a factor XIII concentrate. Levels of 1% to 2% are usually adequate for hemostasis.

Platelet Functional Defects

Inherited platelet functional defects include abnormalities of platelet surface proteins, abnormalities of platelet granules, and enzyme defects. The major surface protein abnormalities are thrombasthenia and Bernard-Soulier syndrome. Thrombasthe-nia, or Glanzmann thrombasthenia, is a rare genetic platelet disorder, inherited in an autosomal recessive pattern, in which the platelet glycoprotein Ib/IIa (GP Ib/IIa) complex is either lacking or present but dysfunctional. This defect leads to faulty platelet aggregation and subsequent bleeding. The disorder was first described by Dr. Eduard Glanzmann in 1918.8 Bleeding in thrombasthenic patients must be treated with platelet transfusions. Bernard-Soulier syndrome is caused by a defect in the GP Ib/IX/V receptor for vWF, which is necessary for platelet adhesion to the subendothelium. Transfusion of normal platelets is required for bleeding in these patients.
The most common intrinsic platelet defect is storage pool disease. It involves loss of dense granules (storage sites for ADP, adenosine triphosphate [ATP], Ca²⁺, and inorganic phosphate) and α-granules. Dense granule deficiency is the most prevalent of these. It may be an isolated defect or occur with partial albinism in Hermansky-Pudlak syndrome. Bleeding is variable, depending on the severity of the granule defect. Bleeding is caused by the decreased release of ADP from these platelets. A few patients have been reported who have decreased numbers of both dense and α-granules. They have a more severe bleeding disorder. Patients with mild bleeding as a consequence of a form of storage pool disease can be treated with DDAVP. It is likely that the high levels of vWF in the plasma after DDAVP somehow compensate for the intrinsic platelet defect. With more severe bleeding, platelet transfusion is required.

**ACQUIRED HEMOSTATIC DEFECTS**

**Platelet Abnormalities**

Acquired congenital abnormalities of platelets are much more common than acquired defects and may be quantitative or qualitative, although some patients have both types of defects. Quantitative defects may be a result of failure of production, shortened survival, or sequestration. Failure of production is generally a result of bone marrow disorders such as leukemia, myelodysplastic syndrome, severe vitamin B₁₂ or folate deficiency, chemotherapy drugs, radiation, acute ethanol intoxication, or viral infection. If a quantitative abnormality exists and treatment is indicated either due to symptoms or the need for an invasive procedure, platelet transfusion is utilized. The etiologies of both qualitative and quantitative defects are reviewed in Table 4-1.

**Quantitative Defects.** Shortened platelet survival is seen in immune thrombocytopenia, disseminated intravascular coagulation, or disorders characterized by platelet thrombi such as thrombotic thrombocytopenic purpura and hemolytic uremic syndrome. Immune thrombocytopenia may be idiopathic or associated with other autoimmune disorders or low-grade B-cell malignancies, and it may also be secondary to viral infections (including HIV) or drugs. Secondary immune thrombocytopenia often presents with a very low platelet count, petechiae and purpura, and epistaxis. Large platelets are seen on peripheral smear. Initial treatment consists of corticosteroids, intravenous gamma globulin, or anti-D immunoglobulin in patients who are Rh positive. Both gamma globulin and anti-D immunoglobulin are rapid in onset. Platelet transfusions are not usually needed unless central nervous system bleeding or active bleeding from other sites occurs. Survival of the transfused platelets is usually short.

Primary immune thrombocytopenia is also known as idiopathic thrombocytopenic purpura (ITP). In children, it is usually acute in onset, short lived, and typically follows a viral illness. In contrast, ITP in adults is gradual in onset, chronic in nature, and has no identifiable cause. Because the circulating platelets in ITP are young and functional, bleeding is less for a given platelet count than when there is failure of platelet production. The pathophysiology of ITP is believed to involve both impaired platelet production and T cell–mediated platelet destruction.¹⁰ Management options are summarized in Table 4-2.¹¹ Treatment of drug-induced immune thrombocytopenia may simply entail withdrawal of the offending drug, but corticosteroids, gamma globulin, and anti-D immunoglobulin may hasten recovery of the count.¹²,¹³

**Table 4-1**

<table>
<thead>
<tr>
<th>Etiology of acquired platelet disorders</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Quantitative Disorders</td>
</tr>
<tr>
<td>1. Failure of production: related to impairment in bone marrow function</td>
</tr>
<tr>
<td>a. Leukemia</td>
</tr>
<tr>
<td>b. Myeloproliferative disorders</td>
</tr>
<tr>
<td>c. B₁₂ or folate deficiencies</td>
</tr>
<tr>
<td>d. Chemotherapy or radiation therapy</td>
</tr>
<tr>
<td>e. Acute alcohol intoxication</td>
</tr>
<tr>
<td>f. Viral infections</td>
</tr>
<tr>
<td>2. Decreased survival</td>
</tr>
<tr>
<td>a. Immune-mediated</td>
</tr>
<tr>
<td>1) Idiopathic thrombocytopenia (ITP)</td>
</tr>
<tr>
<td>2) Heparin-induced thrombocytopenia</td>
</tr>
<tr>
<td>3) Autoimmune disorders or B-cell malignancies</td>
</tr>
<tr>
<td>4) Secondary thrombocytopenia</td>
</tr>
<tr>
<td>b. Disseminated intravascular coagulation (DIC)</td>
</tr>
<tr>
<td>c. Related to platelet thrombi</td>
</tr>
<tr>
<td>1) Thrombocytopenic purpura (TTP)</td>
</tr>
<tr>
<td>2) Hemolytic uremic syndrome (HUS)</td>
</tr>
<tr>
<td>3. Sequestration</td>
</tr>
<tr>
<td>a. Portal hypertension</td>
</tr>
<tr>
<td>b. Sarcoid</td>
</tr>
<tr>
<td>c. Lymphoma</td>
</tr>
<tr>
<td>d. Gaucher’s Disease</td>
</tr>
<tr>
<td>B. Qualitative Disorders</td>
</tr>
<tr>
<td>1. Massive transfusion</td>
</tr>
<tr>
<td>2. Therapeutic platelet inhibitors</td>
</tr>
<tr>
<td>3. Disease states</td>
</tr>
<tr>
<td>a. Myeloproliferative disorders</td>
</tr>
<tr>
<td>b. Monoclonal gammopathies</td>
</tr>
<tr>
<td>c. Liver disease</td>
</tr>
</tbody>
</table>

**Table 4-2**

<table>
<thead>
<tr>
<th>Management of idiopathic thrombocytopenic purpura (ITP) in adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>First line:</td>
</tr>
<tr>
<td>a. Corticosteroids: Longer courses of corticosteroids are preferred over shorter courses of corticosteroids</td>
</tr>
<tr>
<td>b. Intravenous immunoglobulin (IVIG) or anti-D immunoglobulin: the dose should initially be 1 g/kg as a one-time dose. This dosage may be repeated if necessary</td>
</tr>
<tr>
<td>Second line:</td>
</tr>
<tr>
<td>a. Splenectomy</td>
</tr>
<tr>
<td>b. Rituximab, an anti-CD 20 monoclonal antibody</td>
</tr>
<tr>
<td>c. Thrombopoietin (TPO) receptor agonists</td>
</tr>
<tr>
<td>d. Immunosuppressive agents</td>
</tr>
<tr>
<td>Third line: (failing first and second line therapy)</td>
</tr>
<tr>
<td>a. Thrombopoietin (TPO) receptor agonists</td>
</tr>
<tr>
<td>b. Combination of first and second line therapies</td>
</tr>
<tr>
<td>c. Combination chemotherapy</td>
</tr>
</tbody>
</table>
Heparin-induced thrombocytopenia (HIT) is a form of drug-induced immune thrombocytopenia. It is an immunologic event during which antibodies against platelet factor 4 (PF4) formed during exposure to heparin affect platelet activation and endothelial function with resultant thrombocytopenia and intravascular thrombosis. The platelet count typically begins to fall 5 to 7 days after heparin has been started, but if it is a reexposure, the decrease in count may occur within 1 to 2 days. HIT should be suspected if the platelet count falls to less than 100,000 or if it drops by 50% from baseline in a patient receiving heparin. While HIT is more common with full-dose unfractionated heparin (1% to 3%), it can also occur with prophylactic doses or with low molecular weight heparins. Interestingly, approximately 17% of patients receiving unfractionated heparin and 8% receiving low molecular weight heparin develop antibodies against PF4, yet a much smaller percentage develop thrombocytopenia, and even fewer develop clinical HIT. In addition to mild to moderate thrombocytopenia, this disorder is characterized by a high incidence of thrombosis that may be arterial or venous. Importantly, the absence of thrombocytopenia in these patients does not preclude the diagnosis of HIT. The 4Ts scoring system by Lo et al can be used to assess the pretest probability of HIT and incorporates the timing and magnitude of the platelet count fall, new thrombosis, and the likelihood of other reasons for thrombocytopenia. A low probability 4Ts score is quite accurate in excluding HIT, but patients with intermediate or high probability scores require further evaluation.

Laboratory testing should include an anti–platelet factor 4–heparin enzyme-linked immunosorbent assay (ELISA). Unfortunately, this test, like the 4Ts, has a high negative predictive value but a low positive predictive value. While a negative ELISA essentially rules out HIT, a positive ELISA does not confirm HIT. To increase the specificity of this assay, it can be restricted to IgG antibodies or obtained in conjunction with a functional assay such as the serotonin release assay and the heparin-induced platelet activation test. Both of these are available only at specialized laboratories and should only be used as second-line diagnostic assays.

The initial treatment of suspected HIT is to stop heparin and begin an alternative anticoagulant. Stopping heparin without addition of another anticoagulant is not adequate to prevent thrombosis in this setting. Alternative anticoagulants are primarily thrombin inhibitors. The most recent guideline by the American College of Chest Physicians recommends lepirudin, argatroban, or danaparoid for patients with normal renal function and argatroban for patients with renal insufficiency. Because of warfarin’s early induction of a hypercoagulable state, warfarin should be instituted only once full anticoagulation with an alternative agent has been accomplished and the platelet count has begun to recover.

These are also disorders in which thrombocytopenia is a result of platelet activation and formation of platelet thrombi. In thrombotic thrombocytopenic purpura (TTP), large vWF molecules interact with platelets, leading to activation. These large molecules result from inhibition of a metalloproteinase enzyme, ADAM-S13, which cleaves the large vWF molecules. TTP is classically characterized by thrombocytopenia, microangiopathic hemolytic anemia, fever, and renal and neurologic signs or symptoms. The finding of schistocytes on a peripheral blood smear aids in the diagnosis. Plasma exchange with replacement of FFP is the treatment for acute TTP. Additionally, rituximab, a monoclonal antibody against the CD20 protein on B lymphocytes, is indicated in relapsing and/or refractory TTP.

Hemolytic uremic syndrome (HUS) often occurs secondary to infection by Escherichia coli O157:H7 or other Shiga toxin-producing bacteria. The metalloproteinase is normal in these cases. HUS is usually associated with some degree of renal failure, with many patients requiring renal replacement therapy. Neurologic symptoms are less frequent. A number of patients develop features of both TTP and HUS. This may occur with autoimmune diseases, especially systemic lupus erythematosus and HIV infection, or in association with certain drugs (such as ticlopidine, mitomycin C, gemcitabine) or immunosuppressive agents (such as cyclosporine and tacrolimus). Discontinuation of the involved drug is the mainstay of therapy. Plasmapheresis is frequently used, but it is not clear what etiologic factor is being removed by the pheresis.

Sequestration is another important cause of thrombocytopenia and usually involves trapping of platelets in an enlarged spleen typically related to portal hypertension, sarcoid, lymphoma, or Gaucher’s disease. The total body platelet mass is essentially normal in patients with hypersplenism, but a much larger fraction of the platelets are in the enlarged spleen. Platelet survival is mildly decreased. Bleeding is less than anticipated from the count because sequestered platelets can be mobilized to some extent and enter the circulation. Platelet transfusion does not increase the platelet count as much as it would in a normal person because the transfused platelets are similarly sequestered in the spleen. Spleenectomy is not indicated to correct the thrombocytopenia of hypersplenism caused by portal hypertension.

Thrombocytopenia and platelet dysfunction are the most common abnormalities of hemostasis that result in bleeding in the surgical patient. The patient may have a reduced platelet count as a result of a variety of disease processes, as discussed earlier. In these circumstances, the marrow usually demonstrates a normal or increased number of megakaryocytes. By contrast, when thrombocytopenia occurs in patients with leukemia or uremia and in patients on cytotoxic therapy, there are generally a reduced number of megakaryocytes in the marrow. Thrombocytopenia also occurs in surgical patients as a result of massive blood loss with product replacement deficient in platelets. Thrombocytopenia may also be induced by heparin administration during cardiac and vascular cases, as in the case of HIT, or may be associated with thrombotic and hemorrhagic complications. When thrombocytopenia is present in a patient for whom an elective operation is being considered, management is contingent upon the extent and cause of platelet reduction and extent of platelet dysfunction.

Early platelet administration has now become part of massive transfusion protocols. Platelets are also administered preoperatively to rapidly increase the platelet count in surgical patients with underlying thrombocytopenia or platelet dysfunction. One unit of platelet concentrate contains approximately 5.5 \times 10^{10} platelets and would be expected to increase the circulating platelet count by about 10,000/μL in the average 70-kg person. Fever, infection, hepatosplenomegaly, and the presence of antiplatelet alloantibodies decrease the effectiveness of platelet transfusions. In patients who are refractory to standard platelet transfusion, the use of human leukocyte antigen (HLA)-compatible platelets coupled with special processors has proved effective.
Qualitative Platelet Defects. Impaired platelet function often accompanies thrombocytopenia but may also occur in the presence of a normal platelet count. The importance of this is obvious when one considers that 80% of overall clot strength is related to platelet function. The life span of platelets ranges from 7 to 10 days, placing them at increased risk for impairment by medical disorders and prescription and over-the-counter medications. Impairment of ADP-stimulated aggregation occurs with massive transfusion of blood products. Uremia may be associated with increased bleeding time and impaired aggregation. Defective aggregation and platelet dysfunction are also seen in patients with severe trauma, thrombocytopenia, polycythemia vera, and myelofibrosis.

Drugs that interfere with platelet function include aspirin, clopidogrel, prasugrel, dipyridamole, and GP IIb/IIIa inhibitors. Aspirin, clopidogrel, and prasugrel all irreversibly inhibit platelet function. Clopidogrel and prasugrel do so through selective irreversible inhibition of ADP-induced platelet aggregation. Aspirin works through irreversible acetylation of platelet prostaglandin synthase.

There are no prospective randomized trials in general surgical patients to guide the timing of surgery in patients on aspirin, clopidogrel, or prasugrel. The general recommendation is that approximately 5 to 7 days should pass from the time the drug is stopped until an elective procedure is performed. Timing of urgent and emergent surgeries is even more unclear. Preoperative platelet transfusions may be beneficial, but there are no good data to guide their administration. However, functional tests such as thromboelastography (TEG) with platelet mapping are becoming available that may better demonstrate defects in platelet function and may serve to guide the timing of operation or when platelet transfusions might be indicated.

Other disorders associated with abnormal platelet function include uremia, myeloproliferative disorders, monoclonal gammopathies, and liver disease. In the surgical patient, platelet dysfunction of uremia can often be corrected by dialysis or the administration of DDAVP. Platelet transfusion may not be helpful if the patient is uremic when the platelets are given and only serve to increase antibodies. Platelet dysfunction in myeloproliferative disorders is intrinsic to the platelets and usually improves if the platelet count can be reduced to normal with chemotherapy. If possible, surgery should be delayed until the count has been decreased. These patients are at risk for both bleeding and thrombosis. Platelet dysfunction in patients with monoclonal gammopathies is a result of interaction of the monoclonal protein with platelets. Treatment with chemotherapy or, occasionally, plasmapheresis to lower the amount of monoclonal protein improves hemostasis.

Acquired Hypofibrinogenemia

Disseminated Intravascular Coagulation (DIC). DIC is an acquired syndrome characterized by systemic activation of coagulation pathways that result in excessive thrombin generation and the diffuse formation of microthrombi. This disturbance ultimately leads to consumption and depletion of platelets and coagulation factors with the resultant classic picture of diffuse bleeding. Fibrin thrombi developing in the microcirculation may cause microvascular ischemia and subsequent end-organ failure if severe. There are many different conditions that predispose a patient to DIC, and the presence of an underlying condition is required for the diagnosis. For example, injuries resulting in embolization of materials such as brain matter, bone marrow, or amniotic fluid can act as potent thromboplastins that activate the DIC cascade. Additional etiologies include malignancy, organ injury (such as severe pancreatitis), liver failure, certain vascular abnormalities (such as large aneurysms), snake bites, illicit drugs, transfusion reactions, transplant rejection, and sepsis. In fact, DIC frequently accompanies sepsis and may be associated with multiple organ failure. The important interplay between sepsis and coagulation abnormalities was demonstrated by Dhainaut et al who showed that activated protein C was effective in septic patients with DIC, though this has subsequently been disproven. The diagnosis of DIC is made based on an inciting etiology with associated thrombocytopenia, prolongation of the prothrombin time, a low fibrinogen level, and elevated fibrin markers (FPDs, D-dimer, soluble fibrin monomers). A scoring system developed by the International Society for Thrombosis and Hemostasis has been shown to have high sensitivity and specificity for diagnosing DIC as well as a strong correlation between an increasing DIC score and mortality, especially in patients with infections.

The most important facets of treatment are relieving the patient’s causative primary medical or surgical problem and maintaining adequate perfusion. If there is active bleeding, hemostatic factors should be replaced with FFP, which is usually sufficient to correct the hypofibrinogenemia, although cryoprecipitate, fibrinogen concentrates, or platelet concentrates may also be needed. Given the formation of microthrombi in DIC, heparin therapy has also been proposed. Heparin may be indicated in cases where thrombosis predominates, such as arterial or venous thromboembolism and severe purpura fulminans.

Primary Fibrinolysis. Other than due to trauma, an acquired hypofibrinogenic state in the surgical patient can be a result of pathologic fibrinolysis. This may occur in patients following prostate resection when urokinase is released during surgical manipulation of the prostate or in patients undergoing extracorporeal bypass. The severity of fibrinolytic bleeding is dependent on the concentration of breakdown products in the circulation. Antifibrinolytic agents, such as ε-aminocaproic acid and tranexamic acid, interfere with fibrinolysis by inhibiting plasminogen activation.

Myeloproliferative Diseases

Polycythemia, or an excess of red blood cells, places surgical patients at risk. Spontaneous thrombosis is a complication of polycythemia vera, a myeloproliferative neoplasm, and can be explained in part by increased blood viscosity, increased platelet count, and an increased tendency toward stasis. Paradoxically, a significant tendency toward spontaneous hemorrhage also is noted in these patients. Thrombocytosis can be reduced by the administration of low-dose aspirin, phlebotomy, and hydroxyurea.
thrombotic risk. The most common coagulation abnormalities associated with liver dysfunction are thrombocytopenia and impaired humoral coagulation function manifested as prolongation of the prothrombin time and international normalized ratio (INR). The etiology of thrombocytopenia in patients with liver disease is typically related to hypersplenism, reduced production of thrombopoietin, and immune-mediated destruction of platelets. The total body platelet mass is often normal in patients with hypersplenism, but a much larger fraction of the platelets is sequestered in the enlarged spleen. Bleeding may be less than anticipated because sequestered platelets can be mobilized to some extent and enter the circulation. Thrombopoietin, the primary stimulus for thrombopoiesis, may be responsible for some cases of thrombocytopenia in cirrhotic patients, although its role is not well delineated. Finally, immune-mediated thrombocytopenia may also occur in cirrhotics, especially those with hepatitis C and primary biliary cirrhosis. In addition to thrombocytopenia, these patients also exhibit platelet dysfunction via defective interactions between platelets and the endothelium, and possibly due to uremia and changes in endothelial function in the setting of concomitant renal insufficiency. Hypocoagulopathy is further exacerbated with low platelet counts because platelets help facilitate thrombin generation by assembling coagulation factors on their surfaces. In conditions mimicking intravascular flow, low hematocrit and low platelet counts contributed to decreased adhesion of platelets to endothelial cells, although increased vWF, a common finding in cirrhotic patients, may offset this change in patients with cirrhosis. Hypercoagulability of liver disease has recently gained increased attention, with more evidence demonstrating the increased incidence of thromboembolism despite thrombocytopenia and a hypocoagulable state on conventional blood tests. This is attributed to decreased production of liver-synthesized proteins C and S, antithrombin, and plasminogen levels, as well as elevated levels of endothelial-derived vWF and factor VIII, a potent driver of thrombin generation. Given the concomitant hypo- and hypercoagulable features seen in patients with liver disease, conventional coagulation tests may be difficult to interpret, and whole blood functional tests such as thromboelastography (TEG) or ROTEM may be more informative of the status of clot formation and stability in cirrhotic patients. Small studies have indicated that TEG provides a better assessment of bleeding risk than standard tests of hemostasis in patients with liver disease; however, no large studies have directly tested this, and future larger trials are needed.

Before instituting any therapy to ameliorate thrombocytopenia, the actual need for correction should be strongly considered. In general, correction based solely on a low platelet count should be discouraged. Most often, treatment should be withheld for invasive procedures and surgery. When required, platelet transfusions are the mainstay of therapy; however, the effect typically lasts only several hours. Risks associated with transfusions in general and the development of antplatelet antibodies in a patient population likely to need recurrent correction should be considered. A less well-accepted option is splenectomy or splenic embolization to reduce hypersplenism. In addition to the risks associated with these techniques, reduced splenic blood flow can reduce portal vein flow with subsequent portal vein thrombosis. Results are mixed following insertion of a transjugular intrahepatic portosystemic shunt (TIPS). Therefore, treatment of thrombocytopenia should not be the primary indication for a TIPS procedure.

Decreased production or increased destruction of coagulation factors as well as vitamin K deficiency can all contribute to a prolonged PT and INR in patients with liver disease. As liver dysfunction worsens, so does the liver’s synthetic function, which results in decreased production of coagulation factors. Additionally, laboratory abnormalities may mimic those of DIC. Elevated D-dimers have been reported to increase the risk of variceal bleeding. The absorption of vitamin K is dependent on bile production. Therefore, liver patients with impaired bile production and cholestatic disease may be at risk for vitamin K deficiency.

Similar to thrombocytopenia, correction of coagulopathy should be reserved for treatment of active bleeding and prophylaxis for invasive procedures and surgery. Treatment of coagulopathy caused by liver disease is usually done with FFP, but because the coagulopathy is usually not a result of decreased levels of factor V, complete correction is not usually possible. If the fibrinogen is less than 200 mg/dL, administration of cryoprecipitate may be helpful. Cryoprecipitate is also a source of factor VIII for the rare patient with a low factor VIII level.

### Coagulopathy of Trauma

Traditional teaching regarding trauma-related coagulopathy attributed its development to acidosis, hypothermia, and dilution of coagulation factors. Recent data, however, have shown that over one-third of severely injured patients have laboratory-based evidence of coagulopathy at the time of admission, a phenotype called trauma-induced coagulopathy (TIC). TIC is independent of traditional (iatrogenic) causes of posttraumatic coagulopathy such as hemodilution, is precipitated by tissue injury and/or hemorrhagic shock, and is associated with significantly higher risk of mortality, especially in the first 24 hours after injury. Furthermore, TIC is a separate and distinct process from disseminated intravascular coagulopathy with its own specific components of hemostatic failure.

As shown in Fig. 4-5, several non–mutually exclusive mechanisms have been proposed as the etiology of TIC, including activated protein C-mediated clotting factor deactivation, endothelial injury and “auto-heparinization” due to shedding of endothelial heparin sulfate and chondroitin sulfate into the circulation, platelet dysfunction, and hyperfibrinolysis. Hemorrhagic shock was previously thought to be an essential component of TIC, but isolated traumatic brain injury and pulmonary contusions have been shown to induce laboratory-defined TIC in the absence of shock, possibly due to a high proportion of endothelium in these organs. Traumatic brain injury may also induce TIC via a consumptive mechanism by the release of large amounts of tissue factor into the circulation.
However, the relationship between laboratory-based coagulation abnormalities and true clinically evident coagulopathic bleeding is unclear. With the widespread application of damage control resuscitation, the frequency of clinical coagulopathy has decreased.

Interestingly, the converse of hyperfibrinolysis, known as fibrinolytic shutdown, has also been associated with increased mortality after trauma. In a multicenter study of 2540 trauma patients, those with intermediate fibrinolytic activity (“physiologic,” 0.8% to 2.9% lysis) had the lowest mortality (14%). Shutdown (<0.8% lysis) patients had increased mortality (22%), often due to late causes such as multiple organ failure, while patients with hyperfibrinolysis (≥3% lysis) had the greatest mortality (34%) and most often died due to hemorrhage.

**Acquired Coagulation Inhibitors**

Among the most common acquired coagulation inhibitors is the antiphospholipid syndrome (APLS), which includes the lupus anticoagulant and anticardiolipin antibodies. These antibodies may be associated with either venous or arterial thrombosis, or both. In fact, patients presenting with recurrent thrombosis should be evaluated for APLS. Antiphospholipid antibodies are very common in patients with systemic lupus but may also be seen in association with rheumatoid arthritis and Sjögren’s syndrome. There are also individuals who will have no autoimmune disorders but develop transient antibodies in response to infections or those who develop drug-induced APLS. The hallmark of APLS is a prolonged aPTT in vitro but an increased risk of thrombosis in vivo.

**Anticoagulation and Bleeding**

Spontaneous bleeding can be a complication of any anticoagulant therapy whether it is heparin, low molecular weight heparins, warfarin, factor Xa inhibitors, or new direct thrombin inhibitors. The risk of spontaneous bleeding related to heparin is reduced with a continuous infusion technique. Therapeutic anticoagulation is more reliably achieved with a low molecular weight heparin. However, laboratory testing is more challenging with these medications, as they are not detected with conventional coagulation testing. However, their more reliable therapeutic levels (compared to heparin) make them an attractive option for outpatient anticoagulation and more cost-effective for the inpatient setting. If monitoring is required (e.g., in the presence of renal insufficiency or severe obesity), the drug effect should be determined with an assay for anti-Xa activity.

Warfarin is used for long-term anticoagulation in various clinical conditions, including deep vein thrombosis, pulmonary embolism, valvular heart disease, atrial fibrillation, recurrent systemic emboli, recurrent myocardial infarction, prosthetic heart valves, and prosthetic implants. Due to the interaction of the P450 system, the anticoagulant effect of the warfarin is reduced (e.g., increased dose required) in patients receiving barbiturates as well as in patients with diets low in vitamin K. Increased warfarin requirements may also be needed in patients taking contraceptives or estrogen-containing compounds, corticosteroids, and adrenocorticotropic hormone (ACTH). Medications that can alter warfarin requirements are shown in Table 4-4.

Although warfarin use is often associated with a significant increase in morbidity and mortality in acutely injured and emergency surgery patients, with rapid reversal, these complications can be reduced. There are several reversal options that include vitamin K administration, plasma, cryoprecipitate, recombinant factor VIIa, and factor concentrates. The 2012 CHEST guidelines for the Management of Anticoagulant Therapy Antithrombotic Therapy and Prevention of Thrombosis recommends patients with major life-threatening bleeding

**Table 4-4**

<table>
<thead>
<tr>
<th>Medications that can alter warfarin dosing</th>
</tr>
</thead>
<tbody>
<tr>
<td>↓ warfarin effect</td>
</tr>
<tr>
<td>↑ warfarin requirements</td>
</tr>
<tr>
<td>↑ warfarin effect</td>
</tr>
</tbody>
</table>

due to warfarin receive reversal with vitamin K and a rapid reversal agent such as plasma or prothrombin complex concentrate (PCC). Vitamin K is given to sustain the effects of the plasma or PCC due to their short half-lives. In major bleeds, vitamin K 10 mg given as a slow IV infusion is utilized for more rapid onset compared to the oral form. Studies have shown that PCC is superior to plasma for speed of reversal and has decreased risk of fluid overload, but it is equivalent in adverse and thromboembolic events and costlier. Prothrombin complex concentrate is available in two forms: three-factor PCC (factors II, IX, and X) and four-factor PCC (factors II, VII, IX, and X). Four-factor PCCs have shown to have a more reliable correction of INR compared to three-factor PCCs.

Direct oral anticoagulants (DOACs) include direct thrombin inhibitors and factor Xa inhibitors and have no readily available method of detection of the degree of anticoagulation. More concerning is the difficulty in the reversal of these new anticoagulants. Recently, idarucizumab, a humanized monoclonal antibody fragment that binds dabigatran, has been approved for use for reversal of the thrombin inhibitor, dabigatran, and dabigatran-related coagulopathy. Clinical studies have demonstrated normalization of laboratory tests.

Factor Xa inhibitors such as rivaroxaban, apixaban, and edoxaban currently lack a specific antidote. Two novel anticoagulants, andexanet alfa and ciraparantag (PER977), are currently undergoing clinical trials. Andexanet alfa is a recombinant human FXa variant, and ciraparantag is a cationic small molecule. These are both being evaluated for reversal of the factor Xa inhibitors. Until these agents are approved, attempts to reverse Factor Xa inhibitors should include four factor PCCs. In less urgent states, these drugs can be held for 36 to 48 hours prior to surgery without increased risk of bleeding in those with normal renal function. Alternatively, activated clotting time (stand alone or with rapid TEG) or ecarin clotting time can be obtained in those on dabigatran, and anti-factor Xa assays can be obtained in those taking rivaroxaban.

Bleeding complications in patients on anticoagulants include hematuria, soft tissue bleeding, intracerebral bleeding, skin necrosis, and abdominal bleeding. Bleeding secondary to anticoagulation therapy is also a common cause of rectus sheath hematomas.

Surgical intervention may prove necessary in patients receiving anticoagulation therapy. Increasing experience suggests that surgical treatment can be undertaken without full reversal of the anticoagulant, depending on the procedure being performed. When the aPTT is less than 1.3 times control in a heparinized patient, or when the INR is less than 1.5 in a patient on warfarin, reversal of anticoagulation therapy may not be necessary. However, meticulous surgical technique is mandatory, and the patient must be observed closely throughout the postoperative period.

Certain surgical procedures should not be performed in concert with anticoagulation. In particular, cases where even minor bleeding can cause great morbidity, such as the central nervous system and the eye, surgery should be avoided. Emergency operations are occasionally necessary in patients who have been heparinized. The first step in these patients is to discontinue heparin. For more rapid reversal, protamine sulfate is effective. However, significant adverse reactions, especially in patients with severe fish allergies, may be encountered when administering protamine. Symptoms include hypotension, flushing, bradycardia, nausea, and vomiting. Prolongation of the aPTT after heparin neutralization with protamine may also be a result of the anticoagulant effect of protamine. In the elective surgical patient who is receiving warfarin-derivative therapy sufficient to effect anticoagulation, the drug can be discontinued several days before operation and the prothrombin concentration then checked. Rapid reversal of anticoagulation can be accomplished with plasma or prothrombin complex concentrates in the emergent situation. An example of a warfarin reversal guideline using four-factor prothrombin complex concentrate for patients with major or life-threatening bleeding or intracranial bleeding is shown in Fig. 4-6. Parenteral administration of vitamin K also is indicated in elective surgical treatment of patients with biliary obstruction or malabsorption who may be vitamin K deficient. However, if low levels of factors II, VII, IX, and X (vitamin K–dependent factors) exist as a result of hepatocellular dysfunction, vitamin K administration is ineffective.

Cardiopulmonary Bypass. Under normal conditions, homeostasis of the coagulation system is maintained by complex interactions between the endothelium, platelets, and coagulation factors. In patients undergoing cardiopulmonary bypass (CPB), contact with circuit tubing and membranes results in abnormal platelet and clotting factor activation, as well as activation of inflammatory cascades, that ultimately results in excessive fibrinolysis and a combination of both quantitative and qualitative platelet defects. Platelets undergo reversible alterations in morphology and their ability to aggregate, which causes sequestration in the filter, partially degranulated platelets, and platelet fragments. This multifactorial coagulopathy is compounded by the effects of shear stress in the system, induced hypothermia, hemoïdilution, and anticoagulation. While on pump, activated clotting time measurements are obtained along with blood gas measurements; however, conventional coagulation assays and platelet counts are not normally performed until rewarming and after a standard dose of protamine has been given. TEG may give a better estimate of the extent of coagulopathy and may also be used to anticipate transfusion requirements if bleeding is present. Empiric treatment with FFP and cryoprecipitate is often used for bleeding patients; however, there are no universally accepted transfusion thresholds. Platelet concentrates are given for bleeding patients in the immediate postoperative period; however, studies have shown that indiscriminate platelet therapy conferred no therapeutic advantage. It is in these
patients where rapid coagulation testing is required to assist with directed transfusion therapy. Many institutions now use antifibrinolytics, such as ε-aminocaproic acid and tranexamic acid, at the time of anesthesia induction after several studies have shown that such treatment reduced postoperative bleeding and reoperation. Aprotinin, a protease inhibitor that acts as an antifibrinolytic agent, has been shown to reduce transfusion requirements associated with cardiac surgery. Desmopressin acetate stimulates release of factor VIII from endothelial cells and may also be effective in reducing blood loss during cardiac surgery.

**Local Hemostasis.** Significant surgical bleeding is usually caused by ineffective local hemostasis. The goal is therefore to prevent further blood loss from a disrupted vessel that has been incised or transected. Hemostasis may be accomplished by interrupting the flow of blood to the involved area or by direct closure of the blood vessel wall defect.

**Mechanical Procedures.** The oldest mechanical method of bleeding cessation is application of direct digital pressure, either at the site of bleeding or proximally to permit more definitive action. An extremity tourniquet that occludes a major vessel proximal to the bleeding site or the Pringle maneuver for liver bleeding are good examples. Direct digital pressure is very effective and has the advantage of being less traumatic than hemostatic or even “atraumatic” clamps.

When a small vessel is transected, a simple ligature is usually sufficient. However, for larger pulsating arteries, a transfixion suture to prevent slipping is indicated. All sutures represent foreign material, and selection should be based on their intrinsic characteristics and the state of the wound. Direct pressure applied by “packing” a wound with gauze or laparotomy pads affords the best method of controlling diffuse bleeding from large areas, such as in the trauma situation. Packing bone wax on the raw surface to effect pressure can control bleeding from cut bone.

**Thermal Agents.** Heat achieves hemostasis by denaturation of protein that results in coagulation of large areas of tissue. Electrocautery generates heat by induction from an alternating current source, which is then transmitted via conduction from the instrument directly to the tissue. The amplitude setting should be high enough to produce prompt coagulation, but not so high as to set up an arc between the tissue and the cautery tip. This avoids thermal injury outside of the operative field and also prevents exit of current through electrocardiographic leads, other monitoring devices, or permanent pacemakers or defibrillators. A negative grounding plate should be placed beneath the patient to avoid severe skin burns, and caution should be used with certain

---

**Figure 4-6.** Example of a warfarin reversal guideline using four-factor prothrombin complex concentrate for patients with major or life-threatening bleeding or intracranial bleeding.
anesthetic agents (diethyl ether, divinyl ether, ethyl chloride, ethylene, and cyclopropane) because of the hazard of explosion.

A direct current also can result in hemostasis. Because the protein moieties and cellular elements of blood have a negative surface charge, they are attracted to a positive pole where a thrombus is formed. Direct currents in the 20- to 100-mA range have successfully controlled diffuse bleeding from raw surfaces, as has argon gas.

**Topical Hemostatic Agents**

Topical hemostatic agents can play an important role in helping to facilitate surgical hemostasis. These agents are classified based on their mechanism of action, and many act at specific stages in the coagulation cascade and take advantage of natural physiologic responses to bleeding. The ideal topical hemostatic agent has significant hemostatic action, minimal tissue reactivity, nonantigenicity, in vivo biodegradability, ease of sterilization, low cost, and can be tailored to specific needs.

Achneck et al have published a comprehensive overview of absorbable, biologic, and synthetic agents. Absorbable agents include gelatin foams (Gelfoam), oxidized cellulose (Surgicel), and microfibrillar collagens (Avitene). Both gelatin foam and oxidized cellulose provide a physical matrix for clotting initiation, while microfibrillar collagens facilitate platelet adherence and activation. Biologic agents include topical thrombin, fibrin sealants (FloSeal), and platelet sealants (Vitargel). Human or recombinant thrombin derivatives, which facilitate the formation of fibrin clots and subsequent activation of several clotting factors, take advantage of natural physiologic processes, thereby avoiding foreign body or inflammatory reactions. Caution must be taken in judging vessel caliber in the wound because thrombin entry into larger caliber vessels can result in systemic exposure to thrombin with a risk of disseminated intravascular clotting or death. They are particularly effective in controlling capillary bed bleeding when pressure or ligation is insufficient; however, the bovine derivatives should be used with caution due to the potential immunologic response and worsened coagulopathy. Fibrin sealants are prepared from cryoprecipitate (homologous or synthetic) and have the advantage of not promoting inflammation or tissue necrosis. A recent study by Koea et al demonstrated in a prospective multicenter randomized trial that a fibrin sealant patch was safe and highly effective in controlling parenchymal bleeding following hepatectomy regardless of the type of resection. Platelet sealants are a mixture of collagen and thrombin combined with plasma-derived fibrinogen and platelets from the patient, which requires the additional need for centrifugation and processing.

Topical agents are not a substitute for meticulous surgical technique and only function as adjuncts to help facilitate surgical hemostasis. The advantages and disadvantages of each agent must be considered, and use should be limited to the minimum amount necessary to minimize toxicity, adverse reactions, interference with wound healing, and procedural costs.

**Replacement Therapy**

**Typing and Crossmatching.** Serologic compatibility for A, B, O, and Rh groups is established routinely. Crossmatching between the donors’ red blood cells and the recipients’ sera (the major crossmatch) is performed. Rh-negative recipients should be transfused only with Rh-negative red blood cells. However, this group represents only 15% of the population. Therefore, the administration of Rh-positive red blood cells is acceptable if Rh-negative red blood cells blood is not available. However, Rh-positive red blood cells should not be transfused to Rh-negative females who are of childbearing age.

In emergency situations, universal donor type O-negative red blood cells and type AB plasma may be transfused to all recipients. Platelets do not require crossmatching. Due to a shortage of type AB plasma, low anti-B titer type A plasma has become widely adopted for emergency (uncrossmatched) transfusion. In the United States, 85% of individuals are type A or type O, making type AB plasma compatible with the vast majority of potential recipients. Uncrossmatched plasma is routinely transfused as part of platelet transfusions, with major transfusion reactions reported rarely, and type AB plasma currently carries a higher risk of TRALI compared to other plasma types. Many centers have transitioned to low titer type A plasma for emergency transfusions, with no increase in adverse events. O negative and type-specific red blood cells are equally safe for emergency transfusion. In patients known to have clinically significant cold agglutinins, blood should be administered through a blood warmer. If these antibodies are present in high titer, hypothermia is contraindicated.

In patients who have been multiply transfused and who have developed alloantibodies or who have autoimmune hemolytic anemia with pan-red blood cell antibodies, typing and crossmatching is often difficult, and sufficient time should be allotted preoperatively to accumulate blood that might be required during the operation. Crossmatching should always be performed before the administration of dextran because it interferes with the typing procedure.

**Banked Whole Blood.** Interest in whole blood as an ideal therapy for acute traumatic hemorrhagic shock has increased in the last several years with multiple reports of successful use in military and civilian trauma patients. However, there is still limited access in most civilian centers.

**Red Blood Cells and Frozen Red Blood Cells.** Red blood cells are the traditional product of choice for most clinical
situations requiring resuscitation, although deficits in oxygen delivery are rarely related to inadequate red cells. Concentrated suspensions of red blood cells can be prepared by removing most of the supernatant plasma after centrifugation. The preparation reduces but does not eliminate reactions caused by plasma components. With sequential changes in storage solutions, the shelf life of red blood cells is now 42 days. However, recent evidence has demonstrated that the age of red cells may play a significant role in the inflammatory response and incidence of multiple organ failure. The changes in the red blood cells that occur during storage include reduction of intracellular ADP and 2,3-diphosphoglycerate (2,3-DPG), which alters the oxygen dissociation curve of hemoglobin, resulting in a decrease in oxygen transport. Stored RBCs progressively become acidic with elevated levels of lactate, potassium, and ammonia. Additionally, the in vitro hemostatic potential of plasma and platelet products also decrease with storage.

The morphologic and biochemical changes that occur over time in red cells may contribute to worsened outcomes. This limits the ability to bank large amounts of blood, particularly rarer blood types, for use in times of high demand and blood supply shortage, such as on the battlefield and after mass casualty events. Storage solutions, however, do not fully suppress the metabolic and physical changes associated with aging RBCs. Newer evidence suggests that cryopreservation of red blood cells may provide a safe alternative means of storage. Cryopreservation uses the beneficial effects of ultra-low temperatures to suppress molecular motion and arrest metabolic and biochemical reactions. Frozen (cryopreserved) red blood cells have a shelf life of ten years at −80°C with improved cellular viability and maintenance of ATP and 2,3 DPG concentrations. A trial of stable trauma patients randomized to old (>14 storage days) red blood cells, young (≤14 storage days) red blood cells, and cryopreserved red blood cells found that cryopreserved red blood cells were as safe and effective as standard red blood cells. Cryopreserved red blood cells required a thawing and preparation period of about 90 minutes, limiting immediate availability for emergency use. A recent study suggests that the post-thaw characteristics of cryopreserved units may not, however, be comparable to fresh red cells. Additional research needs to be done to optimize the process, but frozen cells likely represent a viable option for storage in the future.

**Leukocyte-Reduced and Leukocyte-Reduced/Washed Red Blood Cells.** These products are prepared by filtration that removes about 99.9% of the white blood cells and most of the platelets (leukocyte-reduced red blood cells) and, if necessary, by additional saline washing (leukocyte-reduced/washed red blood cells). Leukocyte reduction prevents almost all febrile, nonhemolytic transfusion reactions (fever and/or rigors), alloimmunization to HLA class I antigens, and platelet transfusion refractoriness and cytomegalovirus transmission. In most Western nations, it is the standard red blood cell transfusion product. Supporters of universal leukocyte reduction argue that allogenic transfusion of white cells predisposes to postoperative bacterial infection and multiorgan failure. Reviews of randomized trials and meta-analyses have not provided convincing evidence either way, although a large Canadian retrospective study suggests a decrease in mortality and infections.

**Platelet Concentrates.** The indications for platelet transfusion include thrombocytopenia caused by massive blood loss and replacement with platelet-poor products, thrombocytopenia caused by inadequate production, and qualitative platelet disorders. Platelets are stored at room temperature under constant agitation to prevent clumping and have a shelf life of 5 days from time of donation due to risk of bacterial overgrowth. One unit of platelet concentrate has a volume of approximately 50 mL. Platelet preparations are capable of transmitting infectious diseases and can account for allergic reactions similar to those caused by red blood cell transfusion. A therapeutic level of platelets is in the range of 50,000 to 100,000/μL, but is very dependent on the clinical situation. Recent evidence suggests that earlier use of platelets may improve outcomes in bleeding patients.

In rare cases, in patients who become alloimmunized through previous transfusion or patients who are refractory from sensitization through prior pregnancies, HLA-matched platelets can be used.

**Plasma.** Plasma is the usual source of the vitamin K–dependent factors, the only source of factor V, and carries similar infectious risks as other component therapies. Several plasma products are available. Fresh frozen plasma (FFP) is frozen within hours of donation and can be stored for up to two years at −18°C, but requires 20 to 30 minutes to thaw prior to use, limiting immediate availability. Thawed FFP can be relabeled as thawed plasma, which is immediately transfusable and can be stored for up to 5 days at 2°C to 4°C. Liquid plasma is never frozen and can be stored for up to 26 days at 2°C to 4°C. In vitro studies demonstrate that liquid plasma has a better hemostatic profile than thawed plasma. Freeze-dried (lyophilized) plasma (FDP) has been recently “rediscovered” as an ideal resuscitation product for patients in remote and austere environments. FDP is distributed as a powder that is shelf-stable for up to 2 years at room temperature and relatively stable at temperature extremes. It was used extensively as a primary resuscitation fluid during World War II, but production was stopped due to risk of disease transmission. FDP is currently manufactured by updated processes in France, Germany, and South Africa. Several noncomparative studies in the literature have documented its ease of use, rapid reconstitution within minutes, clinical efficacy similar to other plasma products, and lack of apparent adverse events. The Israeli Defense Force has reported successful use of FDP at the point of injury, just as it was used in World War II. Beside limited use by U.S. Special Forces under the U.S. Federal Drug Administration’s (FDA) Investigational New Drug (IND) program, no FDP product is currently approved for general use in the United States. These products have the advantage of being pathogen reduced, have expanded storage capabilities, and can be quickly reconstituted.

**Tranexamic Acid.** Tranexamic acid (TXA; trade name: Cyclokapron) is an antifibrinolytic that inhibits both plasminogen activation and plasmin activity, thus preventing clot breakdown rather than promoting new clot formation. It occupies the lysine-binding sites on plasminogen, thus preventing its binding to lysine residues on fibrin. This reduces plasminogen activation to plasmin. Similarly, blockade of lysine-binding sites on circulating plasmin prevents binding to fibrin and thus prevents clot breakdown. TXA is 10 times more potent in vitro than aminocaproic acid. At therapeutically relevant concentrations, TXA does not affect platelet count or aggregation or coagulation parameters. It is excreted largely unchanged in urine and has a half-life of about 2 hours in circulation. It has been used to decrease bleeding and the need for blood transfusions in coronary artery
bypass grafting (CABG), orthotopic liver transplantation, hip and knee arthroplasty, and other surgical settings. TXA has been used to treat injured patients in both civilian and military settings. A recent practice guideline by the Eastern Association for the Surgery of Trauma (EAST) conditionally recommended the use of tranexamic acid as a hemostatic adjunct in severely injured patients when used early after injury. There is some controversy if its use should be empiric in patients with hemorrhage or based on documented hyperfibrinolysis. Results of prehospital studies with TXA are not yet available, but there are five ongoing trials. The true risk of venous thrombotic events is also not well established. Therefore, tranexamic acid should not be used with active intravascular clotting and should not be given with activated prothrombin complex concentrate or factor IX complex concentrates.

Indications for Replacement of Blood and Its Elements

Improvement in Oxygen-Carrying Capacity. Oxygen-carrying capacity is primarily a function of the red blood cells. Thus, transfusion of red blood cells should augment oxygen-carrying capacity. Additionally, hemoglobin is fundamental to arterial oxygen content and thus oxygen delivery. Despite this obvious association, there is little evidence that actually supports the premise that transfusion of red blood cells equates with enhanced cellular delivery and utilization. The reasons for this apparent discrepancy are related to changes that occur with storage of blood. The decrease in 2,3-DPG and P50 impair oxygen offloading, and deformation of the red cells impairs microcirculatory perfusion.

Treatment of Anemia: Transfusion Triggers. The concept of transfusion triggers refers primarily to the nonactively bleeding ICU patient. A 1988 National Institutes of Health Consensus Report challenged the dictum that a hemoglobin value of less than 10 g/dL or a hematocrit level less than 30% indicates a need for preoperative red blood cell transfusion. This was verified in a prospective randomized controlled trial in critically ill patients that compared a restrictive transfusion threshold to a more liberal strategy and demonstrated that maintaining hemoglobin levels between 7 and 9 g/dL had no adverse effect on mortality. In fact, patients with APACHE II scores of ≤20 or patients age <55 years actually had a lower mortality.

One unresolved issue related to transfusion triggers is the safety of maintaining a hemoglobin of 7 g/dL in a patient with ischemic heart disease. Data on this subject are mixed, and many studies have significant design flaws, including their retrospective nature. However, the majority of the published data favors a restrictive transfusion trigger for patients with non-ST-elevation acute coronary syndrome, with many reporting worse outcomes in those patients receiving transfusions. Recent guidelines from the American Association of Blood Banks (AABB) recommend a minimum threshold of 7 g/dL for hemodynamically stable patients and 8 g/dL for patients undergoing cardiac surgery, orthopedic surgery, and those with pre-existing cardiovascular disease. However, both the SCCM/EAST and AABB guidelines recommend taking into account patient-specific characteristics and the overall clinical context when considering RBC transfusions in non-acuteley hemorrhaging patients. Patients with symptomatic anemia should be transfused one RBC unit at a time, and isolated asymptomatic anemia in and of itself is rarely an indication for RBC transfusion.

Volume Replacement

The most common indication for blood transfusion in surgical patients is the replenishment of the blood volume; however, the quantification of actual intravascular volume deficit is often difficult to accurately and quickly determine. Measurements of hemoglobin or hematocrit levels are frequently used to assess blood loss, but can be occasionally misleading in the face of acute loss. Both the amount and the rate of bleeding are factors in the development of signs and symptoms of blood loss.

Loss of blood in the operating room can be roughly evaluated by estimating the amount of blood in the wound and on the drapes, weighing the sponges, and quantifying blood suctioned from the operative field. Significant blood loss will require a balanced resuscitation including red blood cells, FFP, and platelets (detailed later in this chapter) (Table 4-5).

New Concepts in Resuscitation

Traditional resuscitation algorithms were sequentially based on crystalloid followed by red blood cells and then plasma and platelet transfusions, and they have been in widespread use since the 1970s. No quality clinical data supported this concept. Recently the damage control resuscitation (DCR) strategy, with simultaneous measures to acquire mechanical hemorrhage control, has become the standard for treatment of substantial traumatic hemorrhage. DCR emphasizes rapid maneuvers that promote hemostasis (balanced resuscitation with early delivery of plasma and platelets) while limiting iatrogenic insults that exacerbate bleeding (i.e., minimization of crystalloid and artificial colloid, permissive hypotension), combined with multiple adjuncts for hemorrhage control.

Rationale. In urban civilian trauma systems, nearly half of all deaths happen before a patient reaches the hospital. Patients who survive to an emergency center have a high incidence of truncal hemorrhage, and deaths in this group of patients may be potentially preventable. Truncal hemorrhage patients in shock often present with the early coagulopathy of trauma in the emergency department and are at significant risk of dying. Many of these patients have suffered substantial bleeding, generally defined as requiring the administration of ≥3 units of red blood cells within any hour of admission, and may have received a massive transfusion (MT), traditionally defined as ≥10 units of red blood cells in 24 hours. The traditional definition is admittedly arbitrary and fails to identify many patients who truly receive large volume transfusions in a short period of time, further promoting survival bias. Newer definitions evaluating massive transfusion do so by taking into account both volume transfused as well as the rate at which transfusions are given. The critical administration threshold (CAT) has been prospectively validated and shown to be a superior predictor of mortality when compared to the conventional definition of MT. By this measure, CAT-positive status is defined by transfusion of 3 units of red blood cells within a 60-minute period, and this is additive for each additional time this measure is reached. CAT-positive status is associated with a two-fold increase in risk of mortality. CAT is more sensitive than traditional definitions of bleeding and allows for both earlier and more accurate identification of injured patients at greatest risk of death.

Although 25% of all severely injured trauma admissions receive a unit of blood early after admission, only a small
## Table 4-5: Replacement of clotting factors

<table>
<thead>
<tr>
<th>FACTOR</th>
<th>NORMAL LEVEL</th>
<th>LIFE SPAN IN VIVO (HALF-LIFE)</th>
<th>FATE DURING COAGULATION</th>
<th>LEVEL REQUIRED FOR SAFE HEMOSTASIS</th>
<th>IDEAL AGENT ACD BANK BLOOD (4°C [39.2°F])</th>
<th>IDEAL AGENT FOR REPLACING DEFICIT</th>
</tr>
</thead>
<tbody>
<tr>
<td>I (fibrinogen)</td>
<td>200–400 mg/100 mL</td>
<td>72 h</td>
<td>Consumed</td>
<td>60–100 mg/100 mL</td>
<td>Very stable</td>
<td>Bank blood; concentrated fibrinogen</td>
</tr>
<tr>
<td>II (prothrombin)</td>
<td>20 mg/100 mL (100% of normal level)</td>
<td>72 h</td>
<td>Consumed</td>
<td>15%–20%</td>
<td>Stable</td>
<td>Bank blood; concentrated preparation</td>
</tr>
<tr>
<td>V (proaccelerin, accelerator globulin, labile factor)</td>
<td>100% of normal level</td>
<td>36 h</td>
<td>Consumed</td>
<td>5%–20%</td>
<td>Labile (40% of normal level at 1 wk)</td>
<td>Fresh frozen plasma; blood</td>
</tr>
<tr>
<td>VII (proconvertin, serum prothrombin conversion accelerator, stable factor)</td>
<td>100% of normal level</td>
<td>5 h</td>
<td>Survives</td>
<td>5%–30%</td>
<td>Stable</td>
<td>Bank blood; concentrated preparation</td>
</tr>
<tr>
<td>VIII (antihemophilic factor, antihemophilic globulin)</td>
<td>100% of normal level (50%–150% of normal level)</td>
<td>6–12 h</td>
<td>Consumed</td>
<td>30%</td>
<td>Labile (20%–40% of normal level at 1 wk)</td>
<td>Fresh frozen plasma; concentrated antihemophilic factor; cryoprecipitate</td>
</tr>
<tr>
<td>IX (Christmas factor, plasma thromboplastin component)</td>
<td>100% of normal level</td>
<td>24 h</td>
<td>Survives</td>
<td>20%–30%</td>
<td>Stable</td>
<td>Fresh-frozen plasma; bank blood; concentrated preparation</td>
</tr>
<tr>
<td>X (Stuart-Prower factor)</td>
<td>100% of normal level</td>
<td>40 h</td>
<td>Survives</td>
<td>15%–20%</td>
<td>Stable</td>
<td>Bank blood; concentrated preparation</td>
</tr>
<tr>
<td>XI (plasma thromboplastin antecedent)</td>
<td>100% of normal level</td>
<td>Unknown</td>
<td>Survives</td>
<td>Deficit produces no bleeding tendency</td>
<td>Stable</td>
<td>Replacement not required</td>
</tr>
<tr>
<td>XII (Hageman factor)</td>
<td>100% of normal level</td>
<td>40 h</td>
<td>Survives</td>
<td>Probably &lt;1%</td>
<td>Stable</td>
<td>Bank blood</td>
</tr>
<tr>
<td>XIII (fibrinase, fibrin-stabilizing factor)</td>
<td>100% of normal level</td>
<td>4–7 d</td>
<td>Survives</td>
<td>Probably &lt;1%</td>
<td>Stable</td>
<td>Bank blood</td>
</tr>
<tr>
<td>Platelets</td>
<td>150,000–400,000/μL</td>
<td>8–11 d</td>
<td>Consumed</td>
<td>60,000–100,000/μL</td>
<td>Very labile (40% of normal level at 20 h; 0 at 48 h)</td>
<td>Fresh blood or plasma; fresh platelet concentrate (not frozen plasma)</td>
</tr>
</tbody>
</table>

ACD = acid-citrate-dextrose.

percentage of patients receive a massive transfusion. In the military setting, the percentage of massive transfusion patients almost doubles.111

**Damage Control Resuscitation.** Prior to DCR, resuscitation guidelines advocated volume replacement with crystalloid, followed by packed red blood cell and only later plasma or platelets.112 This conventional massive transfusion practice was based on a several small uncontrolled retrospective studies that used blood products containing increased amounts of plasma, which are no longer available.113 Because of the known early coagulopathy of trauma, the current approach to managing the exsanguinating patient involves early implementation of DCR.

Although most of the attention to hemorrhagic shock resuscitation has centered on higher ratios of plasma and platelets, DCR is actually composed of four basic components: permissive hypotension, minimizing crystalloid-based resuscitation, the immediate release and administration of predefined balanced blood products (red blood cells, plasma, and platelets) in ratios similar to those of whole blood, and the use of hemostatic adjuncts.

The shift to DCR began in earnest in 2007 when a retrospective study of 246 military casualties reported that patients with high plasma:RBC ratio (median 1:1.4) had substantially reduced mortality (19% vs. 65%) compared to patients with low plasma:RBC ratio (median 1:8).114 Subsequent observational studies among civilian and military trauma patients corroborated these findings.115-118 In particular, the prospective, observational, multicenter, major trauma transfusion (PROMMTT) study119 found that hemorrhagic death occurred rapidly (median of 2 to 3 hours after hospital arrival) and that plasma:RBC and platelet:RBC ratios significantly varied during massive transfusion. Increased plasma:RBC (adjusted hazard ratio [HR] 0.31, 95% confidence interval [CI] 0.16-0.58) and increased platelet:RBC (adjusted HR 0.55, 95% CI 0.31-0.98) were associated with reduced 6-hour mortality, when risk of hemorrhagic death was highest. After 6 hours, however, increasing plasma:RBC and platelet:RBC were no longer associated with reduced mortality due to increasing competing risk for non-hemorrhagic death (e.g., traumatic brain injury). The Pragmatic Randomized Optimal Platelet and Plasma Ratios (PROPR) trial120 randomized 680 bleeding trauma patients across 12 highest-level trauma centers to resuscitation with 1:1:1 vs. 1:1:2 plasma to platelets to RBCs. Although there was no significant difference in mortality at 24 hours (13% vs. 17%) or 30 days (22% vs. 26%), the 1:1:1 group had significantly decreased mortality due to hemorrhage at 24 hours (9% vs. 15%) and more patients achieving hemostasis (86% vs. 78%). Despite fears that resuscitation with increased plasma volumes would lead to more inflammatory complications, there were no between-group differences in 23 prespecified secondary outcomes, including acute respiratory distress syndrome, sepsis, multiple organ failure, and venous thromboembolism. A recent systematic review/meta-analysis and practice management guideline from EAST reported reduced mortality (31% vs. 38%) in 5292 patients receiving high (≥1:1) versus low (<1:2) plasma to RBC, and reduced mortality (28% vs. 43%) in 1607 patients receiving high versus low platelet to RBC.121 The authors therefore recommend high and balanced ratio (≥1:1) of plasma and platelet to RBC for resuscitation of severely injured trauma patients.

The mechanism for these benefits are unclear. While correction of hypovolemia as well as augmentation of the patient’s hemostatic potential with clotting factors and platelets are important, other plasma proteins likely play key roles as well. Recently, plasma resuscitation has been shown to reverse endothelial injury in animal models of hemorrhagic shock, particularly by repair of the endothelial glycocalyx layer (EGL).122,123 The EGL is the primary determinant of vascular permeability.122 Hemorrhage results in shedding of EGL components and vascular permeability. Crystalloid and artificial colloid-based resuscitation increases the hydrostatic pressure without repairing the EGL, which likely contributes to the myriad of edema-related complications seen in the pre-DCR era. Plasma, on the other hand, repairs the EGL, limiting extravascular leakage and edema. However, the exact protein moieties that mediate these benefits have yet to be identified and remain an area of investigation. Nevertheless, several studies have reported decreased inflammatory and edema-related complications with increased plasma and decreased crystalloid utilization. In trauma patients, there are strong correlations between increasing circulating levels of glycocalyx components such as syndecan-1 and trauma severity, coagulopathy, and mortality,124-126 although it remains unclear if these relationships are causative or merely associative. Finally, the use of DCR principles to guide transfusion of substantial nontraumatic hemorrhage is intuitive, although there is little evidence in the literature to support this practice.

It is essential that the trauma center has an established mechanism to deliver these products quickly and in the correct amounts to these critically injured patients.99 An example of an adult massive transfusion clinical guideline specifying the early use of component therapy is shown in Table 4-6. Specific recommendations for the administration of component therapy during a massive transfusion are shown in Table 4-7.

Because only a small percentage of trauma patients require a massive transfusion and because blood products in general are in short supply, there is a need for early prediction models.127 A comparison of results from existing models in both civilian and military studies is shown in Table 4-8.128-132 While compelling, many of these models require laboratory data, complicated injury severity scores, or calculated values that are not readily available or feasible to obtain in the urgent setting of bleeding. The Assessment of Blood Consumption (ABC) score is a simplified score to predict massive transfusion after trauma using immediately available data (heart rate, blood pressure, FAST exam, mechanism of injury).128 The ABC score has been validated across multiple trauma centers; however, it may be limited in some centers by the variable use of and operator-dependent FAST examination. In using the ABC score as it was intended, less than 5% of patients who will require massive transfusion will be missed; and 85% of all major trauma patients will be correctly identified.

**Prehospital Transfusion**

In bleeding patients, earlier initiation of appropriate therapy improves outcomes. For example, decreased overall blood product use and increased 30-day survival was observed after moving four units of universal donor, ready-to-transfuse plasma from the blood bank to the emergency department and using the plasma as a primary resuscitation fluid.133 A prehospital retrospective study that analyzed 1677 severely injured trauma patients who were transported by helicopter found that in-flight plasma transfusion was associated with less deranged physiology on admission and reduced early mortality in the most critically ill patients.134 Prehospital RBC transfusion has also been...
A. Initial Transfusion of Red Blood Cells (RBCs):
1. Notify blood bank immediately of urgent need for RBCs.
   - O negative uncrossmatched (available immediately).
   - As soon as possible, switch to O negative for females and O positive for males.
   - Type-specific uncrossmatched (available in approximately 5–10 min).
   - Completely crossmatched (available in approximately 40 min).
2. A blood sample must be sent to blood bank for a type and cross.
3. The Emergency Release of Blood form must be completed. If the blood type is not known and blood is needed immediately, O-negative RBCs should be issued.
4. RBCs will be transfused in the standard fashion. All patients must be identified (name and number) prior to transfusion.
5. Patients who are unstable or receive 1–2 RBCs and do not rapidly respond should be considered candidates for the massive transfusion (MT) guideline.

B. Adult Massive Transfusion Guideline:
1. The Massive Transfusion Guideline (MTG) should be initiated as soon as it is anticipated that a patient will require massive transfusion. The blood bank should strive to deliver plasma, platelets, and RBCs in a 1:1:1 ratio. To be effective and minimize further dilutional coagulopathy, the 1:1:1 ratio must be initiated early, ideally with the first 2 units of transfused RBCs. Crystalloid infusion should be minimized.
2. Once the MTG is activated, the blood bank will have 6 RBCs, 6 FFP, and a 6-pack of platelets packed in a cooler available for rapid transport. If 6 units of thawed FFP are not immediately available, the blood bank will issue units that are ready and notify appropriate personnel when the remainder is thawed. Every attempt should be made to obtain a 1:1:1 ratio of plasma:platelets:RBCs.
3. Once initiated, the MT will continue until stopped by the attending physician. MT should be terminated once the patient is no longer actively bleeding.
4. No blood components will be issued without a pickup slip with the recipient’s medical record number and name.
5. Basic laboratory tests should be drawn immediately on ED arrival and optimally performed on point-of-care devices, facilitating timely delivery of relevant information to the attending clinicians. These tests should be repeated as clinically indicated (e.g., after each cooler of products has been transfused). Suggested laboratory values are:
   - CBC
   - INR, fibrinogen
   - pH and/or base deficit
   - TEG, where available

CBC = complete blood count; ED = emergency department; FFP = fresh frozen plasma; INR = international normalized ratio; TEG = thromboelastography.

---

Table 4-7

<table>
<thead>
<tr>
<th>Component therapy administration during massive transfusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fresh frozen plasma (FFP)</td>
</tr>
<tr>
<td>Platelets</td>
</tr>
<tr>
<td>Cryoprecipitate</td>
</tr>
</tbody>
</table>

Table 4-8

<table>
<thead>
<tr>
<th>Comparison of massive transfusion prediction studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHOR</td>
</tr>
<tr>
<td>McLaughlin et al</td>
</tr>
<tr>
<td>Yücel et al</td>
</tr>
<tr>
<td>Moore et al</td>
</tr>
<tr>
<td>Schreiber et al</td>
</tr>
<tr>
<td>Cotton et al</td>
</tr>
</tbody>
</table>

AUC = area under the curve; BD = base deficit; FAST = focused assessment with sonography for trauma; Hct = hematocrit; Hgb = hemoglobin; HR = heart rate; INR = international normalized ratio; ISS = injury severity score; ROC = receiver operating characteristic; SBP = systolic blood pressure.
associated with similar findings. In the military setting, implementation of prehospital transfusion protocols in conjunction with other measures, including more rapid transport times, was also associated with reduced mortality.

Whole Blood Resuscitation
Military experience with whole blood for the resuscitation of traumatic hemorrhage is extensive, going back to the American Civil War. In the modern era, more than 10,000 whole blood units were transfused during Operations Enduring Freedom and Iraqi Freedom. One key advantage of whole blood versus component therapy is that platelets are often unavailable in the remote and austere settings. Two retrospective studies of military casualties treated at a combat support hospital and forward surgical teams found that whole blood was associated with improved survival compared to component (plasma and RBC) therapy. Whole blood has higher hematocrit, clotting factor activity, and platelet count compared to 1:1:1 component therapy due to relatively less diluent volume in whole blood. During the Vietnam War, low anti-A and anti-B titer whole blood was transfused universally with a low incidence of hemolytic reactions (1 per 9600 units). An in vitro study found that the hemostatic potential of whole blood was preserved for up to 14 days with cold storage. Pilot trials have reported successful use of crossmatched modified whole blood (leukoreduced and platelet-poor) and uncrossmatched low-titer whole blood (leukoreduced, containing platelets) in the initial resuscitation of civilian trauma patients. In the future, whole blood may return as the therapy of choice for the initial resuscitation of substantial hemorrhage.

Fibrinogen Replacement
Fibrinogen is the first coagulation factor to fall to critically low levels during major hemorrhage, and low systemic concentrations of fibrinogen are associated with increased severity of injury and coagulopathy and are independently predictive of mortality. Additionally, fibrinogen levels drop in the prehospital phase of injury, suggesting early administration by fibrinogen concentrate (not FDA-approved) or cryoprecipitate is needed. Fibrinogen concentrate is stored as a lyophilized powder at room temperature and can be reconstituted quickly allowing for rapid administration without delays for thawing or crossmatching. In contrast to plasma, viral inactivation steps are routinely included in the manufacturing process for fibrinogen concentrate, thus minimizing the risk of viral transmission. A pilot trial of massively transfused trauma patients randomized to a massive transfusion protocol or a massive transfusion protocol with early cryoprecipitate found that early cryoprecipitate delivery was feasible and that these patients had higher fibrinogen levels at all time points during resuscitation, although there was no mortality difference. A randomized control trial in Austria of prehospital fibrinogen concentrate versus placebo has been completed with publication of results pending.

Complications of Transfusion (Table 4-9)
Transfusion-related complications are primarily related to blood-induced proinflammatory responses. Transfusion-related events are estimated to occur in approximately 10% of all transfusions, but less than 0.5% are serious in nature. Transfusion-related deaths, although exceedingly rare, do occur and are related primarily to transfusion-related acute lung injury (TRALI), ABO hemolytic transfusion reactions, and bacterial contamination of platelets.

Nonhemolytic Reactions. Febrile, nonhemolytic reactions are defined as an increase in temperature (>1°C) associated with a transfusion and are fairly common (approximately 1% of all transfusions). Preformed cytokines in donated blood and recipient antibodies reacting with donated antibodies are postulated etiologies. The incidence of febrile reactions can be greatly reduced by the use of leukocyte-reduced blood products. Pretreatment with acetaminophen reduces the severity of the reaction. Bacterial contamination of infused blood is rare. Gram-negative organisms, which are capable of growth at 4°C, are the most common cause. Most cases, however, are associated with the administration of platelets that are stored at 20°C or, even more commonly, with apheresis platelets stored at room temperature. Cases from FFP thawed in contaminated water baths have also been reported. Bacterial contamination can result in sepsis and death in up 25% of patients. Clinical manifestations include systemic signs such as fever and chills, tachycardia and hypotension, and gastrointestinal symptoms (abdominal cramps, vomiting, and diarrhea). If the diagnosis is suspected, the transfusion should be discontinued and the blood cultured. Emergency treatment includes oxygen, adrenergic blocking agents, and antibiotics.

Allergic Reactions. Allergic reactions are relatively frequent, occurring in about 1% of all transfusions. Reactions are usually mild and consist of rash, urticaria, and flushing. In rare instances, anaphylactic shock develops. Allergic reactions are caused by the transfusion of antibodies from hypersensitive donors or the transfusion of antigens to which the recipient is hypersensitive. Allergic reactions can occur after the administration of any blood product but are commonly associated with FFP and platelets. Treatment and prophylaxis consist of the administration of antihistamines. In more serious cases, epinephrine or steroids may be indicated.

Respiratory Complications. Respiratory compromise may be associated with transfusion-associated circulatory overload (TACO), which is an avoidable complication. It can occur with rapid infusion of blood, plasma expanders, and crystalloids, particularly in older patients with underlying heart disease. Central venous pressure monitoring should be considered whenever large amounts of fluid are administered. Overload is manifested by a rise in venous pressure, dyspnea, and cough. Rales can generally be heard at the lung bases. Treatment consists of diuresis, slowing the rate of blood administration, and minimizing fluids while blood products are being transfused.

The syndrome of TRALI is defined as noncardiogenic pulmonary edema related to transfusion. It can occur with the administration of any plasma-containing blood product. Symptoms are similar to circulatory overload with dyspnea and associated hypoxemia. However, TRALI is characterized as noncardiogenic and is often accompanied by fever, rigors, and bilateral pulmonary infiltrates on chest X-ray. It most commonly occurs within 1 to 2 hours after the onset of transfusion but virtually always before 6 hours. Toy et al reported a decrease in the incidence of TRALI with the reduction transfusion of plasma from female donors, due to a combination of reduced transfusion of strong cognate HLA class II antibodies and HNA antibodies in patients with risk factors for acute lung injury. TRALI now occurs less than 1 in 10,000 units transfused and is usually self-limited with supportive therapy. Treatment of TRALI entails discontinuation of any transfusion, notification of the transfusion service, and pulmonary support, which may vary from supplemental oxygen to mechanical ventilation.
Hemolytic Reactions. Hemolytic reactions can be classified as either acute or delayed. Acute hemolytic reactions occur with the administration of ABO-incompatible blood and can be fatal in up to 6% of cases. Contributing factors include errors in the laboratory of a technical or clerical nature or the administration of the wrong blood type. Immediate hemolytic reactions are characterized by intravascular destruction of red blood cells and consequent hemoglobinemia and hemoglobinuria. DIC can be initiated by antibody-antigen complexes activating factor XII and complement, leading to activation of the coagulation cascade. Finally, acute renal insufficiency results from the toxicity associated with free hemoglobin in the plasma, resulting in tubular necrosis and precipitation of hemoglobin within the tubules.

Delayed hemolytic transfusion reactions occur 2 to 10 days after transfusion and are characterized by extravascular hemolysis, mild anemia, and indirect (unconjugated) hyperbilirubinemia. They occur when an individual has a low antibody titer at the time of transfusion, but the titer increases after transfusion as a result of an anamnestic response. Reactions to non-ABO antigens involve immunoglobulin G-mediated clearance by the reticuloendothelial system.

If the patient is awake, the most common symptoms of acute transfusion reactions are pain at the site of transfusion, facial flushing, and back and chest pain. Associated symptoms include fever, respiratory distress, hypotension, and tachycardia. In anesthetized patients, diffuse bleeding and hypotension are the hallmarks. A high index of suspicion is needed to make the diagnosis. The laboratory criteria for a transfusion reaction are hemoglobinuria and serologic criteria that show incompatibility of the donor and recipient blood. A positive Coombs’ test indicates
transfused cells coated with patient antibody and is diagnostic. Delayed hemolytic transfusions may also be manifested by fever and recurrent anemia. Jaundice and decreased haptoglobin usually occur, and low-grade hemoglobinemia and hemoglobinuria may be seen. The Coombs’ test is usually positive, and the blood bank must identify the antigen to prevent subsequent reactions.

If an immediate hemolytic transfusion reaction is suspected, the transfusion should be stopped immediately, and a sample of the recipient’s blood drawn and sent along with the suspected unit to the blood bank for comparison with the pretransfusion samples. Urine output should be monitored and adequate hydration maintained to prevent precipitation of hemoglobin within the tubules. Delayed hemolytic transfusion reactions do not usually require specific intervention.

**Transmission of Disease.** Malaria, Chagas’ disease, brucellosis, and, very rarely, syphilis are among the diseases that have been transmitted by transfusion. Malaria can be transmitted by all blood components. The species most commonly implicated is *Plasmodium malariae*. The incubation period ranges from 8 to 100 days; the initial manifestations are shaking chills and spiking fever. Cytomegalovirus (CMV) infection resembling infectious mononucleosis also has occurred.

Transmission of hepatitis C and HIV-1 has been dramatically minimized by the introduction of better antibody and nucleic acid screening for these pathogens. The residual risk among allogeneic donations is now estimated to be less than 1 per 1,000,000 donations. The residual risk of hepatitis B is approximately 1 per 300,000 donations. Hepatitis A is very rarely transmitted because there is no asymptomatic carrier state. Improved donor selection and testing are responsible for the decreased rates of transmission. Recent concerns about the rare transmission of these and other pathogens, such as West Nile virus, are being addressed by current trials of “pathogen inactivation systems” that reduce infectious levels of all viruses and bacteria known to be transmittable by transfusion. Prion disorders (e.g., Creutzfeldt-Jakob disease) also are transmissible by transfusion, but there is currently no information on inactivation of prions in blood products for transfusion.

Recently, there is heightened concern of transmission of Zika virus by blood product transfusion. Studies in endemic areas have shown rates of Zika infection detected in donor blood as high as 2.8%. Although no such cases have been reported in the United States, transmission of Zika virus via platelet products have been reported in Brazil. Zika virus may result in serious birth defects including microcephaly when infection occurs in pregnant women. Because the majority of cases in adults produce nonspecific or no symptoms, Zika screening cannot be accomplished by questionnaires. The Centers for Disease Control and Prevention has issued guidelines for screening of Zika virus in donated blood. Although no tests have been FDA-approved, laboratory testing is currently being performed under the FDA’s IND program.

**Tests of Hemostasis and Blood Coagulation**

The initial approach to assessing hemostatic function is a careful review of the patient’s clinical history (including previous abnormal bleeding or bruising), drug use, and basic laboratory testing.

**Conventional Coagulation Tests.** Common screening laboratory testing includes platelet count, PT or INR, and aPTT. Platelet dysfunction can occur at either extreme of platelet count. The normal platelet count ranges from 150,000 to 400,000/μL. Whereas a platelet count greater than 1,000,000/μL may be associated with bleeding or thrombotic complications, increased bleeding complications may be observed with major surgical procedures when the platelets are below 50,000/μL and with minor surgical procedures when counts are below 30,000/μL, and spontaneous hemorrhage can occur when the counts fall below 20,000/μL. Despite a lack of evidence supporting their use, platelet transfusions are still recommended in ophthalmologic and neurosurgical procedures when the platelet count is less than 100,000/μL.

The PT and aPTT are variations of plasma recalcification times initiated by the addition of a thromboplastin agent. The PT reagent contains thromboplastin and calcium that, when added to plasma, leads to the formation of a fibrin clot. The PT test measures the function of factors I, II, V, VII, and X. Factor VII is part of the extrinsic pathway, and the remaining factors are part of the common pathway. Factor VII has the shortest half-life of the coagulation factors, and its synthesis is vitamin K dependent. The PT test is best suited to detect abnormal coagulation caused by vitamin K deficiencies and warfarin therapy.

Due to variations in thromboplastin activity, it can be difficult to accurately assess the degree of anticoagulation on the basis of PT alone. To account for these variations, the INR is now the method of choice for reporting PT values. The International Sensitivity Index (ISI) is unique to each batch of thromboplastin and is furnished by the manufacturer to the hematology laboratory. Human brain thromboplastin has an ISI of 1, and the optimal reagent has an ISI between 1.3 and 1.5.

The INR is a calculated number derived from the following equation:

\[
\text{INR} = \frac{\text{measured PT}}{\text{normal PT}}^{\text{ISI}}
\]

The aPTT reagent contains a phospholipid substitute, activator, and calcium, which in the presence of plasma leads to fibrin clot formation. The aPTT measures function of factors I, II, and V of the common pathway and factors VIII, IX, X, and XII of the intrinsic pathway. Heparin therapy is often monitored by following aPTT values with a therapeutic target range of 1.5 to 2.5 times the control value (approximately 50 to 80 seconds). Low molecular weight heparins are selective Xa inhibitors that may mildly elevate the aPTT, but therapeutic monitoring is not routinely recommended.

Additional medications may significantly impair hemostatic function, such as antiplatelet agents (clopidogrel and GP IIb/IIIa inhibitors), anticoagulant agents (hirudin, chondroitin sulfate, dextran sulfate), and thrombolytic agents (streptokinase, tPA). If abnormalities in any of the coagulation studies cannot be explained by known medications, congenital abnormalities of coagulation or comorbid disease should be considered.

Unfortunately, while conventional coagulation tests (PT, aPTT) capture the classic intrinsic and extrinsic coagulation cascade, they do not reflect the complexity of in vivo coagulation. Although they are useful to follow warfarin and heparin therapies, they poorly reflect the status of actively bleeding patients. This is not surprising given that these tests use only plasma and not whole blood to provide their assessment of the patient’s clotting status. To better assess the complex and rapidly changing hemostatic function of an actively bleeding patient, many centers have moved to whole blood viscoelastic testing.
Viscoelastic Assays. Viscoelastic assays, such as TEG or rotational thromboelastometry (ROTEM), monitor hemostasis as a dynamic process rather than revealing information from isolated conventional coagulation screen.

**PART I** Viscoelastic Assays. Viscoelastic assays, such as TEG or rotational thromboelastometry (ROTEM), monitor hemostasis as a dynamic process rather than revealing information from isolated conventional coagulation screen.

Both tests measure the viscoelastic properties of blood as clotting is induced under a low-shear environment. The patterns of change in shear elasticity enable determination of the kinetics of clot formation and growth as well as the strength and stability of the formed clot. The strength and stability provide information about the ability of the clot to perform the work of hemostasis, while the kinetics determines the adequacy of quantitative factors available for clot formation.

Continuous improvements in this technique have made this test a valuable tool for medical personnel interested in coagulation. A sample of celite-activated whole blood is placed into a prewarmed cuvette, and the clotting process is activated with reagents, such as kaolin for standard TEG, and kaolin plus tissue factor for rapid TEG. Both TEG and ROTEM employ a vertical pin which is lowered into the activated blood sample. In TEG, the cuvette oscillates in an arc around the stationary pin. As the blood clots, fibrin strands and platelet aggregates form between the pin and inner walls of the cuvette. The resulting torque on the pin is measured and converted to an electrical signal. In ROTEM, the cuvette is stationary while the pin oscillates within the sample. The extent to which the pin can oscillate is reduced as the blood clots, and this is measured by the angle of deflection of a beam of light directed at the pin.

In TEGs, the strength of a clot is graphically represented over time as a characteristic cigar-shaped figure (Fig. 4-7).

Several parameters are generated from the TEG tracing. The r-value (reaction time) represents the time between the start of the assay and initial clot formation. This reflects clotting factor activity and initial fibrin formation and is increased with factor deficiency or severe hemodilution. The k-time (clot kinetics) is the time needed to reach specified clot strength and represents the interactions of clotting factors and platelets. As such, the k-time is prolonged with hypofibrinogenemia and significant factor deficiency. Prolonged r-value and k-time are commonly addressed with plasma transfusions. The alpha or angle ($\alpha$) is the slope of the tracing and reflects clot acceleration. The angle reflects the interactions of clotting factors and platelets. The slope is decreased with hypofibrinogenemia and platelet dysfunction. Decreased angles are treated with cryoprecipitate transfusion or fibrinogen administration. The maximal amplitude (mA) is the greatest height of the tracing and represents clot strength. Its height is reduced with dysfunction or deficiencies in platelets or fibrinogen. Decreased mA is addressed with platelet transfusion and, in cases where the angle is also decreased, with cryoprecipitate (or fibrinogen) as well. The G-value is a parametric measure derived from the mA value and reflects overall clot strength or firmness. An increased G-value is associated with hypercoagulability, whereas a decrease is seen with hypo-coagulable states. Finally, the LY30 is the amount of lysis occurring in the clot, and the value is the percentage of amplitude reduction at 30 minutes after mA is achieved. The LY30 represents clot stability and when increased fibrinolysis is present.

TEG and ROTEM are the only tests measuring all dynamic steps of clot formation until eventual clot lysis or retraction. TEG has also been shown to identify on admission those patients likely to develop thromboembolic complications after injury and postoperatively.

Recent trauma data have shown TEG to be useful in predicting early transfusion of red blood cells, plasma, platelets, and cryoprecipitate. TEG can also predict the need for lifesaving interventions shortly after arrival, 24-hour and 30-day mortality, and can be used to guide administration of TXA to injured patients with hyperfibrinolysis. Lastly, some centers have demonstrated that the graphic display options allow for more rapid return of results and may be less expensive than standard coagulation panels. Given the strong association of viscoelastic tests with clinical outcomes, some centers now use TEG rather than conventional coagulation tests to evaluate injured patients in the emergency department.

**EVALUATION OF EXCESSIVE INTRAOPERATIVE OR POSTOPERATIVE BLEEDING**

Excessive bleeding during or after a surgical procedure may be the result of ineffective hemostasis, blood transfusion, undetected hemostatic defect, consumptive coagulopathy, and/or fibrinolysis. Excessive bleeding from the operative field associated with bleeding from other sites usually suggests inadequate mechanical hemostasis.

Massive blood transfusion is a well-known cause of thrombocytopenia. Bleeding following massive transfusion can occur because of hypothermia, dilutional coagulopathy, platelet dysfunction, fibrinolysis, or hypofibrinogenemia. Another cause of hemostatic failure related to the administration of blood is a hemolytic transfusion reaction. The first sign of a transfusion reaction may be diffuse bleeding. The pathogenesis of this bleeding is thought to be related to the release of ADP from hemolyzed red blood cells, resulting in diffuse platelet aggregation, after which the platelet clumps are removed out of the circulation.

Transfusion purpura occurs when the donor platelets are of the uncommon HPA-1 group. This is an uncommon cause of thrombocytopenia and associated bleeding after transfusion. The platelets sensitize the recipient, who makes antibody to the foreign platelet antigen. The foreign platelet antigen does not completely disappear from the recipient circulation but attaches to the recipient’s own platelets. The antibody then destroys the recipient’s own platelets. The resultant thrombocytopenia and bleeding may continue for several weeks. This uncommon cause of thrombocytopenia should be considered if bleeding follows transfusion by 5 or 6 days. Platelet transfusions are of little help in the management of this syndrome because the new donor platelets usually are subject to the binding of antigen and damage from the antibody. Corticosteroids may be of some help in reducing the bleeding tendency. Posttransfusion purpura is self-limited, and the passage of several weeks inevitably leads to subsidence of the problem.

DIC is characterized by systemic activation of the coagulation system, which results in the deposition of fibrin clots and microvascular ischemia and may contribute to the development of this syndrome.
of multiorgan failure. Consumption and subsequent exhaustion of coagulation proteins and platelets due to the ongoing activation of the coagulation system may induce severe bleeding complications.

Lastly, severe hemorrhagic disorders due to thrombocytopenia have occurred as a result of gram-negative sepsis. Defibrination and hemostatic failure also may occur with meningococcemia, Clostridium perfringens sepsis, and staphylococcal sepsis. Hemolysis appears to be one mechanism in sepsis leading to defibrination.

REFERENCES

Entries highlighted in bright blue are key references.


37. Tripodi A, Mannucci PM. The coagulopathy of chronic liver disease. N Engl J Med. 2011;365:147. This is a review of the pathophysiology abnormalities in coagulation associated with chronic liver disease. It challenges the traditional teaching that these patients are coagulopathic and presents evidence and rational for why they are actually hypercoaguable.


52. Goldstein JN, Refaai MA, Milling TJ Jr, et al. Four-factor prothrombin complex concentrate versus plasma for rapid vitamin K antagonist reversal in patients needing urgent surgical or invasive interventions: a phase 3b, open-label, non-inferiority, randomised trial. Lancet. 2015;385(9982):2077-2087. This is a multicenter phase 3b noninferiority study of patients requiring rapid reversal of vitamin K antagonist-induced anticoagulation for urgent surgical procedures. Four-factor PCC was not just noninferior but in fact superior to plasma for rapid reversal and hemostasis.


66. Douketis JD, Spyropoulos AC, Kaatz S, et al. Perioperative bridging anticoagulation in patients with atrial fibrillation. N Engl J Med. 2015;373(9):823-833. This is an important trial that demonstrated forgoing bridging anticoagulation was noninferior to perioperative bridging with low-molecular-weight heparin for the prevention of arterial thromboembolism and decreased the risk of major bleeding. It has changed management of these patients.


96. Pusateri AE, Given MB, Schreiber MA, et al. Dried plasma: state of the science and recent developments. Transfusion. 2016 Apr;56 Suppl 2:S128-S30. Due to challenges in using fresh frozen plasma in emergent and austere conditions, a variety of dried plasma products have been and are being developed. The history, state of the science, use of dried products in other countries, and recent developments are reviewed.


120. Holcomb JB, Tilley BC, Baraniuk S, et al. Transfusion of plasma, platelets, and red blood cells in a 1:1:1 vs a 1:1:2 ratio and mortality in patients with severe trauma: the PROPPR randomized clinical trial. JAMA. 2015;313:471-482. This is a prospective randomized study at 12 level 1 trauma centers of massively bleeding trauma patients, comparing two balanced transfusion groups. Fewer patients bled to death in the 1:1:1 group.


This page intentionally left blank
“Shock is the manifestation of the rude unhinging of the machinery of life.”

—Samuel V. Gross, 1872

EVOLUTION IN UNDERSTANDING SHOCK

Overview

Shock, at its most rudimentary definition and regardless of its etiology, is the failure to meet the metabolic needs of the cell and the consequences that ensue. The initial cellular injury that occurs is reversible; however, the injury will become irreversible if tissue perfusion is prolonged or severe enough such that, at the cellular level, compensation is no longer possible. Our understanding of shock and the disease processes that result in shock made its most significant advances throughout the 20th century as our appreciation for the physiology and pathophysiology of shock matured. Most notably, our comprehension of the sympathetic and neuroendocrine stress responses on the cardiovascular system has flourished. The clinical manifestations of these physiologic responses are most often what lead practitioners to the diagnosis of shock as well as guide the management of patients in shock. However, hemodynamic parameters such as blood pressure and heart rate are relatively insensitive measures of shock, and additional considerations must be used to help aid in early diagnosis and treatment of patients in shock. The general approach to the management of patients in shock has been empiric: assuring a secure airway with adequate ventilation, control of hemorrhage in the bleeding patient, and restoration of vascular volume and tissue perfusion.

Historical Background

Integral to our understanding of shock is the appreciation that our bodies attempt to maintain a state of homeostasis. Claude Bernard suggested in the mid-19th century that the organism attempts to maintain constancy in the internal environment against external forces that attempt to disrupt the milieu interieur. Walter B. Cannon carried Bernard’s observations further and introduced the term homeostasis, emphasizing that an organism’s ability to survive was related to maintenance of homeostasis. The failure of physiologic systems to buffer the organism against external forces results in organ and cellular dysfunction, what is clinically recognized as shock. He first described the “fight or flight response,” generated by elevated levels of catecholamines in the bloodstream. Cannon’s observations on the battlefields of World War I led him to propose that the initiation of shock was due to a disturbance of the nervous system that resulted in vasodilation and hypotension. He proposed that secondary shock, with its attendant capillary permeability leak, was caused by a “toxic factor” released from the tissues.

In a series of critical experiments, Alfred Blalock documented that the shock state in hemorrhage was associated with reduced cardiac output due to volume loss, not a “toxic factor.” In 1934, Blalock proposed four categories of shock: hypovolemic, vasogenic, cardiogenic, and neurogenic. Hypovolemic shock, the most common type, results from loss of circulating blood volume. This may result from loss of whole blood (hemorrhagic shock), plasma, interstitial fluid (bowel obstruction), or a combination. Vasogenic shock results from decreased resistance within capacitance vessels, usually seen in sepsis. Neurogenic shock is a form of vasogenic shock in which spinal cord injury or spinal anesthesia causes vasodilation due to acute loss of sympathetic vascular tone. Cardiogenic shock results from failure of the heart as a pump, as in arrhythmias or acute myocardial infarction (MI).

This categorization of shock based on etiology persists today (Table 5-1). In recent clinical practice, further classification has described six types of shock: hypovolemic, septic (vasodilatory), neurogenic, cardiogenic, obstructive, and traumatic shock. Obstructive shock is a form of cardiogenic shock that results from mechanical impediment to circulation leading to depressed cardiac output rather than primary cardiac failure.
Key Points

1. Shock is defined as a failure to meet the metabolic demands of cells and tissues and the consequences that ensue.

2. A central component of shock is decreased tissue perfusion. This may be a direct consequence of the etiology of shock, such as in hypovolemic/hemorrhagic, cardiogenic, or neurogenic etiologies, or may be secondary to elaborated or released molecules or cellular products that result in endothelial/cellular activation, such as in septic shock or traumatic shock.

3. Physiologic responses to shock are based upon a series of afferent (sensing) signals and efferent responses that include neuroendocrine, metabolic, and immune/inflammatory signaling.

4. The mainstay of treatment of hemorrhagic/hypovolemic shock includes volume resuscitation with blood products. In the case of hemorrhagic shock, timely control of bleeding is essential and influences outcome.

5. Prevention of hypothermia, acidemia, and coagulopathy are essential in the management of patients in hemorrhagic shock.

6. The mainstay of treatment of septic shock is fluid resuscitation, initiation of appropriate antibiotic therapy, and control of the source of infection. This includes drainage of infected fluid collections, removal of infected foreign bodies, and debridement of devitalized tissues.

7. A combination of physiologic parameters and markers of organ perfusion/tissue oxygenation are used to determine if patients are in shock and to follow the efficacy of resuscitation.

8. Prevention of hypothermia, acidemia, and coagulopathy are essential in the management of patients in hemorrhagic shock.

9. Core principles in the management of patients in hemorrhagic shock include: (a) control of active hemorrhage must occur promptly (delay in control of bleeding increases mortality and recent battlefield data would suggest that in the young and otherwise healthy population commonly injured in combat, that control of bleeding is the paramount priority); (b) volume resuscitation with blood products (red blood cells, plasma, and platelets) with limited volume of crystalloid must occur while operative control of bleeding is achieved; (c) unrecognized or inadequately corrected hypoperfusion increases morbidity and mortality (i.e., inadequate resuscitation results in avoidable early deaths from shock); and (d) excessive fluid resuscitation may exacerbate bleeding (i.e., uncontrolled resuscitation is harmful). Thus, both inadequate or unrestrained volume resuscitation are harmful.

Likewise, observations in the management of septic shock have led to consensus statements and evolving guideline based management. Core principles in the management of patients in septic shock include: (a) septic shock is an emergency, and treatment/resuscitation should begin as early as possible; (b) specific anatomic diagnosis of infection requiring emergent source control be identified or excluded as rapidly as possible and that any required source control intervention be implemented as soon as medically and logistically practical; (c) initiation of broad spectrum antibiotics within 1 hour of diagnosis; (d) in the resuscitation from sepsis-induced hypoperfusion, at any time of the day or night, the patient must be moved from the bed to the operating theater in a surgical gown, and the patient must be placed on a high-flow anesthesia machine within 15 minutes of the beginning of surgery.

Table 5-1

<table>
<thead>
<tr>
<th>Classification of shock</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypovolemic</td>
</tr>
<tr>
<td>Cardiogenic</td>
</tr>
<tr>
<td>Septic (vasogenic)</td>
</tr>
<tr>
<td>Neurogenic</td>
</tr>
<tr>
<td>Traumatic</td>
</tr>
<tr>
<td>Obstructive</td>
</tr>
</tbody>
</table>

This includes etiologies such as pulmonary embolism or tension pneumothorax. In traumatic shock, soft tissue and bony injury lead to the activation of inflammatory cells and the release of circulating factors, such as cytokines and intracellular molecules that modulate the immune response. Recent investigations have revealed that the inflammatory mediators released in response to tissue injury (damage-associated molecular patterns [DAMPs]) are recognized by many of the same cellular receptors (pattern recognition receptors [PRRs]) and activate similar signaling pathways as do bacterial products elaborated in sepsis (pathogen-associated molecular patterns), such as lipopolysaccharide. These effects of tissue injury are combined with the effects of hemorrhage, creating a more complex and amplified deviation from homeostasis.

In the mid-to later 20th century, further development of experimental models contributed significantly to the understanding of the pathophysiology of shock. In 1947, Wiggers developed a sustainable, irreversible model of hemorrhagic shock based on uptake of shed blood into a reservoir to maintain a set level of hypotension. G. Tom Shires added further understanding of hemorrhagic shock with a series of clinical studies demonstrating that a large extracellular fluid deficit, greater than could be attributed to vascular refilling alone, occurred in severe hemorrhagic shock. The phenomenon of fluid redistribution after major trauma involving blood loss was termed third spacing and described the translocation of intravascular volume into the peritoneum, bowel, burned tissues, or crush injury sites. These seminal studies form the scientific basis for the current treatment of hemorrhagic shock with red blood cells and lactated Ringer’s solution or isotonic saline.

As resuscitation strategies evolved and patients survived the initial consequences of hemorrhage, new challenges of sustained shock became apparent. During the Vietnam War, aggressive fluid resuscitation with red blood cells and crystalloid solution or plasma resulted in survival of patients who previously would have succumbed to hemorrhagic shock. Renal failure became a less frequent clinical problem; however, a new disease process, acute fulminant pulmonary failure, appeared as an early cause of death after seemingly successful surgery to control hemorrhage. Initially called DaNang lung or shock lung, the clinical problem became recognized as acute respiratory distress syndrome (ARDS). This led to new methods of prolonged mechanical ventilation. Our current concept of ARDS is a component in the spectrum of multiple organ system failure.

Studies and clinical observations over the past two decades have extended the early observations of Canon, that “restoration of blood pressure prior to control of active bleeding may result in loss of blood that is sorely needed,” and challenged the appropriate endpoints in resuscitation of uncontrolled hemorrhage. Core principles in the management of patients in hemorrhagic shock include: (a) control of active hemorrhage must occur promptly (delay in control of bleeding increases mortality and recent battlefield data would suggest that in the young and otherwise healthy population commonly injured in combat, that control of bleeding is the paramount priority); (b) volume resuscitation with blood products (red blood cells, plasma, and platelets) with limited volume of crystalloid must occur while operative control of bleeding is achieved; (c) unrecognized or inadequately corrected hypoperfusion increases morbidity and mortality (i.e., inadequate resuscitation results in avoidable early deaths from shock); and (d) excessive fluid resuscitation may exacerbate bleeding (i.e., uncontrolled resuscitation is harmful). Thus, both inadequate or unrestrained volume resuscitation are harmful.
least 30 mL/kg of intravenous crystalloid fluid be given within the first 3 hours, and additional fluids be guided by frequent reassessment of hemodynamic status; (e) vaspressors (norepinephrine) should be added to achieve a mean arterial pressure of 65 mmHg if fluid resuscitation is inadequate.

**Current Definitions and Challenges**

A modern definition and approach to shock acknowledges that shock consists of inadequate tissue perfusion marked by decreased delivery of required metabolic substrates and inadequate removal of cellular waste products. This involves failure of oxidative metabolism that can involve defects of oxygen (O₂) delivery, transport, and/or utilization. Current challenges include moving beyond fluid resuscitation based upon endpoints of tissue oxygenation and using therapeutic strategies at the cellular and molecular level. This approach will help to identify compensated patients or patients early in the course of their disease, initiate appropriate treatment, and allow for continued evaluation for the efficacy of resuscitation and adjuncts.

Current investigations focus on determining the cellular events that often occur in parallel to result in organ dysfunction, shock irreversibility, and death. This chapter will review our current understanding of the pathophysiology and cellular responses of shock states. Current and experimental diagnostic and therapeutic modalities for the different categories of shock are reviewed, with a focus on hemorrhagic/hypovolemic shock and septic shock.

**PATHOPHYSIOLOGY OF SHOCK**

Regardless of etiology, the initial physiologic responses in shock are driven by tissue hypoperfusion and the developing cellular energy deficit. This imbalance between cellular supply and demand leads to neuroendocrine and inflammatory responses, the magnitude of which is usually proportional to the degree and duration of shock. The specific responses will differ based on the etiology of shock, as certain physiologic responses may be limited by the inciting pathology. For example, the cardiovascular response driven by the sympathetic nervous system is markedly blunted in neurogenic or septic shock. Additionally, decreased perfusion may occur as a consequence of cellular activation and dysfunction, such as in septic shock and to a lesser extent traumatic shock (Fig. 5-1). Many of the organ-specific responses are aimed at maintaining perfusion in the cerebral and coronary circulation. These are regulated at multiple levels including (a) stretch receptors and baroreceptors in the heart and vasculature (carotid sinus and aortic arch), (b) chemoreceptors, (c) cerebral ischemia responses, (d) release of endogenous vasoconstrictors, (e) shifting of fluid into the intravascular space, and (f) renal reabsorption and conservation of salt and water.

Furthermore, the pathophysiologic responses vary with time and in response to resuscitation. In hemorrhagic shock, the body can compensate for the initial loss of blood volume primarily through the neuroendocrine response to maintain hemodynamics. This represents the compensated phase of shock. With continued hypoperfusion, which may be unrecognized, cellular death and injury are ongoing, and the decompensation phase of shock ensues. Microcirculatory dysfunction, parenchymal tissue damage, and inflammatory cell activation can perpetuate hypoperfusion. Ischemia/reperfusion injury will often exacerbate the initial insult. These effects at the cellular level, if untreated, will lead to compromise of function at the organ system level, thus leading to the “vicious cycle” of shock (Fig. 5-2). Persistent hypoperfusion results in further hemodynamic derangements and cardiovascular collapse. This has been termed the irreversible phase of shock and can develop quite insidiously and may only be obvious in retrospect. At this point, there has occurred extensive enough parenchymal and microvascular injury such that volume resuscitation fails to reverse the process, leading to death of the patient. In experimental animal models of hemorrhagic shock (modified Wiggers model), this is

---

**Figure 5-1.** Pathways leading to decreased tissue perfusion and shock. Decreased tissue perfusion can result directly from hemorrhage/hypovolemia, cardiac failure, or neurologic injury. Decreased tissue perfusion and cellular injury can then result in immune and inflammatory responses. Alternatively, elaboration of microbial products during infection or release of endogenous cellular products from tissue injury can result in cellular activation to subsequently influence tissue perfusion and the development of shock. HMGB1 = high mobility group box 1; LPS = lipopolysaccharide; RAGE = receptor for advanced glycation end products.
Afferent Signals

Afferent impulses transmitted from the periphery are processed within the central nervous system (CNS) and activate the reflexive effector responses or efferent impulses. These effector responses are designed to expand plasma volume, maintain peripheral perfusion and tissue O<sub>2</sub> delivery, and restore homeostasis. The afferent impulses that initiate the body’s intrinsic adaptive responses and converge in the CNS originate from a variety of sources. The initial inciting event usually is loss of circulating blood volume. Other stimuli that can produce the neuroendocrine response include pain, hypoxemia, hypercarbia, acidosis, infection, change in temperature, emotional arousal, or hypoglycemia. The sensation of pain from injured tissue is transmitted via the spinothalamic tracts, resulting in activation of the hypothalamic-pituitary-adrenal axis, as well as activation of the autonomic nervous system (ANS) to induce direct sympathetic stimulation of the adrenal medulla to release catecholamines.

Baroreceptors also are an important afferent pathway in initiation of adaptive responses to shock. Volume receptors, sensitive to changes in both chamber pressure and wall stretch, are present within the atria of the heart. They become activated with low volume hemorrhage or mild reductions in right atrial pressure. Receptors in the aortic arch and carotid bodies respond to alterations in pressure or stretch of the arterial wall, responding to larger reductions in intravascular volume or pressure. These receptors normally inhibit induction of the ANS. When activated, these baroreceptors diminish their output, thus disinhibiting the effect of the ANS. The ANS then increases its output, principally via sympathetic activation at the vasomotor centers of the brain stem, producing centrally mediated constriction of peripheral vessels.

Chemoreceptors in the aorta and carotid bodies are sensitive to changes in O<sub>2</sub> tension, H<sup>+</sup> ion concentration, and carbon dioxide (CO<sub>2</sub>) levels. Stimulation of the chemoreceptors results in vasodilation of the coronary arteries, slowing of the heart rate, and vasoconstriction of the splanchnic and skeletal circulation. In addition, a variety of protein and nonprotein mediators are produced at the site of injury as part of the inflammatory response, and they act as afferent impulses to induce a host response. These mediators include histamine, cytokines,
eicosanoids, and endothelins, among others that are discussed in greater detail later in this chapter in the “Immune and Inflammatory Responses” section.

**Efferent Signals**

**Cardiovascular Response.** Changes in cardiovascular function are a result of the neuroendocrine response and ANS response to shock, and they constitute a prominent feature of both the body’s adaptive response mechanism and the clinical signs and symptoms of the patient in shock. Hemorrhage results in diminished venous return to the heart and decreased cardiac output. This is compensated by increased cardiac heart rate and contractility, as well as venous and arterial vasoconstriction. Stimulation of sympathetic fibers innervating the heart leads to activation of β-adrenergic receptors that increase heart rate and contractility in this attempt to increase cardiac output. Increased myocardial O₂ consumption occurs as a result of the increased workload; thus, myocardial O₂ supply must be maintained, or myocardial dysfunction will develop. The cardiovascular response in hemorrhage/hypovolemia differs from the responses elicited with the other etiologies of shock. These are compared in Table 5-2.

Direct sympathetic stimulation of the peripheral circulation via the activation of α-adrenergic receptors on arterioles induces vasoconstriction and causes a compensatory increase in systemic vascular resistance and blood pressure. The arterial vasoconstriction is not uniform; marked redistribution of blood flow results. Selective perfusion to tissues occurs due to regional variations in arteriolar resistance, with blood shunted away from less essential organ beds such as the intestine, kidney, and skin. In contrast, the brain and heart have autoregulatory mechanisms that attempt to preserve their blood flow despite a global decrease in cardiac output. Direct sympathetic stimulation also induces constriction of venous vessels, decreasing the capacitance of the circulatory system and accelerating blood return to the central circulation.

Increased sympathetic output induces catecholamine release from the adrenal medulla. Catecholamine levels peak within 24 to 48 hours of injury and then return to baseline. Persistent elevation of catecholamine levels beyond this time suggests ongoing noxious afferent stimuli. The majority of the circulating epinephrine is produced by the adrenal medulla, while norepinephrine is derived from synapses of the sympathetic nervous system. Catecholamine effects on peripheral tissues include stimulation of hepatic glycogenolysis and gluconeogenesis to increase circulating glucose availability to peripheral tissues, an increase in skeletal muscle glycogenolysis, suppression of insulin release, and increased glucagon release.

**Hormonal Response.** The stress response includes activation of the ANS as discussed previously in the “Afferent Signals” section, as well as activation of the hypothalamic-pituitary-adrenal axis. Shock stimulates the hypothalamus to release corticotropin-releasing hormone, which results in the release of adrenocorticotropic hormone (ACTH) by the pituitary. ACTH subsequently stimulates the adrenal cortex to release cortisol. Cortisol acts synergistically with epinephrine and glucagon to induce a catabolic state. Cortisol stimulates gluconeogenesis and insulin resistance, resulting in hyperglycemia as well as muscle cell protein breakdown and lipolysis to provide substrates for hepatic gluconeogenesis. Cortisol causes retention of sodium and water by the nephrons of the kidney. In the setting of severe hypovolemia, ACTH secretion occurs independently of cortisol negative feedback inhibition.

The renin-angiotensin system is activated in shock. Decreased renal artery perfusion, β-adrenergic stimulation, and increased renal tubular sodium concentration cause the release of renin from the juxtaglomerular cells. Renin catalyzes the conversion of angiotensinogen (produced by the liver) to angiotensin I, which is then converted to angiotensin II by angiotensin-converting enzyme (ACE) produced in the lung. While angiotensin I has no significant functional activity, angiotensin II is a potent vasconstrictor of both splanchnic and peripheral vascular beds, and also stimulates the secretion of aldosterone, ACTH, and antiuretic hormone (ADH). Aldosterone, a mineralocorticoid, acts on the nephron to promote reabsorption of sodium, and as a consequence, water. Potassium and hydrogen ions are lost in the urine in exchange for sodium.

The pituitary also releases vasopressin or ADH in response to hypovolemia, changes in circulating blood volume sensed by baroreceptors and left atrial stretch receptors, and increased plasma osmolality detected by hypothalamic osmoreceptors. Epinephrine, angiotensin II, pain, and hyperglycemia increase production of ADH. ADH levels remain elevated for about 1 week after the initial insult, depending on the severity and persistence of the hemodynamic abnormalities. ADH acts on the distal tubule and collecting duct of the nephron to increase water permeability, decrease water and sodium losses, and preserve intravascular volume. Also known as arginine vasopressin, ADH acts as a potent mesenteric vasoconstrictor, shunting circulating blood away from the splanchnic organs during hypovolemia. This may contribute to intestinal ischemia and predispose to intestinal mucosal barrier dysfunction in

<table>
<thead>
<tr>
<th>Table 5-2: Hemodynamic responses to different types of shock</th>
</tr>
</thead>
<tbody>
<tr>
<td>TYPE OF SHOCK</td>
</tr>
<tr>
<td>Hypovolemic</td>
</tr>
<tr>
<td>Septic</td>
</tr>
<tr>
<td>Cardiogenic</td>
</tr>
<tr>
<td>Neurogenic</td>
</tr>
</tbody>
</table>

The hemodynamic responses are indicated by arrows to show an increase (↑), severe increase (↑↑), decrease (↓), severe decrease (↓↓), varied response (↑↓), or little effect (→). CVP = central venous pressure; PCWP = pulmonary capillary wedge pressure; $SVO_2$ = mixed venous oxygen saturation; SVR = systemic vascular resistance.
Vasopressin also increases hepatic gluconeogenesis and increases hepatic glycolysis.

In septic states, endotoxin directly stimulates arginine vasopressin secretion independently of blood pressure, osmotic, or intravascular volume changes. Proinflammatory cytokines also contribute to arginine vasopressin release. Interestingly, patients on chronic therapy with ACE inhibitors are more at risk of developing hypotension and vasodilatory shock with open heart surgery. Low plasma levels of arginine vasopressin were confirmed in these patients.

**Circulatory Homeostasis**

**Preload.** At rest, the majority of the blood volume is within the venous system. Venous return to the heart generates ventricular end-diastolic wall tension, a major determinant of cardiac output. Gravitational shifts in blood volume distribution are quickly corrected by alterations in venous capacity. With decreased arteriolar inflow, there is active constriction of the venous smooth muscle and passive elastic recoil in the thin-walled systemic veins. This increases venous return to the heart, thus maintaining ventricular filling.

Most alterations in cardiac output in the normal heart are related to changes in preload. Increases in sympathetic tone have a minor effect on skeletal muscle beds but produce a dramatic reduction in splanchnic blood volume, which normally holds 20% of the blood volume.

The normal circulating blood volume is maintained within narrow limits by the kidney’s ability to manage salt and water balance with external losses via systemic and local hemodynamic changes and hormonal effects of renin, angiotensin, and ADH. These relatively slow responses maintain preload by altering circulating blood volume. Acute responses to intravascular volume include changes in venous tone, systemic vascular resistance, and intrathoracic pressure, with the slower hormonal changes less important in the early response to volume loss. Furthermore, the net effect of preload on cardiac output is influenced by cardiac determinants of ventricular function, which includes coordinated atrial activity and tachycardia.

**Ventricular Contraction.** The Frank-Starling curve describes the force of ventricular contraction as a function of its preload. This relationship is based on force of contraction being determined by initial muscle length. Intrinsic cardiac disease will shift the Frank-Starling curve and alter mechanical performance of the heart. In addition, cardiac dysfunction has been demonstrated experimentally in burns and in hemorrhagic, traumatic, and septic shock.

**Afterload.** Afterload is the force that resists myocardial work during contraction. Arterial pressure is the major component of afterload influencing the ejection fraction. This vascular resistance is determined by precapillary smooth muscle sphincters. Blood viscosity also will increase vascular resistance. As afterload increases in the normal heart, stroke volume can be maintained by increases in preload. In shock, with decreased circulating volume and therefore diminished preload, this compensatory mechanism to sustain cardiac output is impeded. The stress response with acute release of catecholamines and sympathetic nerve activity in the heart increases contractility and heart rate.

**Microcirculation.** The microvascular circulation plays an integral role in regulating cellular perfusion and is significantly influenced in response to shock. The microvascular bed is innervated by the sympathetic nervous system and has a profound effect on the larger arterioles. Following hemorrhage, larger arterioles vasoconstrict; however, in the setting of sepsis or neurogenic shock, these vessels vasodilate. Additionally, a host of other vasoactive proteins, including vasopressin, angiotensin II, and endothelin-1, also lead to vasoconstriction to limit organ perfusion to organs such as skin, skeletal muscle, kidneys, and the GI tract to preserve perfusion of the myocardium and CNS.

Flow in the capillary bed is heterogeneous in shock states, which likely is secondary to multiple local mechanisms, including endothelial cell swelling, dysfunction, and activation marked by the recruitment of leukocytes and platelets. Together, these mechanisms lead to diminished capillary perfusion that may persist after resuscitation. In hemorrhagic shock, correction of hemodynamic parameters and restoration of O2 delivery generally leads to restoration of tissue O2 consumption and tissue O2 levels. In contrast, regional tissue dysoxia often persists in sepsis, despite similar restoration of hemodynamics and O2 delivery. Whether this defect in O2 extraction in sepsis is the result of heterogeneous impairment of the microcirculation (intraparenchymal shunting) or impaired tissue parenchymal cell oxidative phosphorylation and O2 consumption by the mitochondria is not resolved.13 Interesting data suggest that in sepsis the response to limit O2 consumption by the tissue parenchymal cells is an adaptive response to the inflammatory signaling and decreased perfusion.

An additional pathophysiologic response of the microcirculation to shock is failure of the integrity of the endothelium of the microcirculation and development of capillary leak, intracellular swelling, and the development of an extracellular fluid deficit. Seminal work by Shires helped to define this phenomenon.8,17 There is decreased capillary hydrostatic pressure secondary to changes in blood flow and increased cellular uptake of fluid. The result is a loss of extracellular fluid volume. The cause of intracellular swelling is multifactorial, but dysfunction of energy-dependent mechanisms, such as active transport by the sodium-potassium pump contributes to loss of membrane integrity.

Capillary dysfunction also occurs secondary to activation of endothelial cells by circulating inflammatory mediators generated in septic or traumatic shock. This exacerbates endothelial cell swelling and capillary leak, as well as increases leukocyte adherence. This results in capillary occlusion, which may persist after resuscitation, and is termed no-reflow. Further ischemic injury ensues as well as release of inflammatory cytokines to compound tissue injury. Experimental models have shown that neutrophil depletion in animals subjected to hemorrhagic shock produces fewer capillaries with no-reflow and lower mortality.14

---

**METABOLIC EFFECTS**

Cellular metabolism is based primarily on the hydrolysis of adenosine triphosphate (ATP). The splitting of the phosphoanhydride bond of the terminal or γ-phosphate from ATP is the source of energy for most processes within the cell under normal conditions. The majority of ATP is generated in our bodies through aerobic metabolism in the process of oxidative phosphorylation in the mitochondria. This process is dependent on the availability of O2 as a final electron acceptor in the electron transport chain. As O2 tension within a cell decreases, there is a decrease in oxidative phosphorylation, and the generation
of ATP slows. When O₂ delivery is so severely impaired such that oxidative phosphorylation cannot be sustained, the state is termed *dysxia*. When oxidative phosphorylation is insufficient, the cells shift to anaerobic metabolism and glycolysis to generate ATP. This occurs via the breakdown of cellular glycogen stores to pyruvate. Although glycolysis is a rapid process, it is not efficient, allowing for the production of only 2 mol of ATP from 1 mol of glucose. This is compared to complete oxidation of 1 mol of glucose that produces 38 mol of ATP. Additionally, under hypoxic conditions in anaerobic metabolism, pyruvate is converted into lactate, leading to an intracellular metabolic acidosis.

There are numerous consequences secondary to these metabolic changes. The depletion of ATP potentially influences all ATP-dependent cellular processes. This includes maintenance of cellular membrane potential, synthesis of enzymes and proteins, cell signaling, and DNA repair mechanisms. Decreased intracellular pH also influences vital cellular functions such as normal enzyme activity, cell membrane ion exchange, and cellular metabolic signaling. These changes also will lead to changes in gene expression within the cell. Furthermore, acidosis leads to changes in calcium metabolism and calcium signaling. Compounded, these changes may lead to irreversible cell injury and death.

Epinephrine and norepinephrine have a profound impact on cellular metabolism. Hepatic glycogenolysis, gluconeogenesis, ketogenesis, skeletal muscle protein breakdown, and adipose tissue lipolysis are increased by catecholamines. Cortisol, glucagon, and ADH also contribute to the catabolism during shock. Epinephrine induces further release of glucagon, while inhibiting the pancreatic β-cell release of insulin. The result is a catabolic state with glucose mobilization, hyperglycemia, protein breakdown, negative nitrogen balance, lipolysis, and insulin resistance during shock and injury. The relative underuse of glucose by peripheral tissues preserves it for the glucose-dependent organs such as the heart and brain.

**Cellular Hypoperfusion**

Hypoperfused cells and tissues experience what has been termed *oxygen debt*, a concept first proposed by Crowell in 1961. The O₂ debt is the deficit in tissue oxygenation over time that occurs during shock. When O₂ delivery is limited, O₂ consumption can be inadequate to match the metabolic needs of cellular respiration, creating a deficit in O₂ requirements at the cellular level. The measurement of O₂ deficit uses calculation of the difference between the estimated O₂ demand and the actual value obtained for O₂ consumption. Under normal circumstances, cells can “repay” the O₂ debt during reperfusion. The magnitude of the O₂ debt correlates with the severity and duration of hypoperfusion. Surrogate values for measuring O₂ debt include base deficit and lactate levels and are discussed later in the “Hypovolemic/Hemorrhagic” section.

In addition to induction of changes in cellular metabolic pathways, shock also induces changes in cellular gene expression. The DNA binding activity of a number of nuclear transcription factors is altered by hypoxia and the production of O₂ radicals or nitrogen radicals that are produced at the cellular level by shock. Expression of other gene products such as heat shock proteins, vascular endothelial growth factor, inducible nitric oxide synthase (iNOS), heme oxygenase-1, and cytokines also are clearly increased by shock. Many of these shock-induced gene products, such as cytokines, have the ability to subsequently alter gene expression in specific target cells and tissues. The involvement of multiple pathways emphasizes the complex, integrated, and overlapping nature of the response to shock.

**Immune and Inflammatory Responses**

The inflammatory and immune responses are a complex set of interactions between circulating soluble factors and cells that can arise in response to trauma, infection, ischemia, toxic, or autoimmune stimuli. The processes are well regulated and can be conceptualized as an ongoing surveillance and response system that undergoes a coordinated escalation following injury to heal disrupted tissue or restore host-microbe equilibrium, as well as active suppression back to baseline levels. Failure to adequately control the activation, escalation, or suppression of the inflammatory response can lead to systemic inflammatory response syndrome and potentiate multiple organ failure.

Both the innate and adaptive branches of the immune system work in concert to rapidly respond in a specific and effective manner to challenges that threaten an organism’s well-being. Each arm of the immune system has its own set of functions, defined primarily by distinct classes of effector cells and their unique cell membrane receptor families. Alterations in the activity of the innate host immune system can be responsible for both the development of shock (i.e., septic shock following severe infection and traumatic shock following tissue injury with hemorrhage) and the pathophysiologic sequelae of shock such as the proinflammatory changes seen following hypoperfusion (see Fig. 5-1). When the predominantly paracrine mediators gain access to the systemic circulation, they can induce a variety of metabolic changes that are collectively referred to as the host inflammatory response. Understanding of the intricate, redundant, and interrelated pathways that comprise the inflammatory response to shock continues to expand. Despite limited understanding of how our current therapeutic interventions impact the host response to illness, inappropriate or excessive inflammation appears to be an essential event in the development of ARDS, multiple organ dysfunction syndrome (MODS), and posttraumatic immunosuppression that can prolong recovery.

Following direct tissue injury or infection, there are several mechanisms that lead to the activation of the active inflammatory and immune responses. These include release of bioactive peptides by neurons in response to pain and the release of intracellular molecules by broken cells, such as heat shock proteins, mitochondrial products, heparan sulfate, high mobility group box 1, and RNA. Only recently has it been realized that the release of intracellular products from damaged and injured cells can have paracrine and endocrine-like effects on distant tissues to activate the inflammatory and immune responses. This hypothesis, which was first proposed by Matzinger, is known as *danger signaling*. Under this novel paradigm of immune function, endogenous molecules are capable of signaling the presence of danger to surrounding cells and tissues. These molecules that are released from cells are known as *damage associated molecular patterns* (DAMPs, Table 5-3). DAMPs are recognized by cell surface receptors to effect intracellular signaling that primes and amplifies the immune response. These receptors are known as *pattern recognition receptors* (PRRs) and include the Toll-like receptors (TLRs) and the receptor for advanced glycation end products. Interestingly, TLRs and PRRs were first recognized for their role in signaling as part of the immune response to the
entry of microbes and their secreted products into a normally sterile environment. These bacterial products, including lipopolysaccharide, are known as pathogen-associated molecular patterns. The salutary consequences of PRR activation most likely relate to the initiation of the repair process and the mobilization of antimicrobial defenses at the site of tissue disruption. However, in the setting of excessive tissue damage, the inflammation itself may lead to further tissue damage, amplifying the response both at the local and systemic level. PR activation leads to intracellular signaling and release of cellular products, including cytokines (Fig. 5-4).

Before the recruitment of leukocytes into sites of injury, tissue-based macrophages or mast cells act as sentinel responders, releasing histamines, eicosanoids, tryptases, and cytokines (Fig. 5-5). Together these signals amplify the immune response by further activation of neurons and mast cells, as well as increasing the expression of adhesion molecules on the endothelium. Furthermore, these mediators cause leukocytes to release platelet-activating factor, further increasing the stickiness of the endothelium. Additionally, the coagulation and kinin cascades impact the interaction of endothelium and leukocytes.

**Cytokines/Chemokines**

The immune response to shock encompasses the elaboration of mediators with both proinflammatory and anti-inflammatory properties (Table 5-4). Furthermore, new mediators, new relationships between mediators, and new functions of known mediators are continually being identified. As new pathways are uncovered, understanding of the immune response to injury and the potential for therapeutic intervention by manipulating the immune response following shock will expand. What seems clear at present, however, is that the innate immune response can help restore homeostasis, or if it is excessive, promote cellular and organ dysfunction.

Multiple mediators have been implicated in the host immune response to shock. It is likely that some of the most important mediators have yet to be discovered and the roles of many known mediators have not been defined. A comprehensive description of all of the mediators and their complex interactions is beyond the scope of this chapter. For a general overview, a brief description of the more extensively studied mediators, as well as some of the known effects of these substances, see the following discussion. A more comprehensive review can be found in Chapter 2, “Systemic Response to Injury and Metabolic Support.”

Tumor necrosis factor alpha (TNF-α) was one of the first cytokines to be described, and it is one of the earliest cytokines released in response to injurious stimuli. Monocytes, macrophages, and T cells release this potent proinflammatory cytokine. TNF-α levels peak within 90 minutes of stimulation and return frequently to baseline levels within 4 hours. Release of TNF-α may be induced by bacteria or endotoxin and leads to the development of shock and hypoperfusion, most commonly observed in septic shock. Production of TNF-α also may be induced following other insults, such as hemorrhage and ischemia. TNF-α levels correlate with mortality in animal models of hemorrhage. In contrast, the increase in serum TNF-α levels reported in trauma patients is far less than that seen in septic patients. Once released, TNF-α can produce peripheral vaso-dilation, activate the release of other cytokines, induce procoagulant activity, and stimulate a wide array of cellular metabolic changes. During the stress response, TNF-α contributes to the muscle protein breakdown and cachexia.

Interleukin-1 (IL-1) has actions similar to those of TNF-α. IL-1 has a very short half-life (6 minutes) and primarily acts in a paracrine fashion to modulate local cellular responses. Systemically, IL-1 produces a febrile response to injury by activating prostaglandins in the posterior hypothalamus, and causes anorexia by activating the satiety center. This cytokine also augments the secretion of ACTH, glucocorticoids, and β-endorphins. In conjunction with TNF-α, IL-1 can stimulate the release of other cytokines such as IL-2, IL-4, IL-6, IL-8, granulocyte-macrophage colony-stimulating factor, and interferon-γ.

IL-2 is produced by activated T cells in response to a variety of stimuli and activates other lymphocyte subpopulations and natural killer cells. The lack of clarity regarding the role of IL-2 in the response to shock is intimately associated with that of understanding immune function after injury. Some investigators have postulated that increased IL-2 secretion promotes shock-induced tissue injury and the development of shock. Others have demonstrated that depressed IL-2 production is associated with,

---

**Table 5-3**

<table>
<thead>
<tr>
<th>Endogenous damage-associated molecular pattern molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mitochondrial DNA</td>
</tr>
<tr>
<td>Hyaluronan oligomers</td>
</tr>
<tr>
<td>Heparan sulfate</td>
</tr>
<tr>
<td>Extra domain A of fibronectin</td>
</tr>
<tr>
<td>Heat shock proteins 60, 70, Gp96</td>
</tr>
<tr>
<td>Surfactant Protein A</td>
</tr>
<tr>
<td>β-Defensin 2</td>
</tr>
<tr>
<td>Fibrinogen</td>
</tr>
<tr>
<td>Biglycan</td>
</tr>
<tr>
<td>High mobility group box 1</td>
</tr>
<tr>
<td>Uric acid</td>
</tr>
<tr>
<td>Interleukin-1α</td>
</tr>
<tr>
<td>S-100s</td>
</tr>
<tr>
<td>Nucleolin</td>
</tr>
</tbody>
</table>

---

**Figure 5-4.** A schema of information flow between immune cells in early inflammation following tissue injury and infection. Cells require multiple inputs and stimuli before activation of a full response. DAMPs = damage associated molecular patterns; HMGB1 = high mobility group box 1; TNF = tumor necrosis factor.
and perhaps contributes to, the depression in immune function after hemorrhage that may increase the susceptibility of patients who develop shock to suffer infections.\textsuperscript{26,27} It has been postulated that overly exuberant proinflammatory activation promotes tissue injury, organ dysfunction, and the subsequent immune dysfunction/suppression that may be evident later.\textsuperscript{22} Emphasizing the importance of temporal changes in the production of mediators, both the initial excessive production of IL-2 and later depressed IL-2 production are probably important in the progression of shock.

*Figure 5-5.* Signaling via the pattern recognition receptor TLR4. LPS signaling via TLR4 requires the cofactors LPS binding protein (LBP), MD-2, and CD14. Endogenous danger signals released from a variety of sources also signal in a TLR4-dependent fashion, although it is as yet unknown what cofactors may be required for this activity. Once TLR4 is activated, an intracellular signaling cascade is initiated that involves both a MyD88-dependent and independent pathway. DAMP = damage associated molecular pattern; LPS = lipopolysaccharide; MD-2 = myeloid differentiation factor-2; MyD88 = myeloid differentiation primary response gene 88; NF-kB = nuclear factor kB; TLR4 = Toll-like receptor-4. (Reproduced with permission from Mollen KP, Levy RM, Prince JM, et al. Systemic inflammation and end organ damage following trauma involves functional TLR4 signaling in both bone marrow-derived cells and parenchymal cells, J Leukoc Biol. 2008 Jan;83(1):80-88.)
IL-6 is elevated in response to hemorrhagic shock, major operative procedures, or trauma. Elevated IL-6 levels correlate with mortality in shock states. IL-6 contributes to lung, liver, and gut injury after hemorrhagic shock.\(^{28}\) Thus, IL-6 may play a role in the development of diffuse alveolar damage and ARDS. IL-6 and IL-1 are mediators of the hepatic acute phase response to injury, and enhance the expression and activity of complement, C-reactive protein, fibrinogen, haptoglobin, amyloid A, and alpha\(_1\)-antitrypsin, and promote neutrophil activation.\(^{29}\)

IL-10 is considered an anti-inflammatory cytokine that may have immunosuppressive properties. Its production is increased after shock and trauma, and it has been associated with depressed immune function clinically, as well as an increased susceptibility to infection.\(^{30}\) IL-10 is secreted by T cells, monocytes, and macrophages, and inhibits proinflammatory cytokine secretion, O\(_2\) radical production by phagocytes, adhesion molecule expression, and lymphocyte activation.\(^{30,31}\) Administration of IL-10 depresses cytokine production and improves some aspects of immune function in experimental models of shock and sepsis.\(^{32,33}\)

Recent studies point to the importance of chemokines, a specific set of cytokines, that have the ability to induce chemotaxis of leukocytes. Chemokines bind to specific chemokine receptors and transduce chemotactic signals to leukocytes. The significance of this large family of chemoattractant cytokines in immunology is difficult to understand, as almost every facet of the immune system is influenced by chemokines, including immune system development, immune surveillance, immune priming, effector responses, and immune regulation.\(^{34}\)

### Complement

The complement cascade can be activated by injury, shock, and severe infection, and contributes to host defense and proinflammatory activation. Significant complement consumption occurs after hemorrhagic shock.\(^{35}\) In trauma patients, the degree of complement activation is proportional to the magnitude of injury and may serve as a marker for severity of injury. Patients in septic shock also demonstrate activation of the complement pathway, with elevations of the activated complement proteins C3a and C5a. Activation of the complement cascade can contribute to the development of organ dysfunction. Activated complement factors C3a, C4a, and C5a are potent mediators of increased vascular permeability, smooth muscle cell contraction, histamine and arachidonic acid by-product release, and adherence of neutrophils to vascular endothelium. Activated complement acts synergistically with endotoxin to induce the release of TNF-\(\alpha\) and IL-1. The development of ARDS and MODS in trauma patients correlates with the intensity of complement activation.\(^{36}\) Complement and neutrophil activation may correlate with mortality in multiply injured patients.

### Neutrophils

Neutrophil activation is an early event in the upregulation of the inflammatory response; neutrophils are the first cells to be recruited to the site of injury. Polymorphonuclear leukocytes (PMNs) remove infectious agents, foreign substances that have penetrated host barrier defenses, and nonviable tissue through phagocytosis. However, activated PMNs and their products may also produce cell injury and organ dysfunction. Activated PMNs generate and release a number of substances that may induce cell or tissue injury, such as reactive O\(_2\) species, lipid-peroxidation products, proteolytic enzymes (elastase, cathepsin G), and vasoactive mediators (leukotrienes, eicosanoids, and platelet-activating factor). Oxygen-free radicals, such as superoxide anion, hydrogen peroxide, and hydroxyl radical, are released and induce lipid peroxidation, inactivate enzymes, and consume antioxidants (such as glutathione and tocopherol). Ischemia-reperfusion activates PMNs and causes PMN-induced organ injury. In animal models of hemorrhagic shock, activation of PMNs correlates with irreversibility of shock and mortality, and neutrophil depletion prevents the pathophysiologic sequelae of hemorrhagic and septic shock. Human data corroborate the activation of neutrophils in trauma and shock and suggest a role in the development of MODS.\(^{37}\) Plasma markers of PMN activation, such as elastase, correlate with severity of injury in humans.

Interactions between endothelial cells and leukocytes are important in the inflammatory process. The vascular endothelium contributes to regulation of blood flow, leukocyte adherence, and the coagulation cascade. Extracellular ligands such as intercellular adhesion molecules, vascular cell adhesion molecules, and the selectins (E-selectin, P-selectin) are expressed on the surface of endothelial cells, and are responsible for leukocyte adhesion to the endothelium. This interaction allows activated neutrophils to migrate into the tissues to combat infection, but also can lead to PMN-mediated cytotoxicity and microvascular and tissue injury.

### Cell Signaling

A host of cellular changes occur following shock. Although many of the intracellular and intercellular pathways that are important in shock are being elucidated, undoubtedly there are many more that have yet to be identified. Many of the mediators produced during shock interact with cell surface receptors on target cells to alter target cell metabolism. These signaling pathways may be altered by changes in cellular oxygenation, redox state, high-energy phosphate concentration, gene expression, or intracellular electrolyte concentration induced by shock. Cells communicate with their external environment through the use of cell surface membrane receptors, which, once bound by a ligand, transmit their information to the interior of the cell through a variety of signaling cascades. These signaling pathways may subsequently alter the activity of specific enzymes, the expression or breakdown of important proteins, or affect intracellular energy metabolism. Intracellular calcium (Ca\(^{2+}\)) homeostasis and regulation represents one such pathway. Intracellular Ca\(^{2+}\) concentrations regulate many aspects of cellular metabolism; many important enzyme systems require Ca\(^{2+}\) for
full activity. Profound changes in intracellular Ca\(^{2+}\) levels and Ca\(^{2+}\) transport are seen in models of shock.\(^{38}\) Alterations in Ca\(^{2+}\) regulation may lead to direct cell injury, changes in transcription factor activation, alterations in the expression of genes important in homeostasis, and the modulation of the activation of cells by other shock-induced hormones or mediators.\(^{39,40}\)

A proximal portion of the intracellular signaling cascade consists of a series of kinases that transmit and amplify the signal through the phosphorylation of target proteins. The O\(_2\) radicals produced during shock and the intracellular redox state are known to influence the activity of components of this cascade, such as protein tyrosine kinases, mitogen activated kinases, and protein kinase C.\(^{41,42}\) Either through changes in these signaling pathways, changes in the activation of enzyme systems through Ca\(^{2+}\)-mediated events, or direct conformational changes to oxygen-sensitive proteins, O\(_2\) radicals also regulate the activity of a number of transcription factors that are important in gene expression, such as nuclear factor \(\kappa\)B, APETALA1, and hypoxia-inducible factor 1.\(^{43,44}\) It is therefore becoming increasingly clear that oxidant-mediated direct cell injury is merely one consequence of the production of O\(_2\) radicals during shock.

The study of the effects of shock on the regulation of gene expression as an important biologic effect was stimulated by the work of Buchman and colleagues.\(^{46}\) The effects of shock on the expression and regulation of numerous genes and gene products has been studied in both experimental animal models and human patients. These studies include investigations into single genes of interest as well as large-scale genomic and proteomic analysis.\(^{55-57}\) Changes in gene expression are critical for adaptive and survival cell signaling. Polymorphisms in gene promoters that lead to a differential level of expression of gene products are also likely to contribute significantly to varied responses to similar insults.\(^{58,59}\)

In a recent study, the genetic responses to traumatic injury in humans or endotoxin delivery to healthy human volunteers demonstrated that severe stresses produce a global re prioritization affecting >80% of the cellular functions and pathways.\(^{60}\) The similarities in genomic responses among different injuries revealed a fundamental human response to stressors involving dysregulated immune responses (Fig. 5-6). Furthermore, in traumatic injury patients, complications like nosocomial infections and organ failure were not associated with any genomic evidence of a second hit and differed only in the magnitude and duration of this genomic reprioritization.

**FORMS OF SHOCK**

**Hypovolemic/Hemorrhagic**

The most common cause of shock in the surgical or trauma patient is loss of circulating volume from hemorrhage. Acute blood loss results in reflexive decreased baroreceptor stimulation from stretch receptors in the large arteries, resulting in decreased inhibition of vasoconstrictor centers in the brain stem, increased chemoreceptor stimulation of vasomotor centers, and diminished output from atrial stretch receptors. These changes increase vasoconstriction and peripheral arterial resistance. Hypovolemia also induces sympathetic stimulation, leading to epinephrine and norepinephrine release, activation of the renin-angiotensin cascade, and increased vasopressin release. Peripheral vasoconstriction is prominent, while lack of sympathetic effects on cerebral and coronary vessels and local autoregulation promote maintenance of cardiac and CNS blood flow.

---

**Figure 5-6.** The concurrent dysregulated innate immune responses that promote inflammation and dysregulated adaptive immune responses that result in immunosuppression occur in patients following traumatic injury. However, these genetic responses can result in complicated outcomes in trauma patients if the magnitude or duration of these responses are pronounced. (Reproduced with permission from Xiao W, Mirdinos MN, Seok J, et al. A genomic storm in critically injured humans. J Exp Med. 2011 Dec 19;208(13):2581-2590.)

**Diagnosis.** Treatment of shock is initially empiric. A secure airway must be confirmed or established in obtunded patients. The priority is the initiation of volume infusion while the search for the cause of the hypotension is pursued. Shock in a trauma patient or postoperative patient should be presumed to be due to hemorrhage until proven otherwise. The clinical signs of shock may be evidenced by agitation, cool clammy extremities, tachycardia, weak or absent peripheral pulses, and hypotension. Such apparent clinical shock results from at least 25% to 30% loss of the blood volume. However, substantial volumes of blood may be lost before the classic clinical manifestations of shock are evident. Thus, when a patient is significantly tachycardic or hypotensive, this represents both significant blood loss and physiologic decompensation. The clinical and physiologic response to hemorrhage has been classified according to the magnitude of volume loss. Loss of up to 15% of the circulating volume (700 to 750 mL for a 70-kg patient) may produce little in terms of obvious symptoms, while loss of up to 30% of the circulating volume (1.5 L) may result in mild tachycardia, tachypnea, and anxiety. Hypotension, marked tachycardia (i.e., pulse greater than 110 to 120 bpm), and confusion may not be evident until more than 30% of the blood volume has been lost; loss of 40% of circulating volume (2 L) is immediately life threatening and generally requires operative control of bleeding (Table 5-5). Young healthy patients with vigorous compensatory mechanisms may tolerate larger volumes of blood loss while manifesting fewer clinical signs despite the presence of significant peripheral hypoperfusion. These patients may maintain a near-normal blood pressure until a precipitous cardiovascular collapse occurs. Elderly patients may be taking medications that either promote bleeding (e.g., warfarin or aspirin), or mask the compensatory responses to bleeding (e.g., β-blockers). In addition, atherosclerotic vascular disease, diminishing cardiac compliance with age, inability to elevate heart rate or cardiac contractility in response to hemorrhage, and overall decline in physiologic reserve decrease the elderly...
patient’s ability to tolerate hemorrhage. Recent data in trauma patients suggest that a systolic blood pressure (SBP) of less than 110 mmHg is a clinically relevant definition of hypotension and hypoperfusion based upon an increasing rate of mortality below this pressure (Fig. 5-7).55

In addressing the sensitivity of vital signs and identifying major thoracoabdominal hemorrhage, a study retrospectively identified patients with injury to the trunk and an abbreviated injury score of 3 or greater who required immediate surgical intervention and transfusion of at least 5 units of blood within the first 24 hours. Ninety-five percent of patients had a heart rate greater than 80 bpm at some point during their postinjury course. However, only 59% of patients achieved a heart rate greater than 120 bpm. Ninety-nine percent of all patients had a recorded blood pressure of less than 120 mmHg at some point. Ninety-three percent of all patients had a recorded SBP of less than 110 mmHg.54 A more recent study corroborated that tachycardia was not a reliable sign of hemorrhage following trauma and was present in only 65% of hypotensive patients.55

Serum lactate and base deficit are measurements that are helpful to both estimate and monitor the extent of bleeding and shock. The amount of lactate that is produced by anaerobic respiration is an indirect marker of tissue hypoperfusion, cellular O2 debt, and the severity of hemorrhagic shock. Several studies have demonstrated that the initial serum lactate and serial lactate levels are reliable predictors of morbidity and mortality with hemorrhage following trauma (Fig. 5-8).56 Similarly, base deficit values derived from arterial blood gas analysis provide clinicians with an indirect estimation of tissue acidosis from hypoperfusion. Davis and colleagues stratified the extent of base deficit into mild (−3 to −5 mmol/L), moderate (−6 to −9 mmol/L), and severe (less than −10 mmol/L), and from this established a correlation between base deficit upon admission with transfusion requirements, the development of multiple organ failure, and death (Fig. 5-9).57 Both base deficit and lactate correlate with the extent of shock and patient outcome, but interestingly do not firmly correlate with each other.58-60 Evaluation of both values may be useful in trauma patients with hemorrhage.

Although hematocrit changes may not rapidly reflect the total volume of blood loss, admission hematocrit has been shown to be associated with 24 hour fluid and transfusion requirements and more strongly associated with PRBC transfusion than either tachycardia, hypotension, or acidosis.61 It must be noted that lack of a depression in the initial hematocrit does not rule out substantial blood loss or ongoing bleeding.

In management of trauma patients, understanding the patterns of injury of the patient in shock will help direct the evaluation and management. Identifying the sources of blood loss in patients with penetrating wounds is relatively simple because potential bleeding sources will be located along the known or suspected path of the wounding object. Patients with penetrating injuries who are in shock usually require operative intervention. Patients who suffer multisystem injuries from blunt trauma have multiple sources of potential hemorrhage. Blood loss sufficient to cause shock is generally of a large volume, and there are a limited number of sites that can harbor sufficient extravascular blood volume to induce hypotension (e.g., external, intrathoracic, infra-abdominal, retroperitoneal, and long bone fractures). In the nontrauma patient, the GI tract must always be considered as a site for blood loss. Substantial blood loss externally may be suspected from prehospital medical reports documenting a substantial blood loss at the scene of an accident, history of massive

![Figure 5-7](https://example.com/figure5-7.png)
blood loss from wounds, visible brisk bleeding, or presence of a large hematoma adjacent to an open wound. Injuries to major arteries or veins with associated open wounds may cause massive blood loss rapidly. Direct pressure must be applied and sustained to minimize ongoing blood loss. Tourniquets should be used for extremity bleeding stopped by direct pressure and applied in the prehospital setting as needed. Persistent bleeding from uncontrolled smaller vessels can, over time, precipitate shock if inadequately treated.

When major blood loss is not immediately visible in the setting of trauma, internal (intracavitary) blood loss should be suspected. Each pleural cavity can hold 2 to 3 L of blood and can therefore be a site of significant blood loss. Diagnostic and therapeutic tube thoracostomy may be indicated in unstable patients based on clinical findings and clinical suspicion. In a more stable patient, a chest radiograph may be obtained to look for evidence of hemothorax. Major retroperitoneal hemorrhage typically occurs in association with pelvic fractures, which is confirmed by pelvic radiography in the resuscitation bay. Intraperitoneal hemorrhage is probably the most common source of blood loss that induces shock. The physical exam for detection of substantial blood loss or injury is insensitive and unreliable; large volumes of intraperitoneal blood may be present before physical examination findings are apparent. Findings with intra-abdominal hemorrhage include abdominal distension, abdominal tenderness, or visible abdominal wounds. Hemodynamic abnormalities generally stimulate a search for blood loss before the appearance of obvious abdominal findings. Adjunctive tests are essential in the diagnosis of intraperitoneal bleeding; intraperitoneal blood may be rapidly identified by diagnostic ultrasound or diagnostic peritoneal lavage. Furthermore, patients that have sustained high-energy blunt trauma that are hemodynamically stable or that have normalized their vital signs in response to initial volume resuscitation should undergo computed tomography scans to assess for head, chest, and/or abdominal bleeding.

**Treatment.** Control of ongoing hemorrhage is an essential component of the resuscitation of the patient in shock. As mentioned in the previous “Diagnosis” section, treatment of hemorrhagic shock is instituted concurrently with diagnostic evaluation to identify a source. Patients who fail to respond to initial resuscitative efforts should be assumed to have ongoing active hemorrhage from large vessels and require prompt operative intervention. Based on trauma literature, patients with ongoing hemorrhage demonstrate increased survival if the elapsed time between the injury and control of bleeding is decreased. Although there are no randomized controlled trials, retrospective studies provide compelling evidence in this regard. To this end, Clarke and colleagues demonstrated that trauma

---

**Figure 5-8.** Progressive increases in serum lactate, muscle lactate, and liver lactate in a baboon model of hemorrhagic shock. (Reproduced with permission from Petzman AB, Corbett WA, Shires GT, et al: Cellular function in liver and muscle during hemorrhagic shock in primates, Surg Gynecol Obstet. 1985 Nov;161(5):419-424.)

**Figure 5-9.** The relationship between base deficit (negative base excess) and mortality in trauma patients. BEA = base excess arterial; ECF = extracellular fluid. (Reproduced with permission from Siegel JH, Rivkind AI, Dalal S, et al: Early physiologic predictors of injury severity and death in blunt multiple trauma, Arch Surg. 1990 Apr;125(4):498-508.)
patients with major injuries isolated to the abdomen requiring emergency laparotomy had an increased probability of death with increasing length of time in the emergency department for patients who were in the emergency department for 90 minutes or less. This probability increased approximately 1% for each 3 minutes in the emergency department.

The appropriate priorities in these patients are as follows: (a) control the source of blood loss, (b) perform IV volume resuscitation with blood products in the hypotensive patient, and (c) secure the airway. In trauma, identifying the body cavity harboring active hemorrhage will help focus operative efforts; however, because time is of the essence, rapid treatment is essential, and diagnostic laparotomy or thoracotomy may be indicated. The actively bleeding patient cannot be resuscitated until control of ongoing hemorrhage is achieved. Our current understanding has led to the management strategy known as damage control resuscitation.65 This strategy begins in the emergency department, continues into the operating room, and into the intensive care unit (ICU). Initial resuscitation is limited to keep SBP around 80 to 90 mmHg. This prevents renewed bleeding from recently clotted vessels. Resuscitation and intravascular volume resuscitation is accomplished with blood products and limited crystalloids, which is addressed further later in this section. Too little volume allowing persistent severe hypotension and hypoperfusion is dangerous, yet too vigorous of a volume resuscitation may be just as deleterious. Control of hemorrhage is achieved in the operating room (or angiography suite once surgical causes of hemorrhage have been ruled out), and efforts to warm patients and to prevent coagulopathy using multiple blood products and pharmacologic agents are used in both the operating room and ICU.

Cannon and colleagues first made the observation that attempts to increase blood pressure in soldiers with uncontrolled sources of hemorrhage is counterproductive, with increased bleeding and higher mortality.6 The work was the foundation for the “hypotensive resuscitation” strategies. Several laboratory studies confirmed the observation that attempts to restore normal blood pressure with fluid infusion or vasopressors was rarely achievable and resulted in more bleeding and higher mortality.66 A prospective, randomized clinical study compared delayed fluid resuscitation (upon arrival in the operating room) with standard fluid resuscitation (with arrival by the paramedics) in hypotensive patients with penetrating torso injury.57 The authors reported that delayed fluid resuscitation resulted in lower patient mortality. Further laboratory studies demonstrated that fluid restriction in the setting of profound hypotension resulted in early deaths from severe hypoperfusion. These studies also showed that aggressive crystalloid resuscitation attempting to normalize blood pressure resulted in marked hemodilution, with hematocrits of 5%.66 Reasonable conclusions in the setting of uncontrolled hemorrhage include (a) any delay in surgery for control of hemorrhage increases mortality; (b) with uncontrolled hemorrhage, attempting to achieve normal blood pressure may increase mortality, particularly with penetrating injuries and short transport times; (c) a goal of SBP of 80 to 90 mmHg may be adequate in the patient with penetrating injury; and (d) profound hemodilution should be avoided by early transfusion of red blood cells. For the patient with blunt injury, where the major cause of death is a closed head injury, the increase in mortality with hypotension in the setting of brain injury must be avoided. In this setting, a SBP of 110 mmHg would seem to be more appropriate.

Patients who respond to initial resuscitative effort but then deteriorate hemodynamically frequently have injuries that require operative intervention. The magnitude and duration of their response will dictate whether diagnostic maneuvers can be performed to identify the site of bleeding. However, hemodynamic deterioration generally denotes ongoing bleeding for which some form of intervention (i.e., operation or interventional radiology) is required. Patients who have lost significant intravascular volume, but whose hemorrhage is controlled or has abated, often will respond to resuscitative efforts if the depth and duration of shock have been limited.

A subset of patients exists who fail to respond to resuscitative efforts despite adequate control of ongoing hemorrhage. These patients have ongoing fluid requirements despite adequate control of hemorrhage, have persistent hypotension despite restoration of intravascular volume necessitating vasopressor support, and may exhibit a futile cycle of uncorrectable hypothermia, hypoperfusion, acidosis, and coagulopathy that cannot be interrupted despite maximum therapy. These patients have deteriorated to decompensated or irreversible shock with peripheral vasodilation and resistance to vasopressor infusion. Mortality is almost inevitable once the patient manifests shock in its terminal stages. Unfortunately, this is often diagnosed in retrospect.

Fluid resuscitation is a major adjunct to physically controlling hemorrhage in patients with shock. The ideal type of fluid to be used continues to be debated; however, crystalloids continue to be the mainstay fluid of choice. Several studies have demonstrated increased risk of death in bleeding trauma patients treated with colloid compared to patients treated with crystalloid.68 In patients with severe hemorrhage, restoration of intravascular volume should be achieved with blood products.69

Ongoing studies continue to evaluate the use of hypertonic saline as a resuscitative adjunct in bleeding patients.70 The benefit of hypertonic saline solutions may be immunomodulatory. Specifically, these effects have been attributed to pharmacologic effects resulting in decreased reperfusion-mediated injury with decreased O2 radical formation, less impairment of immune function compared to standard crystalloid solution, and less brain swelling in the multi-injured patient. The reduction of total volume used for resuscitation makes this approach appealing as a resuscitation agent for combat injuries and may contribute to a decrease in the incidence of ARDS and multiple organ failure.

Transfusion of packed red blood cells and other blood products is essential in the treatment of patients in hemorrhagic shock. Current recommendations in stable ICU patients aim for a target hemoglobin of 7 to 9 g/dL.71 72 However, no prospective randomized trials have compared restrictive and liberal transfusion regimens in trauma patients with hemorrhagic shock. The current standard in severely injured patients is termed damage control resuscitation and consists of transfusion with red blood cells, fresh frozen plasma (FFP), and platelet units given in equal number.73 Civilian and military trauma data show that the development of coagulopathy of trauma is predictive of mortality.74 Data collected from a U.S. Army combat support hospital helped to propagate this practice, showing in patients that received massive transfusion of packed red blood cells (>10 units in 24 hours) that a high plasma to RBC ratio (1:1.4 units) was independently associated with improved survival.75 A number of civilian studies have demonstrated similar results.76 Similarly, platelet transfusion is important. The Pragmatic Randomized Optimal Platelet and Plasma Ratios (PROPR)
trial has demonstrated that damage-control resuscitation, a massive transfusion strategy targeting a balanced delivery of plasma-platelet-red blood cell in a ratio of 1:1:1, results in improved survival at 3 hours and a reduction in deaths caused by exsanguination in the first 24 hours. Studies have demonstrated that low platelet counts in trauma patients was associated with increased mortality and that increased platelet use appears to improve outcome. The benefit of platelet transfusion may be most pronounced in trauma patients with brain injury. It has been suggested that platelets should be transfused in the bleeding patient to maintain counts above 50 × 10^9/L. In addition, transfusion of whole blood is gaining in popularity, and ongoing studies are evaluating the benefit of this approach.

There is a potential role for other coagulation factor based products, such as fibrinogen concentrates and prothrombin complex concentrates. Use of these agents may be guided by a drop in fibrinogen levels to less than 1 g/L, or less specifically by thromboelastogram findings to suggest hyperfibrinolysis. Data also support the use of antifibrinolytic agents in bleeding trauma patients, specifically tranexamic acid (a synthetic lysine analogues that acts as a competitive inhibitor of plasmin and plasminogen). The multinational Clinical Randomization of an Antifibrinolytic in Significant Haemorrhage 2 (CRASH-2) trial suggested that early use of tranexamic acid limits rebleeding and reduces mortality (Fig 5-10). In the past, coagulopathy associated with the bleeding patient was presumed to be due solely to dilution and depletion of clotting factors and platelets. We now understand that an acute coagulopathy of trauma occurs as an immediate consequence of injury, with abnormal admission coagulation as a predictor of high mortality. Traditional measurement of platelets, INR, and PTT may not reflect the coagulopathy of trauma or response to therapy effectively. Recently, viscoelastic functional testing such as thromboelastography (TEG) has been utilized as a quicker, more comprehensive determination of coagulopathy and fibrinolysis in the injured patient. Holcomb et al recently reported that TEG predicted patients with substantial bleeding and red cell transfusion better than conventional coagulopathy tests, and it also predicted the need for platelet transfusion better than platelet count and the need for plasma transfusion better than fibrinogen levels.

Additional resuscitative adjuncts in patients with hemorrhagic shock include minimization of heat loss and maintaining normothermia. The development of hypothermia in the bleeding patient is associated with acidosis, hypotension, and coagulopathy. Hypothermia in bleeding trauma patients is an independent risk factor for bleeding and death. This likely is secondary to impaired platelet function and impairments in the coagulation cascade. Several studies have investigated the induction of controlled hypothermia in patients with severe shock based on the hypothesis of limiting metabolic activity and energy requirements, creating a state of “suspended animation.” These studies are promising and continue to be evaluated in large trials.

**Traumatic Shock**

The systemic response after trauma, combining the effects of soft tissue injury, long bone fractures, and blood loss, is clearly a different physiologic insult than simple hemorrhagic shock. Multiple organ failure, including acute respiratory distress syndrome (ARDS), develops relatively often in the blunt trauma patient, but rarely after pure hemorrhagic shock (such as a GI bleed). The hypoperfusion deficit in traumatic shock is magnified by the proinflammatory activation that occurs following the induction of shock. In addition to ischemia or ischemia-reperfusion, accumulating evidence demonstrates that even simple hemorrhage induces proinflammatory activation that results in many of the cellular changes typically ascribed only to septic shock. At the cellular level, this may be attributable to the release of cellular products termed damage associated molecular patterns (DAMPs, i.e., ribonucleic acid, uric acid, and high mobility group box 1) that activate the same set of cell surface receptors as bacterial products, initiating similar cell signaling. These receptors are termed pattern recognition receptors (PRRs) and include the TLR family of proteins. Examples of traumatic shock include small volume hemorrhage accompanied by soft tissue injury (femur fracture, crush injury), or any combination of hypovolemic, neurogenic, cardiogenic, and obstructive shock that precipitate rapidly progressive proinflammatory activation. In laboratory models of traumatic shock, the addition of a soft tissue or long bone injury to hemorrhage produces lethality with significantly less blood loss when the animals are stressed by hemorrhage. Treatment of traumatic shock is focused on correction of the individual elements to diminish the cascade of proinflammatory activation, and includes prompt control of hemorrhage, adequate volume resuscitation to correct O₂ debt, debridement of nonviable tissue, stabilization of bony injuries, and appropriate treatment of soft tissue injuries.

**Septic Shock (Vasodilatory Shock)**

In the peripheral circulation, profound vasoconstriction is the typical physiologic response to the decreased arterial pressure and tissue perfusion with hemorrhage, hypovolemia, or acute heart failure. This is not the characteristic response in vasodilatory shock. Vasodilatory shock is the result of dysfunction of the endothelium and vasculature secondary to circulating inflammatory mediators and cells or as a response to prolonged and severe hypoperfusion. Thus, in vasodilatory shock, hypotension results from failure of the vascular smooth muscle to constrict appropriately. Vasodilatory shock is characterized by peripheral vasodilation with resultant hypotension and resistance to treatment with vasopressors. Despite the hypotension, plasma catecholamine levels are elevated, and the renin-angiotensin system
is activated in vasodilatory shock. The most frequently encountered form of vasodilatory shock is septic shock. Other causes of vasodilatory shock include hypoxic lactic acidosis, carbon monoxide poisoning, uncompensated and irreversible hemorrhagic shock, terminal cardiogenic shock, and postcardiomyotomy shock (Table 5-6). Thus, vasodilatory shock seems to represent the final common pathway for profound and prolonged shock of any etiology.88

Despite advances in intensive care, the mortality rate for severe sepsis remains at 30% to 50%. In the United States, 750,000 cases of sepsis occur annually, one-third of which are fatal.89 Sepsis accounts for 9.3% of deaths in the United States, as many yearly as MI. Septic shock is a by-product of the body’s response to disruption of the host-microbe equilibrium, resulting in invasive or severe localized infection.

In the attempt to eradicate the pathogens, the immune and other cell types (e.g., endothelial cells) elaborate soluble mediators that enhance macrophage and neutrophil killing effector mechanisms, increase procoagulant activity and fibroblast activity to localize the invaders, and increase microvascular blood flow to enhance delivery of killing forces to the area of invasion. When this response is overly exuberant or becomes systemic rather than localized, manifestations of sepsis may be evident. These findings include enhanced cardiac output, peripheral vasodilation, fever, leukocytosis, hyperglycemia, and tachycardia. In septic shock, the vasodilatory effects are due, in part, to the upregulation of the inducible isoform of nitric oxide synthase (iNOS or NOS 2) in the vessel wall. iNOS produces large quantities of nitric oxide for sustained periods of time. This potent vasodilator suppresses vascular tone and renders the vasculature resistant to the effects of vasoconstricting agents. Additionally, endothelial activation or injury likely contributes to some degree of vascular dysfunction.

Diagnosis. Attempts to standardize terminology have led to the establishment of criteria for the diagnosis of sepsis in the hospitalized adult. These criteria include manifestations of the host response to infection in addition to identification of an offending organism. The terms sepsis and septic shock are used to quantify the magnitude of the systemic inflammatory reaction. Patients with sepsis have evidence of an infection, as well as systemic signs of inflammation (e.g., fever, leukocytosis, and tachycardia). Hypoperfusion with signs of organ dysfunction is termed severe sepsis. Septic shock requires the presence of the aforementioned signs, associated with more significant evidence of tissue hypoperfusion and systemic hypotension. Beyond the hypotension, maldistribution of blood flow and shunting in the microcirculation further compromise delivery of nutrients to the tissue beds.90,91 The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3) have refined the definitions and utilize a Sequential (Sepsis-related) Organ Failure Assessment (SOFA) score to help determine signs of organ dysfunction and to guide management.92

Recognizing septic shock begins with defining the patient at risk. The clinical manifestations of septic shock will usually become evident and prompt the initiation of treatment before bacteriologic confirmation of an organism or the source of an organism is identified. In addition to fever, tachycardia, and tachypnea, signs of hypoperfusion such as confusion, malaise, oliguria, or hypotension may be present. These should prompt an aggressive search for infection, including a thorough physical examination, inspection of all wounds, evaluation of intravascular catheters or other foreign bodies, obtaining appropriate cultures, and adjunctive imaging studies, as needed.

Treatment. Evaluation of the patient in septic shock begins with an assessment of the adequacy of their airway and ventilation. Severely obtunded patients and patients whose work of breathing is excessive require intubation and ventilation to prevent respiratory collapse. Because vasodilation and decrease in total peripheral resistance may produce hypotension, fluid resuscitation and restoration of circulatory volume with balanced salt solutions is essential. The Surviving Sepsis Campaign has updated treatment recommendations and care bundles with a most recent goal for care within the first hour93,94 (Table 5-7). Serum lactate should be measured as a marker of shock. Fluid resuscitation should begin within the first hour and should be at least 30 mL/kg for hypotensive patients. Incremental fluid boluses should be continued based upon endpoint of resuscitation, including clearance of lactate. Starch-based colloid solutions should be avoided as recent evidence suggests that these fluids may be deleterious in the setting of sepsis.90-95,98 Blood cultures should be obtained. Empiric antibiotics must be chosen carefully based on the most likely pathogens (gram-negative rods, gram-positive cocci, and anaerobes) because the portal of entry of the offending organism and its identity may not be evident until culture data return or imaging studies are completed. Knowledge of the bacteriologic profile of infections in an individual unit can be obtained from most hospital infection control departments and will suggest potential responsible organisms. Antibiotics should be tailored to cover the responsible organisms once culture data are available, and if appropriate, the

### Table 5-6

**Causes of septic and vasodilatory shock**

<table>
<thead>
<tr>
<th>Category</th>
<th>Causes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systemic response to infection</td>
<td>Systemic response to infection</td>
</tr>
<tr>
<td>Noninfectious systemic inflammation</td>
<td>Systemic response to infection</td>
</tr>
<tr>
<td>Pancreatitis</td>
<td>Systemic response to infection</td>
</tr>
<tr>
<td>Burns</td>
<td>Systemic response to infection</td>
</tr>
<tr>
<td>Anaphylaxis</td>
<td>Systemic response to infection</td>
</tr>
<tr>
<td>Acute adrenal insufficiency</td>
<td>Systemic response to infection</td>
</tr>
<tr>
<td>Prolonged, severe hypotension</td>
<td>Systemic response to infection</td>
</tr>
<tr>
<td>Hemorrhagic shock</td>
<td>Systemic response to infection</td>
</tr>
<tr>
<td>Cardiogenic shock</td>
<td>Systemic response to infection</td>
</tr>
<tr>
<td>Cardiopulmonary bypass</td>
<td>Systemic response to infection</td>
</tr>
<tr>
<td>Metabolic</td>
<td>Systemic response to infection</td>
</tr>
<tr>
<td>Hypoxic lactic acidosis</td>
<td>Systemic response to infection</td>
</tr>
<tr>
<td>Carbon monoxide poisoning</td>
<td>Systemic response to infection</td>
</tr>
</tbody>
</table>

### Table 5-7

**Surviving Sepsis Campaign bundles of care to be initiated within the first hour after presentation in the patient with sepsis**

<table>
<thead>
<tr>
<th>Step</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Measure serum lactate level. Remeasure if the initial lactate is &gt;2 mmol/L.</td>
</tr>
<tr>
<td>2</td>
<td>Obtain blood culture prior to administration of antibiotics.</td>
</tr>
<tr>
<td>3</td>
<td>Administer broad spectrum antibiotics.</td>
</tr>
<tr>
<td>4</td>
<td>Rapid administration of 30 mL/kg crystalloid for hypotension or lactate ≥4 mmol/L.</td>
</tr>
<tr>
<td>5</td>
<td>Use vasopressors if the patient is hypotensive during or after fluid resuscitation to maintain a mean arterial pressure ≥65 mmHg.</td>
</tr>
</tbody>
</table>
spectrum of coverage narrowed. Long-term, empiric, broad-spectrum antibiotic use should be minimized to reduce the development of resistant organisms and to avoid the potential complications of fungal overgrowth and antibiotic-associated colitis from overgrowth of Clostridium difficile. IV antibiotics will be insufficient to adequately treat the infectious episode in the settings of infected fluid collections, infected foreign bodies, and devitalized tissue. These situations require source control and involve percutaneous drainage and operative management to target a focus of infection. These situations may require multiple operations to ensure proper wound hygiene and healing.

After first-line therapy of the septic patient with antibiotics, IV fluids, and intubation if necessary, vasopressors may be necessary to treat patients with septic shock. Catecholamines are the vasopressors used most often, with norepinephrine being the first-line agent followed by epinephrine. Occasionally, patients with septic shock will develop arterial resistance to catecholamines. Arginine vasopressin, a potent vasoconstrictor, is often efficacious in this setting and is often added to norepinephrine.

The majority of septic patients have hyperdynamic physiology with supranormal cardiac output and low systemic vascular resistance. On occasion, septic patients may have low cardiac output despite volume resuscitation and even vasoppressor support. Dobutamine therapy is recommended for patients with cardiac dysfunction as evidenced by high filling pressures and low cardiac output or clinical signs of hypoperfusion after achievement of restoration of blood pressure following fluid resuscitation. Mortality in this group is high. Despite the increasing incidence of septic shock over the past several decades, the overall mortality rates have changed little. Studies of interventions, including immunotherapy, resuscitation to pulmonary artery endpoints with hemodynamic optimization (cardiac output and O2 delivery, even to supranormal values), and optimization of mixed venous O2 measurements up to 72 hours after admission to the ICU, have not changed mortality.

The advances made in the treatment of patients with sepsis and septic shock and collaborative groups such as the Surviving Sepsis Campaign continue to evaluate, modify, and put forth recommendations based upon data. Negative results from previous studies have led to the suggestion that earlier interventions directed at improving global tissue oxygenation may be of benefit. To this end, Rivers and colleagues reported that goal-directed therapy of septic shock and severe sepsis initiated in the emergency department and continued for 6 hours significantly improved outcome.70 This approach involved adjustment of cardiac preload, afterload, and contractility to balance O2 delivery with O2 demand. They found that goal-directed therapy during the first 6 hours of hospital stay (initiated in the emergency department) had significant effects, such as higher mean venous O2 saturation, lower lactate levels, lower base deficit, higher pH, and decreased 28-day mortality (49.2% vs. 33.3%) compared to the standard therapy group. The frequency of sudden cardiovascular collapse was also significantly less in the group managed with goal-directed therapy (21.0% vs. 10.3%). Interestingly, the goal-directed therapy group received more IV fluids during the initial 6 hours, but the standard therapy group required more IV fluids by 72 hours. The authors emphasize that continued cellular and tissue decompensation is subclinical and often irreversible when obvious clinically. Goal-directed therapy allowed identification and treatment of these patients with insidious illness (global tissue hypoxia in the setting of normal vital signs).

Several multicenter trials have been performed to further refine these finding. In the Protocolized Care for Early Septic Shock (ProCESS) trial, a multicenter, randomized trial in which patients were identified early in the emergency department as having septic shock and received antibiotics and other nonresuscitation aspects of care promptly, the investigators found no significant advantage, with respect to mortality or morbidity, of protocol-based resuscitation over bedside care that was provided according to the treating physician’s judgment. They also found no significant benefit of the mandated use of central venous catheterization and central hemodynamic monitoring in all patients. This last finding was recapitulated in the Protocolised Management in Sepsis (ProMISE) trial.69 Failure of these more recent trials to show a benefit of early goal-directed protocols vs. standard of care may be secondary to the generalized improvement in early recognition of sepsis and institution of protocolized care by efforts such as the Surviving Sepsis Campaign.

Hyperglycemia and insulin resistance are typical in critically ill and septic patients, including patients without underlying diabetes mellitus. A recent study reported significant positive impact of tight glucose management on outcome in critically ill patients.100 The two treatment groups in this randomized, prospective study were assigned to receive intensive insulin therapy (maintenance of blood glucose between 80 and 110 mg/dL) or conventional treatment (infusion of insulin only if the blood glucose level exceeded 215 mg/dL, with a goal between 180 and 200 mg/dL). The mean morning glucose level was significantly higher in the conventional treatment as compared to the intensive insulin therapy group (153 vs. 103 mg/dL). Mortality in the intensive insulin treatment group (4.6%) was significantly lower than in the conventional treatment group (8.0%), representing a 42% reduction in mortality. This reduction in mortality was most notable in the patients requiring longer than 5 days in the ICU. Furthermore, intensive insulin therapy reduced episodes of septicemia by 46%, reduced duration of antibiotic therapy, and decreased the need for prolonged ventilatory support and renal replacement therapy.

Another treatment protocol that has been demonstrated to increase survival in patients with ARDS investigated the use of lower ventilatory tidal volumes compared to traditional tidal volumes.103 The majority of the patients enrolled in this multicenter, randomized trial developed ARDS secondary to pneumonia or sepsis. The trial compared traditional ventilation treatment, which involved an initial tidal volume of 12 mL/kg of predicted body weight, with ventilation with a lower tidal volume, which involved an initial tidal volume of 6 mL/kg of predicted body weight. The trial was stopped after the enrollment of 861 patients because mortality was lower in the group treated with lower tidal volumes than in the group treated with traditional tidal volumes (31.0% vs. 39.8%, P = .007), and the number of days without ventilator use during the first 28 days after randomization was greater in this group (mean ± SD, 12 ± 11 vs. 10 ± 11; P = .007). The investigators concluded that in patients with acute lung injury and ARDS, mechanical ventilation with a lower tidal volume than is traditionally used results in decreased mortality and increases the number of days without ventilator use. Additional strategies in ARDS management include higher levels of positive end expiratory pressure (PEEP), alveolar recruitment maneuvers, and prone positioning. Prone positioning has become part of many standardized protocols. This is supported by several studies, including a
meta-analysis demonstrating that prone positioning is associated with significantly reduced mortality from ARDS in the low tidal volume era.102

The use of corticosteroids in the treatment of sepsis and septic shock has been controversial for decades. The observation that severe sepsis often is associated with adrenal insufficiency or glucocorticoid receptor resistance has generated renewed interest in therapy for septic shock with corticosteroids. A single IV dose of 50 mg of hydrocortisone improved mean arterial blood pressure response relationships to norepinephrine and phenylephrine in patients with septic shock and was most notable in patients with relative adrenal insufficiency. A study evaluated therapy with hydrocortisone (50 mg IV every 6 hours) and fludrocortisone (50 µg orally once daily) vs. placebo for 1 week in patients with septic shock.103 As in earlier studies, the authors performed corticotropin tests on these patients to document and stratify patients by relative adrenal insufficiency. In this study, 7-day treatment with low doses of hydrocortisone and fludrocortisone significantly and safely lowered the risk of death in patients with septic shock and relative adrenal insufficiency. In an international, multicenter, randomized trial of corticosteroids in sepsis (CORTICUS study; 499 analyzable patients), steroids showed no benefit in intent to treat mortality or shock reversal.104 This study suggested that hydrocortisone therapy cannot be recommended as routine adjuvant therapy for septic shock. However, if SBP remains less than 90 mmHg despite appropriate fluid and vasopressor therapy, hydrocortisone at 200 mg/day for 7 days in four divided doses or by continuous infusion should be considered.

Additional adjunctive immune modulation strategies have been developed for the treatment of septic shock. These include the use of antiendotoxin antibodies, anticytokine antibodies, cytokine receptor antagonists, immune enhancers, a non–isoform-specific nitric oxide synthase inhibitor, and O2• radical scavengers. These compounds are each designed to alter some aspect of the host immune response to shock that is hypothesized to play a key role in its pathophysiology. However, most of these strategies have failed to demonstrate efficacy in human patients despite utility in well-controlled animal experiments. It is unclear whether the failure of these compounds is due to poorly designed clinical trials, inadequate understanding of the interactions of the complex host immune response to injury and infection, or animal models of shock that poorly represent the human disease.

Cardiogenic Shock
Cardiogenic shock is defined clinically as circulatory pump failure leading to diminished forward flow and subsequent tissue hypoxia, in the setting of adequate intravascular volume. Hemodynamic criteria include sustained hypotension (i.e., SBP <90 mmHg for at least 30 minutes), reduced cardiac index (<2.2 L/min per square meter), and elevated pulmonary arterial wedge pressure (>15 mmHg).105 Mortality rates for cardiogenic shock are 50% to 80%. Acute, extensive MI is the most common cause of cardiogenic shock; a smaller infarction in a patient with existing left ventricular dysfunction also may precipitate shock. Cardiogenic shock complicates 5% to 10% of acute MIs. Conversely, cardiogenic shock is the most common cause of death in patients hospitalized with acute MI. Although shock may develop early after MI, it typically is not found on admission. Seventy-five percent of patients who have cardiogenic shock complicating acute MIs develop signs of cardiogenic shock within 24 hours after onset of infarction (average 7 hours).

### Table 5-8

<table>
<thead>
<tr>
<th>Causes of cardiogenic shock</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute myocardial infarction</td>
</tr>
<tr>
<td>Pump failure</td>
</tr>
<tr>
<td>Mechanical complications</td>
</tr>
<tr>
<td>Acute mitral regurgitation</td>
</tr>
<tr>
<td>Acute ventricular septal defect</td>
</tr>
<tr>
<td>Free wall rupture</td>
</tr>
<tr>
<td>Pericardial tamponade</td>
</tr>
<tr>
<td>Arrhythmia</td>
</tr>
<tr>
<td>End-stage cardiomyopathy</td>
</tr>
<tr>
<td>Myocarditis</td>
</tr>
<tr>
<td>Severe myocardial contusion</td>
</tr>
<tr>
<td>Left ventricular outflow obstruction</td>
</tr>
<tr>
<td>Aortic stenosis</td>
</tr>
<tr>
<td>Hypertrophic obstructive cardiomyopathy</td>
</tr>
<tr>
<td>Obstruction to left ventricular filling</td>
</tr>
<tr>
<td>Mitral stenosis</td>
</tr>
<tr>
<td>Left atrial myxoma</td>
</tr>
<tr>
<td>Acute mitral regurgitation</td>
</tr>
<tr>
<td>Acute aortic insufficiency</td>
</tr>
<tr>
<td>Metabolic</td>
</tr>
<tr>
<td>Drug reactions</td>
</tr>
</tbody>
</table>
Increased sympathetic stimulation of the heart, either through direct neural input or from circulating catecholamines, increases heart rate, myocardial contraction, and myocardial O₂ consumption, which may not be relieved by increases in coronary artery blood flow in patients with fixed stenoses of the coronary arteries. Diminished cardiac output may also decrease coronary artery blood flow, resulting in a scenario of increased myocardial O₂ demand at a time when myocardial O₂ supply may be limited. Acute heart failure may also result in fluid accumulation in the pulmonary microcirculatory bed, decreasing myocardial O₂ delivery even further.

**Diagnosis.** Rapid identification of the patient with pump failure and institution of corrective action are essential in preventing the ongoing spiral of decreased cardiac output from injury causing increased myocardial O₂ needs that cannot be met, leading to progressive and unremitting cardiac dysfunction. In evaluation of possible cardiogenic shock, other causes of hypotension must be excluded, including hemorrhage, sepsis, pulmonary embolism, and aortic dissection. Signs of circulatory shock include hypotension, cool and mottled skin, depressed mental status, tachycardia, and diminished pulses. Cardiac exam may include dysrhythmia, precordial heave, or distal heart tones. Confirmation of a cardiac source for the shock requires electrocardiogram and urgent echocardiography. Other useful diagnostic tests include chest radiograph, arterial blood gases, electrolytes, complete blood count, and cardiac enzymes. Invasive cardiac monitoring, which generally is not necessary, can be useful to exclude right ventricular infarction, hypovolemia, and possible mechanical complications.

Making the diagnosis of cardiogenic shock involves the identification of cardiac dysfunction or acute heart failure in a susceptible patient. In the setting of blunt traumatic injury, hemorrhagic shock from intra-abdominal bleeding, intrathoracic bleeding, and bleeding from fractures must be excluded, before implicating cardiogenic shock from blunt cardiac injury. Relatively few patients with blunt cardiac injury will develop cardiac pump dysfunction. Those who do generally exhibit cardiogenic shock early in their evaluation. Therefore, establishing the diagnosis of blunt cardiac injury is secondary to excluding other etiologies for shock and establishing that cardiac dysfunction is present. Invasive hemodynamic monitoring with a pulmonary artery catheter may uncover evidence of diminished cardiac output and elevated pulmonary artery pressure.

**Treatment.** After ensuring that an adequate airway is present and ventilation is sufficient, attention should be focused on support of the circulation. Intubation and mechanical ventilation often are required, if only to decrease work of breathing and facilitate sedation of the patient. Rapidly excluding hypovolemia and establishing the presence of cardiac dysfunction are essential. Treatment of cardiac dysfunction includes maintenance of adequate oxygenation to ensure adequate myocardial O₂ delivery and judicious fluid administration to avoid fluid overload and development of cardiogenic pulmonary edema. Electrolyte abnormalities, commonly hypokalemia and hypomagnesemia, should be corrected. Pain is treated with IV morphine sulfate or fentanyl. Significant dysrhythmias and heart block must be treated with antiarrhythmic drugs, pacing, or cardioversion, if necessary. Early consultation with cardiology is essential in current management of cardiogenic shock, particularly in the setting of acute MI.¹⁰⁵

When profound cardiac dysfunction exists, inotropic support may be indicated to improve cardiac contractility and cardiac output. Dobutamine primarily stimulates cardiac β₁-receptors to increase cardiac output but may also vasodilate peripheral vascular beds, lower total peripheral resistance, and lower systemic blood pressure through effects on β₂-receptors. Ensuring adequate preload and intravascular volume is therefore essential prior to instituting therapy with dobutamine. Dopamine stimulates receptors (vasoconstriction), β₁-receptors (cardiac stimulation), and β₂-receptors (vasodilation), with its effects on β-receptors predominating at lower doses. Dopamine may be preferable to dobutamine in treatment of cardiac dysfunction in hypotensive patients. Tachycardia and increased peripheral resistance from dopamine infusion may worsen myocardial ischemia. Titration of both dopamine and dobutamine infusions may be required in some patients.

Epinephrine stimulates α- and β-receptors and may increase cardiac contractility and heart rate; however, it also may have intense peripheral vasoconstrictor effects that impair further cardiac performance. Catecholamine infusions must be carefully controlled to maximize coronary perfusion, while minimizing myocardial O₂ demand. Balancing the beneficial effects of impaired cardiac performance with the potential side effects of excessive reflex tachycardia and peripheral vasoconstriction requires serial assessment of tissue perfusion using indices such as capillary refill, character of peripheral pulses, adequacy of urine output, or improvement in laboratory parameters of resuscitation such as pH, base deficit, and lactate. Invasive monitoring generally is necessary in these unstable patients. The phosphodiesterase inhibitors amrinone and milrinone may be required on occasion in patients with resistant cardiogenic shock. These agents have long half-lives and induce thrombocytopenia and hypotension, and use is reserved for patients unresponsive to other treatment.

Patients whose cardiac dysfunction is refractory to inotropics may require mechanical circulatory support with an intra-aortic balloon pump.¹¹⁰ Intra-aortic balloon pumping increases cardiac output and improves coronary blood flow by reduction of systolic afterload and augmentation of diastolic perfusion pressure. Unlike vasopressor agents, these beneficial effects occur without an increase in myocardial O₂ demand. An intra-aortic balloon pump can be inserted at the bedside in the ICU via the femoral artery through either a cutdown or using the percutaneous approach. Aggressive circulatory support of patients with cardiac dysfunction from intrinsic cardiac disease has led to more widespread application of these devices and more familiarity with their operation by both physicians and critical care nurses.

Preservation of existing myocardium and preservation of cardiac function are priorities of therapy for patients who have suffered an acute MI. Ensuring adequate oxygenation and O₂ delivery, maintaining adequate preload with judicious volume restoration, minimizing sympathetic discharge through adequate relief of pain, and correcting electrolyte imbalances are all straightforward nonspecific maneuvers that may improve existing cardiac function or prevent future cardiac complications. Anticoagulation and aspirin are given for acute MI. Although thrombolytic therapy reduces mortality in patients with acute MI, its role in cardiogenic shock is less clear. Patients in cardiac failure from an acute MI may benefit from pharmacologic or mechanical circulatory support in a manner similar to that of patients with cardiac failure related to blunt cardiac
injury. Additional pharmacologic tools may include the use of β-blockers to control heart rate and myocardial O₂ consumption, nitrates to promote coronary blood flow through vasodilation, and ACE inhibitors to reduce ACE-mediated vasoconstrictive effects that increase myocardial workload and myocardial O₂ consumption.

Current guidelines of the American Heart Association recommend percutaneous transluminal coronary angiography for patients with cardiogenic shock, ST elevation, left bundle-branch block, and age less than 75 years. Early definition of coronary anatomy and revascularization is the pivotal step in treatment of patients with cardiogenic shock from acute MI. When feasible, percutaneous transluminal coronary angioplasty (generally with stent placement) is the treatment of choice. Coronary artery bypass grafting seems to be more appropriate for patients with multiple vessel disease or left main coronary artery disease.

### Obstructive Shock

Although obstructive shock can be caused by a number of different etiologies that result in mechanical obstruction of venous return (Table 5-9), in trauma patients this is most commonly due to the presence of tension pneumothorax. Cardiac tamponade occurs when sufficient fluid has accumulated in the pericardial sac to obstruct blood flow to the ventricles. The hemodynamic abnormalities in pericardial tamponade are due to elevation of intracardiac pressures with limitation of ventricular filling in diastole with resultant decrease in cardiac output. Acutely, the pericardium does not distend; thus small volumes of blood may produce cardiac tamponade. If the effusion accumulates slowly (e.g., in the setting of uremia, heart failure, or malignant effusion), the quantity of fluid producing cardiac tamponade may reach 2000 mL. The major determinant of the degree of hypotension is the pericardial pressure. With either cardiac tamponade or tension pneumothorax, reduced filling of the right side of the heart from either increased intrapleural pressure secondary to air accumulation (tension pneumothorax) or increased intrapericardial pressure precluding atrial filling secondary to blood accumulation (cardiac tamponade) results in decreased cardiac output associated with increased central venous pressure.

### Diagnosis and Treatment

The diagnosis of tension pneumothorax should be made on clinical examination. The classic findings include respiratory distress (in an awake patient), hypotension, diminished breath sounds over one hemithorax, hyper-resonance to percussion, jugular venous distention, and shift of mediastinal structures to the unaffected side with tracheal deviation. In most instances, empiric treatment with pleural decompression is indicated rather than delaying to wait for radiographic confirmation. When a chest tube cannot be immediately inserted, such as in the prehospital setting, the pleural space can be decompressed with a large caliber needle. Immediate return of air should be encountered with rapid resolution of hypotension. Unfortunately, not all of the clinical manifestations of tension pneumothorax may be evident on physical examination. Hyperresonance may be difficult to appreciate in a noisy resuscitation area. Jugular venous distention may be absent in a hypovolemic patient. Tracheal deviation is a late finding and often is not apparent on clinical examination. Practically, three findings are sufficient to make the diagnosis of tension pneumothorax: respiratory distress or hypotension, decreased lung sounds, and hypertympany to percussion. Chest X-ray findings that may be visualized include deviation of mediastinal structures, depression of the hemidiaphragm, and hypo-opacification with absent lung markings. As discussed previously, definitive treatment of a tension pneumothorax is immediate tube thoracostomy. The chest tube should be inserted rapidly, but carefully, and should be large enough to evacuate any blood that may be present in the pleural space. Most recommend placement is in the fourth intercostal space (nipple level) at the anterior axillary line.

Cardiac tamponade results from the accumulation of blood within the pericardial sac, usually from penetrating trauma or chronic medical conditions such as heart failure or uremia. Although precordial wounds are most likely to injure the heart and produce tamponade, any projectile or wounding agent that passes in proximity to the mediastinum can potentially produce tamponade. Blunt cardiac rupture, a rare event in trauma victims who survive long enough to reach the hospital, can produce refractory shock and tamponade in the multiply-injured patient. The manifestations of cardiac tamponade, such as total circulatory collapse and cardiac arrest, may be catastrophic, or they may be more subtle. A high index of suspicion is warranted to make a rapid diagnosis. Patients who present with circulatory arrest from cardiac tamponade require emergency pericardial decompression, usually through a left thoracotomy. The indications for this maneuver are discussed in Chapter 7. Cardiac tamponade also may be associated with dyspnea, orthopnea, cough, peripheral edema, chest pain, tachycardia, muffled heart tones, jugular venous distention, and elevated central venous pressure. Beck’s triad consists of hypotension, muffled heart tones, and neck vein distention. Unfortunately, absence of these clinical findings may not be sufficient to exclude cardiac injury and cardiac tamponade. Muffled heart tones may be difficult to appreciate in a busy trauma center, and jugular venous distention and central venous pressure may be diminished by coexistent bleeding. Therefore, patients at risk for cardiac tamponade whose hemodynamic status permits additional diagnostic tests frequently require additional diagnostic maneuvers to confirm cardiac injury or tamponade.

Invasive hemodynamic monitoring may support the diagnosis of cardiac tamponade if elevated central venous pressure, pulsus paradoxus (i.e., decreased systemic arterial pressure with inspiration), or elevated right atrial and right ventricular pressure by pulmonary artery catheter are present. These hemodynamic profiles suffer from lack of specificity, the duration of time required to obtain them in critically injured patients, and their inability to exclude cardiac injury in the absence of tamponade. Chest radiographs may provide information on the possible trajectory of a projectile, but rarely are diagnostic because

<table>
<thead>
<tr>
<th>Causes of obstructive shock</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pericardial tamponade</td>
<td></td>
</tr>
<tr>
<td>Pulmonary embolus</td>
<td></td>
</tr>
<tr>
<td>Tension pneumothorax</td>
<td></td>
</tr>
<tr>
<td>IVC obstruction</td>
<td></td>
</tr>
<tr>
<td>Deep venous thrombosis</td>
<td></td>
</tr>
<tr>
<td>Gravid uterus on IVC</td>
<td></td>
</tr>
<tr>
<td>Neoplasm</td>
<td></td>
</tr>
<tr>
<td>Increased intrathoracic pressure</td>
<td></td>
</tr>
<tr>
<td>Excess positive end-expiratory pressure</td>
<td></td>
</tr>
<tr>
<td>Neoplasm</td>
<td></td>
</tr>
</tbody>
</table>

IVC = inferior vena cava.
the acutely filled pericardium distends poorly. Echocardiography has become the preferred test for the diagnosis of cardiac tamponade. Good results in detecting pericardial fluid have been reported, but the yield in detecting pericardial fluid depends on the skill and experience of the ultrasonographer, body habitus of the patient, and absence of wounds that preclude visualization of the pericardium. Standard two-dimensional or transesophageal echocardiography are sensitive techniques to evaluate the pericardium for fluid, and are typically performed by examiners skilled at evaluating ventricular function, valvular abnormalities, and integrity of the proximal thoracic aorta. Unfortunately, these skilled examiners are rarely immediately available at all hours of the night, when many trauma patients present; therefore, waiting for this test may result in inordinate delays. In addition, although both ultrasound techniques may demonstrate the presence of fluid or characteristic findings of tamponade (large volume of fluid, right atrial collapse, poor distensibility of the right ventricle), they do not exclude cardiac injury per se. Pericardiocentesis to diagnose pericardial blood and potentially relieve tamponade may be used. Performing pericardiocentesis under ultrasound guidance has made the procedure safer and more reliable. An indwelling catheter may be placed for several days in patients with chronic pericardial effusions. Needle pericardiocentesis may not evacuate clotted blood and has the potential to produce cardiac injury, making it a poor alternative in busy trauma centers.

Diagnostic pericardial window represents the most direct method to determine the presence of blood within the pericardium. The procedure is best performed in the operating room under general anesthesia. It can be performed through either the subxiphoid or transdiaphragmatic approach. Adequate equipment and personnel to rapidly decompress the pericardium, explore the injury, and repair the heart should be present. Once the pericardium is opened and tamponade relieved, hemodynamics usually improve dramatically and formal pericardial exploration can ensue. Exposure of the heart can be achieved by extending the incision to a median sternotomy, performing a left anterior thoracotomy, or performing bilateral anterior thoracotomies (“clamshell”).

**Neurogenic Shock**

Neurogenic shock refers to diminished tissue perfusion as a result of loss of vasomotor tone to peripheral arterial beds. Loss of vasoconstrictor impulses results in increased vascular capacitance, decreased venous return, and decreased cardiac output. Neurogenic shock is usually secondary to spinal cord injuries from vertebral body fractures of the cervical or high thoracic region that disrupt sympathetic regulation of peripheral vascular tone (Table 5-10). Rarely, a spinal cord injury without bony fracture, such as an epidural hematoma impinging on the spinal cord, can produce neurogenic shock. Sympathetic input to the heart, which normally increases heart rate and cardiac contractility, and input to the adrenal medulla, which increases catecholamine release, may also be disrupted, preventing the typical reflex tachycardia that occurs with hypovolemia. Acute spinal cord injury results in activation of multiple secondary injury mechanisms: (a) vascular compromise to the spinal cord with loss of autoregulation, vasospasm, and thrombosis; (b) loss of cellular membrane integrity and impaired energy metabolism; and (c) neurotransmitter accumulation and release of free radicals. Importantly, hypotension contributes to the worsening of acute spinal cord injury as the result of further reduction in blood flow to the spinal cord. Management of acute spinal cord injury with attention to blood pressure control, oxygenation, and hemodynamics, essentially optimizing perfusion of an already ischemic spinal cord, seems to result in improved neurologic outcome. Patients with hypotension from spinal cord injury are best monitored in an ICU and carefully followed for evidence of cardiac or respiratory dysfunction.

**Diagnosis.** Acute spinal cord injury may result in bradycardia, hypotension, cardiac dysrhythmias, reduced cardiac output, and decreased peripheral vascular resistance. The severity of the spinal cord injury seems to correlate with the magnitude of cardiovascular dysfunction. Patients with complete motor injuries are over five times more likely to require vasopressors for neurogenic shock compared to those with incomplete lesions. The classic description of neurogenic shock consists of decreased blood pressure associated with bradycardia (absence of reflexive tachycardia due to disrupted sympathetic discharge), warm extremities (loss of peripheral vasoconstriction), motor and sensory deficits indicative of a spinal cord injury, and radiographic evidence of a vertebral column fracture. Patients with multisystem trauma that includes spinal cord injuries often have head injuries that may make identification of motor and sensory deficits difficult in the initial evaluation. Furthermore, associated injuries may occur that result in hypovolemia, further complicating the clinical presentation. In a subset of patients with spinal cord injuries from penetrating wounds, most of the patients with hypotension had blood loss as the etiology (74%) rather than neurogenic causes, and few (7%) had the classic findings of neurogenic shock. In the multiply injured patient, other causes of hypotension including hemorrhage, tension pneumothorax, and cardiogenic shock, must be sought and excluded.

**Treatment.** After the airway is secured and ventilation is adequate, fluid resuscitation and restoration of intravascular volume often will improve perfusion in neurogenic shock. Most patients with neurogenic shock will respond to restoration of intravascular volume alone, with satisfactory improvement in perfusion and resolution of hypotension. Administration of vasoconstrictors will improve peripheral vascular tone, decrease vascular capacitance, and increase venous return, but should only be considered once hypovolemia is excluded as the cause of the hypotension and the diagnosis of neurogenic shock is established. If the patient’s blood pressure has not responded to what is felt to be adequate volume resuscitation, dopamine may be used first. A pure α-agonist, such as phenylephrine, may be used primarily or in patients unresponsive to dopamine. Specific treatment for the hypotension is often of brief duration, as the need to administer vasoconstrictors typically lasts 24 to 48 hours. On the other hand, life-threatening cardiac dysrhythmias and hypotension may occur up to 14 days after spinal cord injury.

The duration of the need for vasopressor support for neurogenic shock may correlate with the overall prognosis or chances of improvement in neurologic function. Appropriate rapid restoration of blood pressure and circulatory perfusion may improve perfusion to the spinal cord, prevent progressive

<table>
<thead>
<tr>
<th>Table 5-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Causes of neurogenic shock</td>
</tr>
<tr>
<td>Spinal cord trauma</td>
</tr>
<tr>
<td>Spinal cord neoplasm</td>
</tr>
<tr>
<td>Spinal/epidural anesthetic</td>
</tr>
</tbody>
</table>
spinal cord ischemia, and minimize secondary cord injury. Restoration of normal blood pressure and adequate tissue perfusion should precede any operative attempts to stabilize the vertebral fracture.

ENDPOINTS IN RESUSCITATION

Shock is defined as inadequate perfusion to maintain normal organ function. With prolonged anaerobic metabolism, tissue acidosis and O₂ debt accumulate. Thus, the goal in the treatment of shock is restoration of adequate organ perfusion and tissue oxygenation. Resuscitation is complete when O₂ debt is repaid, tissue acidosis is corrected, and aerobic metabolism restored. Clinical confirmation of this endpoint remains a challenge.

Resuscitation of the patient in shock requires simultaneous evaluation and treatment; the etiology of the shock often is not initially apparent. Hemorrhagic shock, septic shock, and traumatic shock are the most common types of shock encountered on surgical services. To optimize outcome in bleeding patients, early control of the hemorrhage and adequate volume resuscitation, including both red blood cells and crystalloid solutions, are necessary. Expedient operative resuscitation is mandatory to limit the magnitude of activation of multiple mediator systems and to abort the microcirculatory changes, which may evolve insidiously into the cascade that ends in irreversible hemorrhagic shock. Attempts to stabilize an actively bleeding patient anywhere but in the operating room are inappropriate. Any intervention that delays the patient’s arrival in the operating room for control of hemorrhage increases mortality, thus the important concept of operating room resuscitation of the critically injured patient.

Recognition by care providers of the patient who is in the compensated phase of shock is equally important, but more difficult based on clinical criteria. Compensated shock exists when inadequate tissue perfusion persists despite normalization of blood pressure and heart rate. Even with normalization of blood pressure, heart rate, and urine output, 80% to 85% of trauma patients have inadequate tissue perfusion, as evidenced by increased lactate or decreased mixed venous O₂ saturation. Persistent, occult hypoperfusion is frequent in the ICU, with a resultant significant increase in infection rate and mortality in major trauma patients. Patients failing to reverse their lactic acidosis within 12 hours of admission (acidosis that was persistent despite normal heart rate, blood pressure, and urine output) developed an infection three times as often as those who normalized their lactate levels within 12 hours of admission. In addition, mortality was fourfold higher in patients who developed infections. Both injury severity score and occult hypotension (lactic acidosis) longer than 12 hours were independent predictors of infection. Thus, recognition of subclinical hypoperfusion requires information beyond vital signs and urinary output.

Endpoints in resuscitation can be divided into systemic or global parameters, tissue-specific parameters, and cellular parameters. Global endpoints include vital signs, cardiac output, pulmonary artery wedge pressure, O₂ delivery and consumption, lactate, and base deficit (Table 5-11).

**Table 5-11**

<table>
<thead>
<tr>
<th>Endpoints in resuscitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systemic/global</td>
</tr>
<tr>
<td>Lactate</td>
</tr>
<tr>
<td>Base deficit</td>
</tr>
<tr>
<td>Cardiac output</td>
</tr>
<tr>
<td>Oxygen delivery and consumption</td>
</tr>
<tr>
<td>Tissue specific</td>
</tr>
<tr>
<td>Gastric tonometry</td>
</tr>
<tr>
<td>Tissue pH, oxygen, carbon dioxide levels</td>
</tr>
<tr>
<td>Near infrared spectroscopy</td>
</tr>
<tr>
<td>Cellular</td>
</tr>
<tr>
<td>Membrane potential</td>
</tr>
<tr>
<td>Adenosine triphosphate</td>
</tr>
</tbody>
</table>

Assessment of Endpoints in Resuscitation

Inability to repay O₂ debt is a predictor of mortality and organ failure; the probability of death has been directly correlated to the calculated O₂ debt in hemorrhagic shock. Direct measurement of the O₂ debt in the resuscitation of patients is difficult. The easily obtainable parameters of arterial blood pressure, heart rate, urine output, central venous pressure, and pulmonary artery occlusion pressure are poor indicators of the adequacy of tissue perfusion. Therefore, surrogate parameters have been sought to estimate the O₂ debt; serum lactate and base deficit have been shown to correlate with O₂ debt.

**Lactate.** Lactate is generated by conversion of pyruvate to lactate by lactate dehydrogenase in the setting of insufficient O₂. Lactate is released into the circulation and is predominantly taken up and metabolized by the liver and kidneys. The liver accounts for approximately 50% and the kidney for about 30% of whole body lactate uptake. Elevated serum lactate is an indirect measure of the O₂ debt, and therefore an approximation of the magnitude and duration of the severity of shock. The admission lactate level, highest lactate level, and time interval to normalize the serum lactate are important prognostic indicators for survival. For example, in a study of 76 consecutive patients, 100% survival was observed among the patients with normalization of lactate within 24 hours, 78% survival when lactate normalized between 24 and 48 hours, and only 14% survival if it took longer than 48 hours to normalize the serum lactate. In contrast, individual variability of lactate may be too great to permit accurate prediction of outcome in any individual case. Base deficit and volume of blood transfusion required in the first 24 hours of resuscitation may be better predictors of mortality than the plasma lactate alone.

**Base Deficit.** Base deficit is the amount of base in millimoles that is required to titrate 1 L of whole blood to a pH of 7.40 with the sample fully saturated with O₂ at 37°C (98.6°F) and a partial pressure of CO₂ of 40 mmHg. It usually is measured by arterial blood gas analysis in clinical practice as it is readily and quickly available. The mortality of trauma patients can be stratified according to the magnitude of base deficit measured in the first 24 hours after admission. In a retrospective study of over 3000 trauma admissions, patients with a base deficit worse than 15 mmol/L had a mortality of 70%. Base deficit can be stratified into mild (3 to 5 mmol/L), moderate (6 to 14 mmol/L), and severe (15 mmol/L) categories, with a trend toward higher mortality with worsening base deficit in patients with trauma. Both the magnitude of the perfusion deficit as indicated by the base deficit and the time required to correct it are major factors determining outcome in shock.

Indeed, when elevated base deficit persists (or lactic acidosis) in the trauma patient, ongoing bleeding is often the...
etiology. Trauma patients admitted with a base deficit greater than 15 mmol/L required twice the volume of fluid infusion and six times more blood transfusion in the first 24 hours compared to patients with mild acidosis. Transfusion requirements increased as base deficit worsened and ICU and hospital lengths of stay increased. Mortality increased as base deficit worsened; the frequency of organ failure increased with greater base deficit.59 The probability of trauma patients developing ARDS has been reported to correlate with severity of admission base deficit and lowest base deficit within the first 24 hours postinjury.59 Persistently high base deficit is associated with abnormal \( O_2 \) utilization and higher mortality. Monitoring base deficit in the resuscitation of trauma patients assists in assessment of \( O_2 \) transport and efficacy of resuscitation.58

Factors that may compromise the utility of the base deficit in estimating \( O_2 \) debt are the administration of bicarbonate, hypothermia, hypcapnia (overventilation), heparin, ethanol, and ketoacidosis. However, the base deficit remains one of the most widely used estimates of \( O_2 \) debt for its clinical relevance, accuracy, and availability.

**Near Infrared Spectroscopy.** Near infrared (NIR) spectroscopy can measure tissue oxygenation and redox state of cytochrome \( a,a_3 \) on a continuous, noninvasive basis. The NIR probe emits multiple wavelengths of light in the NIR spectrum (650 to 1100 nm). Photons are then either absorbed by the tissue or reflected back to the probe. Maximal exercise in laboratory studies resulted in reduction of cytochrome \( a,a_3 \); this correlated with tissue lactate elevation. NIR spectroscopy can be used to compare tissue oxyhemoglobin levels (indicating tissue \( O_2 \) supply to cytochrome \( a,a_3 \), with mitochondrial \( O_2 \) consumption), thus demonstrating flow-independent mitochondrial oxidative dysfunction and the need for further resuscitation. Trauma patients with decoupled oxyhemoglobin and cytochrome \( a,a_3 \) have redox dysfunction and have been shown to have a higher incidence of organ failure (89% vs. 13%).121,122

**Tissue PH, Oxygen, and Carbon Dioxide Concentration.** Tissue probes with optical sensors have been used to measure tissue \( pH \) and partial pressure of \( O_2 \) and \( CO_2 \) in subcutaneous sites, muscle, and the bladder. These probes may use transcutaneous methodology with Clark electrodes or direct percutaneous probes.223,124 The percutaneous probes can be inserted through an 18-gauge catheter and hold promise as continuous monitors of tissue perfusion.

**Right Ventricular End-Diastolic Volume Index.** Right ventricular end-diastolic volume index (RVEDVI) seems to more accurately predict preload for cardiac index than does pulmonary artery wedge pressure.125 Chang and colleagues reported that 50% of trauma patients had persistent splanchnic ischemia that was reversed by increasing RVEDVI. RVEDVI is a parameter that seems to correlate with preload-related increases in cardiac output. More recently, these authors have described left ventricular power output as an endpoint (LVP >320 mmHg L/min per square meter), which is associated with improved clearance of base deficit and a lower rate of organ dysfunction following injury.126

**REFERENCES**

Entries highlighted in bright blue are key references.


64. Clarke JR, Trooskin SZ, Doshi PJ, Greenwald L, Mode CJ. Time to laparotomy for intra-abdominal bleeding from trauma.


101. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute


HISTORICAL BACKGROUND

Although treatment of infection has long been an integral part of the surgeon’s practice, the body of knowledge that led to the present field of surgical infectious disease was derived from the evolution of germ theory and antisepsis. Application of the latter to clinical practice, concurrent with the development of anesthesia, was pivotal in allowing surgeons to expand their repertoire to encompass complex procedures that previously were associated with extremely high rates of morbidity and mortality due to postoperative infections. However, until recently the occurrence of infection related to the surgical wound was the rule rather than the exception. In fact, the development of modalities to effectively prevent and treat infection has occurred only within the last several decades.

A number of observations by 19th century physicians and investigators were critical to our current understanding of the pathogenesis, prevention, and treatment of surgical infections. In 1846, Ignaz Semmelweis, a Magyar physician, took a post at the Allgemein Krankenhaus in Vienna. He noticed that the mortality rate from puerperal (“childbed”) fever was nearly three times higher in the teaching ward than in the ward where patients were delivered by midwives. He also made the observation that women who delivered prior to arrival on the teaching ward had a negligible mortality rate. When a colleague died from overwhelming infection resulting from a knife scratch received during an autopsy of a woman who had died of puerperal fever, Semmelweis observed that pathologic changes in his friend were identical to those of women dying from this postpartum disease. He hypothesized that puerperal fever was caused by putrid material carried on the examining fingers of medical students and physicians who cared for women dying of the disease, and who often went from the autopsy room to the wards. The low mortality rate in the midwives’ ward, Semmelweis realized, was because midwives did not participate in autopsies. Fired with the zeal of his revelation, he posted a notice on the door to the ward requiring all caregivers to rinse their hands thoroughly in chlorine water prior to entering the area. This simple intervention reduced the mortality rate from puerperal fever on the teaching ward to 1.5%, surpassing the record of the midwives. In 1861, he published his classic work on childbed fever based on records from his practice. Unfortunately, Semmelweis’ ideas were not well accepted by the authorities of the time.1 Increasingly frustrated by the indifference of the medical profession, he began writing open letters to well-known obstetricians in Europe and was committed to an asylum due to concerns that he was losing his mind. He died shortly thereafter. His achievements were only recognized after Pasteur’s description of the germ theory of disease.

Louis Pasteur performed a body of work during the latter part of the 19th century that provided the underpinnings of modern microbiology, at the time known as *germ theory*. His work in humans followed experiments identifying infectious agents in silkworms. He was able to elucidate the principle that contagious diseases are caused by specific microbes and that these microbes are foreign to the infected organism. Using this principle, he developed techniques of sterilization critical to oenology and identified several bacteria responsible for human illnesses, including *Staphylococcus* and *Streptococcus pneumoniae* (pneumococcus).

Joseph Lister, the son of a wine merchant, was appointed professor of surgery at the Glasgow Royal Infirmary in 1859. In his early practice, he noted that more than half of his patients undergoing amputation died because of postoperative infection. After hearing of Pasteur’s work, Lister experimented with the use of a solution of carbolic acid, which he knew was being used to treat sewage. He first reported his findings to the British Medical Association in 1867 using dressings saturated with carbolic acid on 12 patients with compound fractures; 10 recovered
without amputation, one survived with amputation, and one died of causes unrelated to the wound. In spite of initial resistance, his methods were quickly adopted throughout much of Europe.

From 1878 until 1880, Robert Koch was the district medical officer for Wollstein, an area in Prussia where anthrax was endemic. Performing experiments in his home, without the benefit of scientific equipment and academic contact, Koch developed techniques for culture of *Bacillus anthracis* and proved the ability of this organism to cause anthrax in healthy animals. He developed the following four postulates to identify the association of organisms with specific diseases: (a) the suspected pathogenic organism should be present in all cases of the disease and absent from healthy animals, (b) the suspected pathogen should be isolated from a diseased host and grown in a pure culture in vitro, (c) cells from a pure culture of the suspected organism should cause disease in a healthy animal, and (d) the organism should be reisolated from the newly diseased animal and shown to be the same as the original. He used these same techniques to identify the organisms responsible for cholera and tuberculosis. During the next century, Koch’s postulates, as they came to be called, became critical to the understanding of surgical infections.

The first intra-abdominal operation to treat infection via “source control” (i.e., surgical intervention to eliminate the source of infection) was appendectomy. This operation was pioneered by Charles McBurney at the New York College of Physicians and Surgeons, among others. McBurney’s classic report on early operative intervention for appendicitis was presented before the New York Surgical Society in 1889. Appendectomy for the treatment of appendicitis, previously an often fatal disease, was popularized after the 1902 coronation of King Edward VII of England was delayed due to his falling ill with appendicitis. Edward insisted on carrying out his schedule, despite worsening abdominal pain. Sir Frederick Treves, a prominent London surgeon, was among the consultants in attendance upon Edward. As the prince’s condition deteriorated, and as he continued to insist that he would go to Westminster Abbey to be crowned, Treves told him, “Then Sire, you will go as a corpse.” Edward relented, Treves drained a large periappendiceal abscess, and the king lived.

During the 20th century the development of effective antimicrobials added a new dimension to modern surgical practice. Sir Alexander Fleming, after serving in the British Army Medical Corps during World War I, continued his work on the natural antibacterial action of the blood and antiseptics. In 1928, while studying influenza virus, he noted a zone of inhibition around a mold colony (*Penicillium notatum*) that serendipitously grew on a plate of *Staphylococcus*, and he named the active substance penicillin. Penicillin, along with the sulfonamide antibiotics, were among the first of hundreds of potent antimicrobials that became a critical component of the armamentarium to prevent and treat aggressive, lethal surgical infections.

Concurrent with the development of antimicrobial agents were advances in the field of clinical microbiology. Many new microbes were identified, including numerous anaerobes. The autochthonous microflora of the skin, gastrointestinal tract, and other parts of the body that the surgeon encountered in the process of an operation were characterized in great detail. However, it remained unclear whether these organisms were commensals or pathogens. Subsequently, the initial clinical observations of surgeons such as Frank Meloney, William Altemeier, and others provided the key when they observed that aerobic and anaerobic host flora could synergize to cause serious soft tissue and severe intra-abdominal infection. Thus, the concepts that resident

---

**Key Points**

1. **Sepsis** is a life-threatening syndrome reflecting both an infection and the systemic host response to it. It has a broad variety of presentations and manifestations that hold in common some form of organ dysfunction. Outcomes in patients with sepsis are improved with an organized approach to therapy that addresses rapid resuscitation, antibiotics, and source control.

2. **Source control** is a key concept in the treatment of most surgically relevant infections. Infected or necrotic material must be drained or removed as part of the treatment plan in this setting. Delays in adequate source control are associated with worsened outcomes.

3. **Principles relevant to appropriate antibiotic prophylaxis for surgery:** (a) select an agent with activity against organisms commonly found at the site of surgery, (b) administer the initial dose of the antibiotic within 30 minutes prior to incision, (c) redose the antibiotic during long operations based upon the half-life of the agent to ensure adequate tissue levels, and (d) limit the antibiotic regimen to no more than 24 hours after surgery for routine prophylaxis.

4. **When using antimicrobial agents for therapy of serious infection,** several principles should be followed: (a) identify likely sources of infection, (b) select an agent (or agents) that will have efficacy against likely organisms for these sources, (c) begin therapy rapidly with broad coverage, as inadequate or delayed antibiotic therapy results in increased mortality, (d) when possible, obtain cultures early and use results to refine therapy, (e) if no infection is identified after 3 days, strongly consider discontinuation of antibiotics, based upon the patient’s clinical course, and (f) discontinue antibiotics after an appropriate course of therapy.

5. **The incidence of surgical site infections can be reduced by appropriate patient preparation, timely perioperative antibiotic administration, maintenance of perioperative normothermia and normoglycemia, and appropriate wound management.**

6. **The keys to good outcomes in patients with necrotizing soft tissue infection are early recognition and appropriate debridement of infected tissue with repeated debridement until no further signs of infection are present.**

7. **Transmission of HIV and other infections spread by blood and body fluids from patient to healthcare worker can be minimized by practicing universal precautions, which include routine use of barriers when anticipating contact with blood or body fluids, washing of hands and other skin surfaces immediately after contact with blood or body fluids, and careful handling and disposal of sharp instruments during and after use.**
microbes were nonpathogenic until they entered a sterile body cavity at the time of surgery, and that many, if not most, surgical infections were polymicrobial in nature, became critical ideas.\(^8,9\) These tenets became firmly established after microbiology laboratories demonstrated the invariable presence of aerobes and anaerobes in peritoneal cultures obtained at the time of surgery for intra-abdominal infection due to perforated viscus or gangrenous appendicitis. Clinical trials provided ample evidence that optimal therapy for these infections required effective source control and the administration of antimicrobial agents directed against both types of pathogens.

William Osler made an observation in 1904 in his treatise *The Evolution of Modern Medicine* that was to have profound implications for the future of treatment of infection: “Except on few occasions, the patient appears to die from the body’s response to infection rather than from it.”\(^10\) The discovery of cytokines began to allow insight into the human organism’s response to infection, and led to an explosion in our understanding of the host inflammatory response. Expanding knowledge of the multiple pathways activated during the response to invasion by infectious organisms has permitted the design of new therapies targeted at modifying the inflammatory response to infection, which seems to cause much of the organ dysfunction and failure. Preventing and treating this process of multiple organ failure during infection is one of the major challenges of modern critical care and surgical infectious disease.

**PATHOGENESIS OF INFECTION**

**Host Defenses**

The mammalian host possesses several layers of endogenous defense mechanisms that serve to prevent microbial invasion, limit proliferation of microbes within the host, and contain or eradicate invading microbes. These defenses are integrated and redundant so that the various components function as a complex, highly regulated system that is extremely effective in coping with microbial invaders. They include site-specific defenses that function at the tissue level, as well as components that freely circulate throughout the body in both blood and lymph. Systemic host defenses invariably are recruited to a site of infection, a process that begins immediately upon introduction of microbes into a sterile area of the body. Perturbation of one or more components of these defenses (e.g., via immunosuppressants, foreign body, chronic illness, or burns) may have substantial negative impact on resistance to infection.

Entry of microbes into the mammalian host is precluded by a number of barriers that possess either an epithelial (integument) or mucosal (respiratory, gut, and urogenital) surface. Barrier function, however, is not solely limited to physical characteristics. Host barrier cells may secrete substances that limit microbial proliferation or prevent invasion. Also, resident or commensal microbes adhere to the physical surface and to each other may preclude invasion, particularly of virulent organisms; this is termed *colonization resistance*.\(^11\)

The most extensive physical barrier is the integument or skin. In addition to the physical barrier posed by the epithelial surface, the skin harbors its own resident microflora that may block the attachment and invasion of noncommensal microbes. Microbes also are held in check by chemicals secreted by sebaceous glands and by the constant shedding of epithelial cells. The endogenous microflora of the integument primarily comprises gram-positive aerobic microbes belonging to the genera *Staphylococcus* and *Streptococcus*, as well as *Corynebacterium* and *Propionibacterium* species. These organisms plus *Enterococcus faecalis* and *faecium*, *Escherichia coli* and other *Enterobacteriaceae*, and yeast such as *Candida albicans* can be isolated from the inframountibial regions of the body. Diseases of the skin (e.g., eczema and dermatitis) are associated with overgrowth of skin commensal organisms, and barrier breaches invariably lead to the introduction of these microbes.

The respiratory tract possesses several host defense mechanisms that facilitate the maintenance of sterility in the distal bronchi and alveoli. In the upper respiratory tract, respiratory mucus traps larger particles, including microbes. This mucus is then passed into the upper airways and oropharynx by ciliated epithelial cells, where the mucus is cleared via coughing. Smaller particles arriving in the lower respiratory tract are cleared via phagocytosis by pulmonary alveolar macrophages. Any process that diminishes these host defenses can lead to development of bronchitis or pneumonia.

The urogenital, biliary, pancreatic ductal, and distal respiratory tracts do not possess resident microflora in healthy individuals, although microbes may be present if these barriers are affected by disease (e.g., malignancy, inflammation, calculi, or foreign body), or if microorganisms are introduced from an external source (e.g., urinary catheter or pulmonary aspiration). In contrast, significant numbers of microbes are encountered in many portions of the gastrointestinal tract, with vast numbers being found within the oropharynx and distal colon or rectum, although the specific organisms differ.

One would suppose that the entire gastrointestinal tract would be populated via those microbes found in the oropharynx, but this is not the case.\(^11\) This is because after ingestion these organisms routinely are killed in the highly acidic, low-motility environment of the stomach during the initial phases of digestion. Thus, only small numbers of microbes populate the gastric mucosa (∼10⁸ to 10⁹ colony-forming units [CFU]/mL). This population expands in the presence of drugs or disease states that diminish gastric acidity. Microbes that are not destroyed within the stomach enter the small intestine, in which a certain amount of microbial proliferation takes place, such that approximately 10⁶ to 10⁸ CFU/mL are present in the terminal ileum. The relatively low-oxygen, static environment of the colon is accompanied by the exponential growth of microbes that comprise the most extensive host endogenous microflora. Anaerobic microbes outnumber aerobic species approximately 100:1 in the distal colon, and approximately 10¹ⁱ to 10¹² CFU/g are present in feces. Large numbers of facultative and strict anaerobes (*Bacteroides fragilis, distasonis, and thetaiotaomicron, Bifidobacterium, Clostridium, Eubacterium, Fusobacterium, Lactobacillus, and Peptostreptococcus* species) as well as several orders of magnitude fewer aerobic microbes (*E. coli* and other *Enterobacteriaceae, E. faecalis and faecium, C. albicans* and other *Candida* spp.) are present. Intriguingly, although colonization resistance on the part of this extensive, well-characterized host microflora effectively prevents invasion of enteric pathogens such as *Salmonella*, *Shigella*, *Vibrio*, and other enteropathogenic bacterial species, these same organisms provide the initial inoculum for infection should perforation of the gastrointestinal tract occur. It is of great interest that only some of these microbial species predominate in established intra-abdominal infections.

Once microbes enter a sterile body compartment (e.g., the pleural or peritoneal cavity) or tissue, additional host defenses act to limit and/or eliminate these pathogens. Initially, several
primitive and relatively nonspecific host defenses act to contain the nidus of infection, which may include microbes as well as debris, devitalized tissue, and foreign bodies, depending on the nature of the injury. These defenses include the physical barrier of the tissue itself, as well as the capacity of proteins such as lactoferrin and transferrin to sequester the critical microbial growth factor iron, thereby limiting microbial growth. In addition, fibrinogen within the inflammatory fluid has the ability to trap large numbers of microbes during the process in which it polymerizes into fibrin. Within the peritoneal cavity, unique host defenses exist, including a diaphragmatic pumping mechanism whereby particles—including microbes—within peritoneal fluid are expelled from the abdominal cavity via specialized structures (stomata) on the undersurface of the diaphragm that lead to thoracic lymphatic channels. Concurrently, containment by the omentum and intestinal ileus serve to wall off infections. However, the latter processes and fibrin trapping have a high likelihood of contributing to the formation of an intra-abdominal abscess.

Microbes also immediately encounter a series of host defense mechanisms that reside within the vast majority of tissues of the body. These include resident macrophages and low levels of complement (C) proteins and immunoglobulins (e.g., antibodies). The response in macrophages is initiated by genome-encoded pattern recognition receptors that respond to invading microbes. With exposure to a foreign organism, these receptors recognize microbial pathogen-associated molecular patterns (PAMPs) and endogenous danger-associated molecular patterns (DAMPs). Toll-like receptors (TLRs) are a well-defined example of a PAMP that plays an important role in pathogen signaling. Resident macrophages secrete a wide array of substances in response to the aforementioned processes, some of which appear to regulate the cellular components of the host defense response. This results in recruitment and proliferation of inflammatory cells. Macrophage cytokine synthesis is upregulated. Secretion of tumor necrosis factor-alpha (TNF-α), of interleukins (IL)-1β, 6, and 8; and of gamma interferon (IFN-γ) occurs within the tissue milieu, and depending on the magnitude of the host defense response, the systemic circulation. Concurrently, a counterregulatory response is initiated consisting of binding protein (TNF-BP), cytokine receptor antagonists (e.g., IL-1ra), and anti-inflammatory cytokines (IL-4 and IL-10).

The interaction of microbes with these first-line host defenses leads to microbial opsonization (C1q, C3bi, and IgFc), phagocytosis, and both extracellular (CSb6-9 membrane attack complex) and intracellular microbial destruction (via cellular ingestion into phagocytic vacuoles). Concurrently, the classical and alternate complement pathways are activated both via direct contact with and via IgM and IgG binding to microbes, leading to the release of a number of different biologically active complement protein fragments (C3a, C4a, C5a), acting to markedly enhance vascular permeability. Bacterial cell wall components and a variety of enzymes expelled from leukocyte phagocytic vacuoles during microbial phagocytosis and killing act in this capacity as well.

Simultaneously, the release of substances to which polymorphonuclear leukocytes (PMNs) in the bloodstream are attracted takes place. These consist of C5a, microbial cell wall peptides containing N-formyl-methionine, and macrophage secretion of cytokines such as IL-8. This process of host defense recruitment leads to further influx of inflammatory fluid into the area of incipient infection and is accompanied by diapedesis of large numbers of PMNs, a process that begins within several minutes and may peak within hours or days. The magnitude of the response and eventual outcome is generally related to several factors: (a) the initial number of microbes, (b) the rate of microbial proliferation in relation to containment and killing by host defenses, (c) microbial virulence, and (d) the potency of host defenses. In regard to the latter, drugs or disease states that diminish any or multiple components of host defenses are associated with higher rates and potentially more grave infections.

**Definitions**

Several possible outcomes can occur subsequent to microbial invasion and the interaction of microbes with resident and recruited host defenses: (a) eradication; (b) containment, often leading to the presence of purulence, the hallmark of chronic infections (e.g., a furuncle in the skin and soft tissue or abscess within the parenchyma of an organ or potential space); (c) locoregional infection (cellulitis, lymphangitis, and aggressive soft tissue infection) with or without distant spread of infection (metastatic abscess); or (d) systemic infection (bacteremia or fungemia). Obviously, the latter represents the failure of resident and recruited host defenses at the local level, and is associated with significant morbidity and mortality. Disease progression commonly occurs such that serious locoregional infection is associated with concurrent systemic infection. A chronic abscess also may intermittently drain and/or be associated with bacteremia.

*Infection* is defined by the presence of microorganisms in host tissue or the bloodstream. The classic findings of *ruber, calor, and dolor* in areas such as the skin or subcutaneous tissue are common at the site of infection. Most infections in normal individuals with intact host defenses are associated with these local manifestations, plus systemic manifestations such as elevated temperature, elevated white blood cell (WBC) count, tachycardia, or tachypnea. The systemic manifestations noted previously comprise what has been termed the *systemic inflammatory response syndrome* (SIRS). SIRS reflects a proinflammatory state in response to a variety of disease processes, including infection, pancreatitis, polytrauma, malignancy, and burns. There are a variety of systemic manifestations of infection, with the classic factors of fever, tachycardia, and tachypnea broadened to include a variety of other variables (Table 6-1).

The definition of *sepsis* is evolving. Earlier models described sepsis as SIRS caused by infection. This was based upon the idea that sepsis is mediated by the production of a cascade of proinflammatory mediators produced in response to exposure to microbial products. These products include lipopolysaccharide (endotoxin, LPS) derived from gram-negative organisms; peptidoglycans and teichoic acids from gram-positive organisms; many different microbial cell wall components, such as mannan from yeast and fungi; and many others.

There are several issues, however, with basing a sepsis diagnosis on the presence of SIRS. One problem is that it is insufficiently specific. Patients can exhibit SIRS criteria without the presence of the more whole-body dysregulation consistent with sepsis, and conversely can suffer from sepsis without meeting SIRS criteria. Patients with SIRS do not necessarily progress to sepsis and do not necessarily have worsened outcomes because of the SIRS diagnosis; in other words, SIRS is not inherently life-threatening. Another issue is that the SIRS criteria can vary and are inconsistently applied. Numerous definitions exist, specifying differing physiologic and laboratory criteria for the
diagnosis. This creates difficulty in clinical, epidemiological, and research settings. Further, sepsis is not a purely inflammatory phenomenon, as both pro- and anti-inflammatory cascades have been shown to be activated in septic patients. Basing a diagnosis upon inflammatory markers alone disregards non-inflammatory organ dysfunction, which may not manifest as SIRS but can contribute to mortality. A final concern is that defining sepsis using SIRS criteria implies that SIRS, sepsis, severe sepsis, and septic shock exist upon a continuum, and while SIRS and sepsis have common features, the former does not necessarily lead to the latter. This being said, SIRS criteria have utility in that they point toward an organism experiencing physiological stress. The presence of SIRS warrants further investigation by the clinician.16

An international consensus panel proposed new definitions of sepsis and septic shock in 2016. What is known as the Sepsis-3 model defines sepsis as life-threatening organ dysfunction caused by a dysregulated host response to infection. Organ dysfunction is quantified by an increase of ≥2 points on the Sequential Organ Failure Assessment (SOFA). The SOFA score looks at PaO₂/FiO₂ ratio, bilirubin, platelet count, mean arterial pressure (MAP), Glasgow Coma Scale (GCS) score, creatinine level, and urine output (Table 6-2). An increase in SOFA score of 2 or more is correlated with a 10% in-hospital mortality risk, which is suggestive of the life-threatening nature of sepsis. An abbreviated version of the scoring system, the quick SOFA (qSOFA) is recommended as a screening and monitoring tool for patients with suspected sepsis. The qSOFA suggests potentially life-threatening sepsis when at least two of the following parameters are met: altered mental status, systolic blood pressure of 100 mmHg or less, and respiratory rate greater than 22 breaths/minute. The qSOFA can readily identify patients at risk of poor outcome from sepsis without reliance upon laboratory or imaging data.16

Under the older nomenclature, severe sepsis was characterized as sepsis combined with the presence of new-onset organ failure. The Sepsis-3 definitions consider the term “severe sepsis” to be redundant, as by this definition all sepsis involves organ dysfunction. Under the Sepsis-3 guidelines, septic shock is a subset of sepsis in which circulatory and cellular metabolic derangements are profound enough to significantly increase the risk of death. Sepsis is the most common cause of death in non-coronary critical care units and the 11th most common cause of death overall in the United States, with a mortality rate of 10.3 cases per 100,000 population in 2010.17 Septic shock is the most severe manifestation of infection, with an attendant mortality rate in excess of 40%. It can be identified by persistent arterial hypotension requiring vasopressors to maintain mean arterial pressure (MAP) ≥65, and by serum lactate >2 mmol/L (18 mg/dL) despite adequate volume resuscitation.16,18,10

### Microbiology of Infectious Agents

A partial list of common pathogens that cause infections in surgical patients is provided in Table 6-3.

**Bacteria**

Bacteria are responsible for the majority of surgical infections. Specific species are identified using Gram stain and growth characteristics on specific media. The Gram stain is an important evaluation that allows rapid classification of bacteria by color. This color is related to the staining characteristics of the bacterial cell wall: gram-positive bacteria stain blue and gram-negative bacteria stain red. Bacteria are classified based upon a number of additional characteristics, including morphology (cocci and bacilli), the pattern of division (single organisms, groups of organisms in pairs [diplococci], clusters [staphylococci], and chains [streptococci]), and the presence and location of spores.

Gram-positive bacteria that frequently cause infections in surgical patients include aerobic skin commensals (Staphylococcus aureus and Enterococcus faecalis and faecium) and staphylococci, and streptococci). Aerobic skin commensals cause a large percentage of surgical site infections (SSIs), either alone or in conjunction with other pathogens; enterococci can cause nosocomial infections (urinary tract infections [UTIs] and bacteremia) in immunocompromised or chronically ill patients, but are of relatively low virulence in healthy individuals.

There are many pathogenic gram-negative bacterial species that are capable of causing infection in surgical patients. Most gram-negative organisms of interest to the surgeon are bacilli belonging to the family Enterobacteriaceae, including Escherichia coli, Klebsiella pneumoniae, Serratia marcescens, and Enterobacter, Citrobacter, and Acinetobacter species. Other gram-negative bacilli of note include Pseudomonas, including P aeruginosa and fluorescens, and Stenotrophomonas species.
Anaerobic organisms divide poorly or are unable to grow in air, as most do not possess the enzyme catalase, which allows for metabolism of reactive oxygen species. Anaerobes are the predominant indigenous flora in many areas of the human body, with the particular species being dependent on the site. For example, Propionibacterium acnes and other species are a major component of the skin microflora and cause the infectious manifestation of acne. As noted previously, large numbers of anaerobes contribute to the microflora of the oropharynx and colon.

Infection due to Mycobacterium tuberculosis was once one of the most common causes of death in Europe, causing one in four deaths in the 17th and 18th centuries. In the 19th and 20th centuries, thoracic surgical intervention was often required for severe pulmonary disease, now an increasingly uncommon occurrence in developed countries. This organism and other related organisms (M avium-intracellulare and M leprae) are known as acid-fast bacilli. Other acid-fast bacilli include Nocardia. These organisms typically are slow growing, sometimes necessitating observation in culture for weeks to months prior to final identification, although deoxyribonucleic acid (DNA)-based analysis is increasingly available to provide a means for preliminary, rapid detection.

**Fungi**

Fungi are typically identified by use of special stains (e.g., potassium hydroxide, India ink, methenamine silver, or Giemsa). Initial identification is assisted by observation of the form of branching and septation in stained specimens or in culture.

**Table 6-2**

<table>
<thead>
<tr>
<th>Table 6-2</th>
<th>Sequential Organ Failure Assessment score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SYSTEM</strong></td>
<td><strong>SCORE</strong></td>
</tr>
<tr>
<td><strong>Respiratory</strong></td>
<td></td>
</tr>
<tr>
<td>PaO2/FiO2, mmHg (kPa)</td>
<td>0</td>
</tr>
<tr>
<td>≥400 (53.3)</td>
<td></td>
</tr>
<tr>
<td>&lt;400 (53.3)</td>
<td></td>
</tr>
<tr>
<td>&lt;300 (40)</td>
<td></td>
</tr>
<tr>
<td>&lt;200 (26.7) with respiratory support</td>
<td></td>
</tr>
<tr>
<td>&lt;100 (13.3) with respiratory support</td>
<td></td>
</tr>
<tr>
<td><strong>Coagulation</strong></td>
<td></td>
</tr>
<tr>
<td>Platelets, ×10^9/μL</td>
<td>≥150</td>
</tr>
<tr>
<td><strong>Hepatic</strong></td>
<td></td>
</tr>
<tr>
<td>Bilirubin, mg/dL (μmol/L)</td>
<td>&lt;1.2 (20)</td>
</tr>
<tr>
<td><strong>Cardiovascular</strong></td>
<td></td>
</tr>
<tr>
<td>MAP ≥70 mmHg</td>
<td>MAP &lt;70 mmHg</td>
</tr>
<tr>
<td><strong>CNS</strong></td>
<td></td>
</tr>
<tr>
<td>GCS score</td>
<td>15</td>
</tr>
<tr>
<td><strong>Renal</strong></td>
<td></td>
</tr>
<tr>
<td>Creatinine, mg/dL (μmol/L)</td>
<td>&lt;1.2 (110)</td>
</tr>
<tr>
<td><strong>Urinary output, mL/24 hours</strong></td>
<td></td>
</tr>
<tr>
<td>&lt;500</td>
<td>&lt;200</td>
</tr>
</tbody>
</table>

MAP = mean arterial pressure; PaO2 = partial pressure of oxygen; FiO2 = fraction of inspired oxygen; CNS = central nervous system; GCS = Glasgow Coma Scale

Catecholamine doses in μg/kg/minute


**Viruses**

Due to their small size and necessity for growth within cells, viruses are difficult to culture, requiring a longer time than is typically optimal for clinical decision making. Previously, viral infection was identified by indirect means (i.e., the host antibody response); modern techniques identify the presence of viral DNA or ribonucleic acid (RNA) using methods such as polymerase chain reaction. Similar to many fungal infections, most clinically relevant viral infections in surgical patients occur in the immunocompromised host, particularly those receiving immunosuppression to prevent rejection of a solid organ allograft. Relevant viruses include adenoviruses, cytomegalovirus, Epstein-Barr virus, herpes simplex virus, and varicella-zoster virus. Surgeons must be aware of the manifestations of hepatitis B and C viruses, as well as human immunodeficiency
virus infections, including their capacity to be transmitted to healthcare workers (see “General Principles”). Prophylactic and therapeutic use of antiviral agents is discussed elsewhere in this textbook.

**PREVENTION AND TREATMENT OF SURGICAL INFECTIONS**

**General Principles**

Maneuvers to diminish the presence of exogenous (surgeon and operating room environment) and endogenous (patient) microbes are termed **prophylaxis** and consist of a variety of mechanical and chemical modalities. The Centers for Disease Control and Prevention (CDC) publishes updated, evidence-based guidelines on best practices for prevention of surgical site infections. Important principles in prophylaxis can be grouped into factors pertaining to skin preparation, antimicrobial therapy, and patient physiological management.

**Patient skin preparation** should begin the night before a planned surgical procedure with a full body bath or shower using soap or an antiseptic agent. Hair removal from an operative site should be performed in the operating room with clippers rather than with a razor, to avoid creating nicks in the skin that could foster bacterial growth. Prior to incision, the skin should be cleansed with an alcohol-based antiseptic agent. There is no clear evidence that use of antimicrobial-containing fluids for either irrigation or soaking prosthetic materials is beneficial in preventing infections. Preoperative antimicrobial therapy should be administered when appropriate, based on clinical guidelines, and occur within a time frame that allows bactericidal concentration of the agent in tissues before the incision is made.

**Physiological management of the intraoperative patient** includes maintenance of euglycemia (serum glucose <200 mg/dL) and normothermia, and optimization of tissue oxygenation.

**Source Control**

The primary precept of surgical infectious disease therapy consists of drainage of all purulent material, debridement of all infected, devitalized tissue and debris, and/or removal of foreign bodies at the site of infection, plus remediation of the underlying cause of infection. This is termed **source control**. A discrete, walled-off purulent fluid collection (i.e., an abscess)
requires drainage, either surgically or via percutaneous drain insertion. An ongoing source of contamination (e.g., bowel perforation) or the presence of an aggressive, rapidly spreading infection (e.g., necrotizing soft tissue infection) invariably requires expedient, aggressive operative intervention, both to remove contaminated material and infected tissue (e.g., radical debridement or amputation) and to remove the initial cause of infection (e.g., bowel resection). Delay in operative intervention, whether due to misdiagnosis or the need for additional diagnostic studies, is associated with increased morbidity and occasional mortality. Other treatment modalities such as antimicrobial agents, albeit critical, are of secondary importance to effective surgery with regard to treatment of surgical infections. Rarely, if ever, can an aggressive surgical infection be cured only by the administration of antibiotics, and never in the face of an ongoing source of contamination.22

### Appropriate Use of Antimicrobial Agents

A classification of antimicrobial agents, mechanisms of action, and spectrums of activity is shown in Table 6-5. As discussed previously, prophylaxis consists of the administration of an antimicrobial agent or agents prior to initiation of certain specific types of surgical procedures in order to reduce the number of microbes that enter the tissue or body cavity. Agents are selected according to their activity against microbes likely to be present at the surgical site, based on knowledge of host microflora. For example, patients undergoing elective colorectal surgery should receive antimicrobial prophylaxis directed against skin flora, gram-negative aerobes, and anaerobic bacteria. There are a wide variety of agents that meet these criteria with recently published guidelines.23

By definition, prophylaxis is limited to the time prior to and during the operative procedure; in the vast majority of cases only a single dose of antibiotic is required, and only for certain types of procedures (see “Surgical Site Infections”). However, patients who undergo complex, prolonged procedures in which the duration of the operation exceeds the serum drug half-life should receive an additional dose or doses of the antimicrobial agent.23 There is no evidence that administration of postoperative doses of an antimicrobial agent provides additional benefit, and this practice should be discouraged, as it is costly and is associated with increased rates of microbial drug resistance. Guidelines for prophylaxis are provided in Table 6-6.

Empiric therapy is the use of antimicrobial agents when the risk of a surgical infection is high, based on the underlying disease process (e.g., ruptured appendicitis), or when significant contamination during surgery has occurred (e.g., inadequate bowel preparation or considerable spillage of colon contents). Obviously, prophylaxis merges into empiric therapy in situations in which the risk of infection increases markedly because of intraoperative findings. Empiric therapy also is often employed in critically ill patients in whom a potential site of infection has been identified and severe sepsis or septic shock occurs. Empiric therapy should be limited to a short course of treatment (3 to 5 days) and should be curtailed as soon as possible based on microbiologic data (i.e., absence of positive cultures) coupled with improvements in the clinical course of the patient.

Empiric therapy can merge into therapy of established infection in some patients. However, among surgical patients, the manner in which therapy is employed, particularly in relation to the use of microbiologic data (culture and antibiotic sensitivity patterns), differs depending on whether the infection is monomicrobial or polymicrobial. Monomicrobial infections frequently are nosocomial infections occurring in postoperative patients, such as UTIs, pneumonia, or bacteremia. Evidence of systemic inflammatory response syndrome (fever, tachycardia, tachypnea, or elevated leukocyte count) in such individuals, coupled with evidence of local infection (e.g., an infiltrate on chest roentgenogram plus a positive Gram stain in bronchoalveolar lavage samples) should lead the surgeon to initiate empiric antibiotic therapy. An appropriate approach to antimicrobial treatment involves de-escalation therapy, where initial antimicrobial selection is broad, with a narrowing of agents based on patient response and culture results. Initial drug selection must be based on initial evidence (gram-positive vs. gram-negative microbes, yeast), coupled with institutional and unit-specific drug sensitivity patterns. It is important to ensure that antimicrobial coverage chosen is adequate, since delay in appropriate antibiotic treatment has been shown to be associated with significant increases in mortality. A critical component of this approach is appropriate collection of culture specimens to allow for thorough analysis, since within 48 to 72 hours culture and sensitivity reports will allow refinement of the antibiotic regimen to select the most efficacious agent.

Although the primary therapeutic modality to treat polymicrobial surgical infections is source control, antimicrobial agents play an important role. Culture results are of lesser importance in managing these types of infections, as it has been repeatedly demonstrated that only a limited cadre of microbes predominate in the established infection, selected from a large number present at the time of initial contamination. Invariably it is difficult to identify all microbes that comprise the initial polymicrobial inoculum. For this reason, the antibiotic regimen should not be modified solely on the basis of culture information, as it is less important than the clinical course of the patient. As long as appropriately broad-spectrum coverage for aerobic and anaerobic microbes is provided, a worsening of the patient’s clinical course should direct the surgeon to investigate whether effective source control has been achieved.24 Duration of antibiotic administration should be decided at the time the drug regimen is prescribed. As mentioned previously, prophylaxis is limited to a single dose administered immediately prior to creating the incision. Empiric therapy should be limited to 3 to 5 days or less and should be curtailed if the presence of a local site or systemic infection is not revealed.25 In fact, prolonged use of empirical antibiotic therapy in culture-negative critically ill patients is associated with increased mortality, highlighting the need to discontinue therapy when there is no proven evidence of infection.26

Therapy for monomicrobial infections follows standard guidelines: 3 to 5 days for UTIs, 7 to 8 days for pneumonia, and 7 to 14 days for bacteremia. Longer courses of therapy in this setting do not result in improved care and are associated with increased risk of superinfection by resistant organisms.27-29 There is some evidence that measuring and monitoring serum procalcitonin trends in the setting of infection allows earlier cessation of antibiotics without decrement in the rate of clinical cure.30 Antibiotic therapy for osteomyelitis, endocarditis, or prosthetic infections in which it is hazardous to remove the device consists of prolonged courses of treatment for 6 to 12 weeks. The specific agents are selected based on analysis of the degree to which the organism is killed in vitro using the minimum inhibitory concentration (MIC) of a standard pure
<table>
<thead>
<tr>
<th>ANTIBIOTIC CLASS, GENERIC NAME</th>
<th>TRADE NAME</th>
<th>MECHANISM OF ACTION</th>
<th>S Pyogenes</th>
<th>MSSA</th>
<th>MRSA</th>
<th>S epidermidis</th>
<th>Enterococcus</th>
<th>VRE</th>
<th>E coli</th>
<th>P aeruginosa</th>
<th>ANAEROBES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Penicillins</td>
<td></td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+/-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Nallpen, Unipen</td>
<td></td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>+/-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Pipracil</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>+/-</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>+/-</td>
<td>0</td>
</tr>
<tr>
<td>β-lactamase inhibitor/β-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lactamase inhibitor combinations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cell wall synthesis inhibitors (bind</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>penicillin-binding protein)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Penicillin G 1000+/-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+/-</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nafcillin Nallpen, Unipen 1100</td>
<td>+/-</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Piperacillin Pipracil 1000</td>
<td>+/-</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>+/-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Penicillin/α-lactamase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>inhibitor combinations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ampicillin/sulbactam Unasyn 1100</td>
<td>+/-</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>+/-</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ticarcillin/clavulanate Timentin 1100</td>
<td>+/-</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>+/-</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Piperacillin/tazobactam Zosyn 1100</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>+/-</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>First-generation cephalosporins</td>
<td></td>
<td>Cell wall synthesis inhibitors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cefazolin, cephalexin Ancef, Keflex 1100</td>
<td>+/-</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Second-generation cephalosporins</td>
<td></td>
<td>Cell wall synthesis inhibitors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cefoxitin Mefoxin 1100</td>
<td>+/-</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cefotetan Cefotan 1100</td>
<td>+/-</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cefuroxime Ceftin 1100</td>
<td>+/-</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Third- and fourth-generation cephalosporins</td>
<td></td>
<td>Cell wall synthesis inhibitors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ceftriaxone Rocephin 1100</td>
<td>+/-</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ceftazidime Fortaz 1100</td>
<td>+/-</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cefepime Maxipime 1100</td>
<td>+/-</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cefotaxime Cefotaxime 1100</td>
<td>+/-</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ceftaroline Teflaro 111100</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

(Continued)
### Bacterial agents (Continued)

<table>
<thead>
<tr>
<th>ANTIBIOTIC CLASS, GENERIC NAME</th>
<th>TRADE NAME</th>
<th>MECHANISM OF ACTION</th>
<th>S Pyogenes</th>
<th>MSSA</th>
<th>MRSA</th>
<th>S epidermidis</th>
<th>Enterococcus</th>
<th>VRE</th>
<th>E coli</th>
<th>P aeruginosa</th>
<th>ANAEROBES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carbapenems</td>
<td>Imipenem-cilastatin</td>
<td>Primaxin</td>
<td>Cell wall synthesis inhibitors</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>+/-</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Meropenem</td>
<td>Merrem</td>
<td>Cell wall synthesis inhibitors</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Ertapenem</td>
<td>Invanz</td>
<td>Cell wall synthesis inhibitors</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Aztreonam</td>
<td>Azactam</td>
<td>Cell wall synthesis inhibitors</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Aminoglycosides</td>
<td>Gentamicin</td>
<td></td>
<td>Alteration of cell membrane, binding and inhibition of 30S ribosomal subunit</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>+/-</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Tobramycin, amikacin</td>
<td></td>
<td>Alteration of cell membrane, binding and inhibition of 30S ribosomal subunit</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>+/-</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Fluoroquinolones</td>
<td>Ciprofloxacin</td>
<td>Cipro</td>
<td>Inhibit topoisomerase II and IV (DNA synthesis inhibition)</td>
<td>+/-</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Levofloxacin</td>
<td>Levaquin</td>
<td>Inhibit topoisomerase II and IV (DNA synthesis inhibition)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>+/-</td>
</tr>
<tr>
<td>Glycopeptides</td>
<td>Vancomycin</td>
<td>Vancocin</td>
<td>Cell wall synthesis inhibition (peptidoglycan synthesis inhibition)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Quinupristin-dalfopristin</td>
<td>Synercid</td>
<td>Inhibits 2 sites on 50S ribosome (protein synthesis inhibition)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Drug</td>
<td>Mechanism</td>
<td>E. coli</td>
<td>MRSA</td>
<td>MSSA</td>
<td>P. aeruginosa</td>
<td>S. epidermidis</td>
<td>S. pyogenes</td>
<td>VRE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------------------------------------------------</td>
<td>---------</td>
<td>------</td>
<td>------</td>
<td>--------------</td>
<td>---------------</td>
<td>-------------</td>
<td>-----</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linezolid</td>
<td>Inhibits 50S ribosomal activity</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ceftriaxone</td>
<td>Binds bacterial membrane, results in depolarization, lysis</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daptomycin</td>
<td>Inhibits DNA-dependent RNA polymerase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>+/–</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rifampin</td>
<td>Inhibits DNA-dependent RNA polymerase</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clindamycin</td>
<td>Inhibits 50S ribosomal activity</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metronidazole</td>
<td>Production of toxic intermediates (free radicals)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macrolides</td>
<td>Inhibits sequential steps of folate metabolism</td>
<td>+/-</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erythromycin</td>
<td>Inhibit 50S ribosomal activity, protein synthesis inhibition</td>
<td>1</td>
<td>+/–</td>
<td>0</td>
<td>+/-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Azithromycin</td>
<td>Inhibit 50S ribosomal activity, protein synthesis inhibition</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clarithromycin</td>
<td>Inhibit 50S ribosomal activity, protein synthesis inhibition</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trimethoprim-</td>
<td>Inhibit 50S ribosomal activity, protein synthesis inhibition</td>
<td>+/-</td>
<td>1</td>
<td>0</td>
<td>+/-</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sulfamethoxazole</td>
<td>Inhibit 50S ribosomal activity, protein synthesis inhibition</td>
<td>+/-</td>
<td>1</td>
<td>0</td>
<td>+/-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tetracyclines</td>
<td>Bind 30S ribosomal unit (protein synthesis inhibition)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minocin</td>
<td>Bind 30S ribosomal unit (protein synthesis inhibition)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doxycycline</td>
<td>Bind 30S ribosomal unit (protein synthesis inhibition)</td>
<td>1</td>
<td>+/–</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tigacycline</td>
<td>Bind 30S ribosomal unit (protein synthesis inhibition)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*E. coli* = Escherichia coli; MRSA = methicillin-resistant *Staphylococcus aureus*; MSSA = methicillin-sensitive *S. aureus*; *P. aeruginosa* = *Pseudomonas aeruginosa*; *S. epidermidis* = *Staphylococcus epidermidis*; *S. pyogenes* = *Streptococcus pyogenes*; VRE = vancomycin-resistant *Enterococcus*

Activity: +/- = variable activity; 0 = no activity.

The sensitivities printed here are generalizations. The clinician should confirm sensitivity patterns at the locale where the patient is being treated since these patterns may vary widely depending on location.
inoculum of $10^5$ CFU/mL of the organism isolated from the site of infection or bloodstream. Sensitivities are reported in relation to the achievable blood level of each antibiotic in a panel of agents. The least toxic, least expensive agent to which the organism is most sensitive should be selected. Serious or recrudescence infection may require therapy with two or more agents, particularly if a multidrug-resistant pathogen is causative, limiting therapeutic options to drugs to which the organism is only moderately sensitive. Commonly, an agent may be administered intravenously for 1 to 2 weeks, followed by treatment with an oral drug. However, this should only be undertaken in patients who demonstrate progressive clinical improvement, and the oral agent should be capable of achieving high serum levels as well (e.g., fluoroquinolones).

The 2016 Surgical Infection Society guidelines on management of intra-abdominal infection recommend antibiotic duration of no more than 24 hours in patients with traumatic bowel perforation who receive surgical treatment within 12 hours, gastroduodenal perforations operated upon within 24 hours, ischemic nonperforated bowel, and gangrenous acute appendicitis or cholecystitis without perforation. More extensive intraperitoneal infection (perforated appendicitis, for example) should have treatment limited to 4 days. Patients with a greater degree of contamination may require longer courses of therapy; as in all facets of clinical practice, the therapeutic plan must be individualized to the patient. In the later phases of postoperative antibiotic treatment of serious intra-abdominal infection, the absence of an elevated white blood cell (WBC) count, lack of band forms of PMNs on peripheral smear, and lack of fever ($<38^\circ$C [100.5$^\circ$F]) provide close to complete assurance that infection has been eradicated. $^1$ There is also emerging data that suggest following a patient’s procalcitonin level may provide the clinician with useful information regarding whether an infection has resolved and allow more expedient cessation of therapy. $^{32,33}$ Patients who do not improve with 5 to 7 days of antibiotic therapy should be reevaluated for inadequate source control or a new extra-abdominal source of infection.

**Allergy** to antimicrobial agents must be considered prior to prescribing them. First, it is important to ascertain whether a patient has had any type of allergic reaction in association with administration of a particular antibiotic. However, one should take care to ensure that the purported reaction consists of true allergic symptoms and signs, such as urticaria, bronchospasm, or other similar manifestations, rather than indigestion or nausea. Penicillin allergy is quite common, the reported incidence ranging from 0.7% to 10%. Although avoiding the use of any β-lactam drug is appropriate in patients who manifest significant allergic reactions to penicillins, the incidence of cross-reactivity appears low for all related agents, with 1% cross-reactivity for carbapenems, 5% to 7% cross-reactivity for cephalosporins, and extremely small or nonexistent cross-reactivity for monobactams. $^{34}$

Severe allergic manifestations, such as anaphylaxis, to a specific class of agents generally preclude the use of any agents in that class, except under circumstances in which use of a certain drug represents a lifesaving measure. In some centers, patients undergo intradermal testing using a dilute solution of a particular antibiotic to determine whether a severe allergic reaction would be elicited by parenteral administration. A pathway, including such intradermal testing, has been effective in reduction of vancomycin use to 16% in surgical patients with reported allergy to penicillin. $^{35}$ This type of testing rarely is employed because it is simpler to select an alternative class of agent. Should administration of a specific agent to which the patient is

### Table 6-6

<table>
<thead>
<tr>
<th>SITE</th>
<th>ANTIBIOTIC</th>
<th>ALTERNATIVE (E.G., PENICILLIN ALLERGIC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiovascular surgery</td>
<td>Cefazolin, cefuroxime</td>
<td>Vancomycin, clindamycin</td>
</tr>
<tr>
<td>Gastrointestinal area</td>
<td>Cefazolin</td>
<td>Clindamycin or vancomycin + aminoglycoside or aztreonem or fluoroquinolone</td>
</tr>
<tr>
<td>Small intestine, nonobstructed</td>
<td>Cefazolin</td>
<td>Clindamycin or vancomycin + aminoglycoside or aztreonem or fluoroquinolone</td>
</tr>
<tr>
<td>Biliary tract: open procedure, laparoscopic high risk</td>
<td>Cefazolin, cefoxitin, cefotetan, ceftriaxone, ampicillin-sulbactam</td>
<td>Clindamycin or vancomycin + aminoglycoside or aztreonem or fluoroquinolone</td>
</tr>
<tr>
<td>Biliary tract: laparoscopic low risk</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Appendectomy, uncomplicated</td>
<td>Cefoxitin, cefotetan, cefazolin + metronidazole</td>
<td>Clindamycin + aminoglycoside or aztreonem or fluoroquinolone Metronidazole + aminoglycoside or fluoroquinolone</td>
</tr>
<tr>
<td>Colorectal surgery, obstructed small intestine</td>
<td>Cefazolin or ceftriaxone plus metronidazole, ertapenem, cefoxitin, cefotetan, ampicillin-sulbactam</td>
<td>Clindamycin + aminoglycoside or aztreonem or fluoroquinolone Metronidazole + aminoglycoside or fluoroquinolone</td>
</tr>
<tr>
<td>Head and neck; clean contaminated</td>
<td>Cefazolin or cefuroxime + metronidazole, ampicillin-sulbactam</td>
<td>Clindamycin</td>
</tr>
<tr>
<td>Neurosurgical procedures</td>
<td>Cefazolin</td>
<td>Clindamycin, vancomycin</td>
</tr>
<tr>
<td>Orthopedic surgery</td>
<td>Cefazolin, ceftriaxone</td>
<td>Clindamycin, vancomycin</td>
</tr>
<tr>
<td>Breast, hernia</td>
<td>Cefazolin</td>
<td>Clindamycin, vancomycin</td>
</tr>
</tbody>
</table>

allergic become necessary, desensitization using progressively higher doses of antibiotic can be undertaken, providing the initial testing does not cause severe allergic manifestations.

Misuse of antimicrobial agents is rampant in both the inpatient and outpatient settings, and is associated with an enormous financial impact on healthcare costs, adverse reactions due to drug toxicity and allergy, the occurrence of new infections such as *Clostridium difficile* colitis, and the development of multiagent drug resistance among nosocomial pathogens. Each of these factors has been directly correlated with overall drug administration. It has been estimated that in the United States in excess of $20 billion is spent on antibiotics each year. The responsible practitioner limits prophylaxis to the period during the operative procedure, does not convert prophylaxis into empiric therapy except under well-defined conditions, sets the duration of antibiotic therapy from the outset, curtails antibiotic administration when clinical and microbiologic evidence does not support the presence of an infection, and limits therapy to a short course in every possible instance. Prolonged treatment associated with drains and tubes has not been shown to be beneficial.

**INFECTIONS OF SIGNIFICANCE IN SURGICAL PATIENTS**

**Surgical Site Infections**

Surgical site infections (SSIs) are infections of the tissues, organs, or spaces exposed by surgeons during performance of an invasive procedure. SSIs are classified into incisional and organ/space infections, and the former are further subclassified into superficial (limited to skin and subcutaneous tissue) and deep incisional categories. The development of SSIs is related to three factors: (a) the degree of microbial contamination of the wound during surgery; (b) the duration of the procedure; and (c) host factors such as diabetes, malnutrition, obesity, immune suppression; and a number of other underlying disease states. Table 6-7 lists risk factors for development of SSIs. By definition, an incisional SSI has occurred if a surgical wound drains purulent material or if the surgeon judges it to be infected and opens it.

Surgical wounds are classified based on the presumed magnitude of the bacterial load at the time of surgery (Table 6-8). Clean wounds (class I) include those in which no infection is present; only skin microflora potentially contaminate the wound, and no hollow viscus that contains microbes is entered. Class I D wounds are similar except that a prosthetic device (e.g., mesh or valve) is inserted. Clean/contaminated wounds (class II) include those in which a hollow viscus such as the respiratory, alimentary, or genitourinary tracts with indigenous bacterial flora is opened under controlled circumstances without significant spillage of contents.

While elective colorectal cases have classically been included as class II cases, a number of studies in the last decade have documented higher SSI rates (9–25%). One study identified two-thirds of infections presenting after discharge from hospital, highlighting the need for careful follow-up of these patients. Infection is also more common in cases involving entry into the rectal space. In a recent single-center quality improvement study using a multidisciplinary approach, one group of clinicians has demonstrated the ability to decrease SSI from 9.8% to 4.0%. The responsible practitioner limits prophylaxis to the period during the operative procedure, does not convert prophylaxis into empiric therapy except under well-defined conditions, sets the duration of antibiotic therapy from the outset, curtails antibiotic administration when clinical and microbiologic evidence does not support the presence of an infection, and limits therapy to a short course in every possible instance. Prolonged treatment associated with drains and tubes has not been shown to be beneficial.

### Table 6-7

**Risk factors for development of surgical site infections**

<table>
<thead>
<tr>
<th>Patient factors</th>
<th>Microbial factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Older age</td>
<td>Chronic skin disease</td>
</tr>
<tr>
<td>Immunosuppression</td>
<td>Carcinoid syndrome</td>
</tr>
<tr>
<td>Obesity</td>
<td>Prolong hospitalization (leading to nosocomial organisms)</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>Toxin secretion</td>
</tr>
<tr>
<td>Chronic inflammatory process</td>
<td>Resistance to clearance (e.g., capsule formation)</td>
</tr>
<tr>
<td>Malnutrition</td>
<td></td>
</tr>
<tr>
<td>Smoking</td>
<td></td>
</tr>
<tr>
<td>Renal failure</td>
<td></td>
</tr>
<tr>
<td>Peripheral vascular disease</td>
<td></td>
</tr>
<tr>
<td>Anemia</td>
<td></td>
</tr>
<tr>
<td>Radiation</td>
<td></td>
</tr>
<tr>
<td>Chronic skin disease</td>
<td></td>
</tr>
<tr>
<td>Carrier state (e.g., chronic <em>Staphylococcus</em> carriage)</td>
<td></td>
</tr>
<tr>
<td>Recent operation</td>
<td></td>
</tr>
</tbody>
</table>

### Table 6-8

**Wound class, representative procedures, and expected infection rates**

<table>
<thead>
<tr>
<th>WOUND CLASS</th>
<th>EXAMPLES OF CASES</th>
<th>EXPECTED INFECTION RATES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clean (class I)</td>
<td>Hernia repair, breast biopsy</td>
<td>1–2%</td>
</tr>
<tr>
<td>Clean/contaminated (class II)</td>
<td>Cholecystectomy, elective GI surgery (not colon)</td>
<td>2.1–9.5%</td>
</tr>
<tr>
<td>Clean/contaminated (class II)</td>
<td>Colorectal surgery</td>
<td>4–14%</td>
</tr>
<tr>
<td>Contaminated (class III)</td>
<td>Penetrating abdominal trauma, large tissue injury, enterotomy during bowel obstruction</td>
<td>3.4–13.2%</td>
</tr>
<tr>
<td>Dirty (class IV)</td>
<td>Perforated diverticulitis, necrotizing soft tissue infections</td>
<td>3.1–12.8%</td>
</tr>
</tbody>
</table>
Contaminated wounds (class III) include open accidental wounds encountered early after injury, those with extensive introduction of bacteria into a normally sterile area of the body due to major breaks in sterile technique (e.g., open cardiac massage), gross spillage of viscus contents such as from the intestine, or incision through inflamed, albeit nonpurulent tissue. Dirty wounds (class IV) include traumatic wounds in which a significant delay in treatment has occurred and in which necrotic tissue is present, those created in the presence of overt infection as evidenced by the presence of purulent material, and those created to access a perforated viscus accompanied by a high degree of contamination. The microbiology of SSIs is reflective of the initial host microflora such that SSIs following creation of a class I wound are invariably caused by skin microbes found on that portion of the body, while SSIs subsequent to a class II wound created for the purpose of elective colon resection may be caused by either skin microbes or colonic microflora, or both.

Surgical management of the wound is a critical determinant of the propensity to develop an SSI. In healthy individuals, classes I and II wounds may be closed primarily, while skin closure of class III and IV wounds is associated with high rates of incisional SSIs (~25–50%). The superficial aspects of these latter types of wounds should be packed open and allowed to heal by secondary intention, although selective use of delayed primary closure has been associated with a reduction in incisional SSI rates. One clear example based on data from clinical trials is that class III wounds in healthy patients undergoing appendectomy for perforated or gangrenous appendicitis can be primarily closed as long as antibiotic therapy directed against aerobes and anaerobes is administered. This practice leads to SSI rates of approximately 3% to 4%.

Recent investigations have studied the effect of additional maneuvers in an attempt to further reduce the rate of SSIs. The adverse effects of hyperglycemia on WBC function have been well described. A number of studies in patients undergoing several different types of surgery describe increased risk of SSI in patients with hyperglycemia, and the 2017 CDC guidelines for prevention of surgical site infection recommend maintaining blood glucose <200 mg/dL (11.1 mmol/L) in all patients during the perioperative period.

The respective effects of body temperature and the level of inhaled oxygen during surgery on SSI rates also have been studied, and both hypothermia and hypoxia during surgery are associated with a higher rate of SSI. There is conflicting evidence regarding whether supplying higher levels of inhaled oxygen to perioperative patients reduces the rate of SSI. Although an initial study provided evidence that patients who received high levels of inhaled oxygen during colorectal surgery developed fewer SSIs, a later meta-analysis suggested that the overall benefit is small and may not warrant use. The 2017 CDC guidelines, however, support administration of increased FiO₂ during surgery and after extubation in patients with normal pulmonary function receiving general anesthesia as there has been some evidence of benefit. Further evaluation via multicenter studies is needed prior to implementation of hyperoxia as standard therapy, but it is clear that intraoperative hypothermia and hypoxia should be prevented.

Effective therapy for incisional SSIs consists solely of incision and drainage without the additional use of antibiotics. Antibiotic therapy is reserved for patients in whom evidence of significant cellulitis is present, or who concurrently manifest a systemic inflammatory response syndrome. The open wound often is allowed to heal by secondary intention, with dressings being changed as the clinical team deems appropriate. The use of topical antibiotics and antiseptics to further wound healing remains unproven, although anecdotal studies indicate their potential utility in complex wounds that do not heal with routine measures. Despite a paucity of prospective studies, vacuum-assisted closure is increasingly used in management of large, complex open wounds and can be applied to wounds in locations that are difficult to manage with dressings (Fig. 6-1). One also should consider obtaining wound cultures in patients who develop SSIs and who have been hospitalized or reside in long-term care facilities due to the increasing incidence of infection caused by multidrug-resistant organisms.

In the United States, hospitals are required to conduct surveillance for the development of SSIs for a period of 30 days.

![Figure 6-1](image_url). Negative pressure wound therapy in a patient after amputation for wet gangrene (A) and in a patient with enterocutaneous fistula (B). It is possible to adapt these dressings to fit difficult anatomy and provide appropriate wound care while reducing frequency of dressing change. It is important to evaluate the wound under these dressings if the patient demonstrates signs of sepsis with an unidentified source, since typical clues of wound sepsis such as odor and drainage are hidden by the suction apparatus.
after the operative procedure. Such surveillance has been associated with greater awareness and a reduction in SSI rates, probably in large part based upon the impact of observation and promotion of adherence to appropriate care standards. Beginning in 2012, all hospitals receiving reimbursement from the Centers for Medicare & Medicaid Services (CMS) are required to report SSIs.

A recent refinement of risk indexes has been implemented through the National Healthcare Safety Network, a secure, web-based system of surveillance used by the CDC for surveillance of healthcare-associated infections. This refinement utilized data reported from 847 hospitals in nearly one million patients over a 2-year period to develop procedure-specific risk indices for SSIs.

SSIs are associated with considerable morbidity and occasional lethality, as well as substantial healthcare costs and patient inconvenience and dissatisfaction. A number of healthcare organizations within the United States are interested in evaluating performance of hospitals and physicians with respect to implementing processes that support delivery of standard of care. One major process of interest is reduction in SSIs, since the morbidity (and subsequent cost) of this complication is high. Several of these organizations are noted in Table 6-9. Appropriate guidelines in this area incorporating the principles discussed previously have been developed and disseminated. However, observers have noted that adherence to these guidelines has been poor. Most experts believe that better adherence to evidence-based practice recommendations and implementing systems of care with redundant safeguards will result in reduction of surgical complications and better patient outcomes. More important, the CMS, the largest third-party insurance payer in the United States, has required reporting by hospitals of many processes related to reduction of surgical infections, including appropriate use of perioperative antibiotics. This information, which is reported publicly by hospitals, has led to significant improvement in reported rates of these process measures. However, the effect of this approach on the incidence of SSIs is not known at this time.

### Intra-Abdominal Infections

Microbial contamination of the peritoneal cavity is termed peritonitis or intra-abdominal infection and is classified according to etiology. Primary microbial peritonitis occurs when microbes invade the normally sterile confines of the peritoneal cavity via hematogenous dissemination from a distant source of infection or direct inoculation. This process is more common among patients who retain large amounts of peritoneal fluid due to ascites, and among those individuals who are being treated for renal failure via peritoneal dialysis. These infections invariably are monomicrobial and rarely require surgical intervention. The diagnosis is established based on identification of risk factors as noted previously, physical examination that reveals diffuse tenderness and guarding without localized findings, absence of a surgically treatable source of infection on an imaging study, and the presence of more than 250 neutrophils/mL in fluid obtained via paracentesis. Cultures typically will demonstrate the presence of gram-positive organisms in patients undergoing peritoneal dialysis. In patients without this risk factor, the most common etiologic organisms are E coli, K pneumoniae, and S pneumoniae. Treatment consists of administration of an antibiotic to which the organism is sensitive; often 14 to 21 days of therapy are required. Removal of indwelling devices, if present, may be required for effective therapy of recurrent infections.

Secondary microbial peritonitis occurs subsequent to contamination of the peritoneal cavity due to perforation or severe inflammation and infection of an intra-abdominal organ. Examples include appendicitis, perforation of any portion of the gastrointestinal tract, or diverticulitis. As noted previously, effective therapy requires source control to resect or repair the diseased organ; debridement of necrotic, infected tissue and debris; and administration of antimicrobial agents directed against aerobes and anaerobes. This type of antibiotic regimen should be chosen because in most patients the precise diagnosis cannot be established until exploratory laparotomy is performed, and the most morbid form of this disease process is colonic perforation, due to the large number of microbes present. A combination of agents or single agents with a broad spectrum of activity can be used for this purpose; conversion of a parenteral to an oral regimen when the patient’s ileus resolves provides results similar to those achieved with intravenous antibiotics. Effective source control and antibiotic therapy is associated with low failure rates and a mortality rate of approximately 5% to 6%; inability to control the source of infection is associated with mortality greater than 40%.

The response rate to effective source control and use of appropriate antibiotics has remained approximately 70% to 90% over the past several decades. Patients in whom standard therapy fails typically develop one or more of the following: an intra-abdominal abscess, leakage from a gastrointestinal anastomosis leading to postoperative peritonitis, or tertiary (persistent) peritonitis. The latter is a poorly understood entity that is more common in immunosuppressed patients in whom peritoneal host defenses do not effectively clear or sequester

<table>
<thead>
<tr>
<th>Table 6-9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality improvement organizations of interest to surgeons in the United States</td>
</tr>
<tr>
<td>ABBREVIATION</td>
</tr>
<tr>
<td>NSQIP</td>
</tr>
<tr>
<td>IHI</td>
</tr>
<tr>
<td>NCQA</td>
</tr>
<tr>
<td>SIS</td>
</tr>
<tr>
<td>CDC</td>
</tr>
</tbody>
</table>
the initial secondary microbial peritoneal infection. Microbes such as \textit{E faecalis} and \textit{faecium}, \textit{S epidermidis}, \textit{C albicans}, and \textit{P aeruginosa} commonly are identified, typically in combination, and their presence may be due to their lack of responsiveness to the initial antibiotic regimen, coupled with diminished activity of host defenses. Unfortunately, even with effective antimicrobial agent therapy, this disease process is associated with mortality rates in excess of 50%.\textsuperscript{14}

Formerly, the presence of an intra-abdominal abscess mandated surgical reexploration and drainage. Today, the vast majority of such abscesses can be effectively diagnosed via abdominal computed tomographic (CT) imaging techniques and drained percutaneously. Surgical intervention is reserved for those individuals who harbor multiple abscesses, those with abscesses in proximity to vital structures such that percutaneous drainage would be hazardous, and those in whom an ongoing source of contamination (e.g., enteric leak) is identified. The necessity of antimicrobial agent therapy and precise guidelines that dictate duration of catheter drainage have not been established. A short course (3 to 5 days) of antibiotics that possess aerobic and anaerobic activity seems reasonable so long as the patient has good clinical response to therapy, and most practitioners leave the drainage catheter in situ until it is clear that cavity collapse has occurred, output is less than 10 to 20 mL/d, no evidence of an ongoing source of contamination is present, and the patient’s clinical condition has improved.\textsuperscript{33}

**Organ-Specific Infections**

Hepatic abscesses are rare, currently accounting for approximately 15 per 100,000 hospital admissions in the United States. Pyogenic abscesses account for approximately 80\% of cases, the remaining 20\% being equally divided among parasitic and fungal forms.\textsuperscript{65} Formerly, pyogenic liver abscesses mainly were caused by phlebitis due to neglected appendicitis or diverticulitis. Today, manipulation of the biliary tract to treat a variety of diseases has become a more common cause, although in nearly 50\% of patients no cause is identified. The most common aerobic bacteria identified in recent series include \textit{E coli}, \textit{K pneumoniae}, and other enteric bacilli, enterococci, and \textit{Pseudomonas} spp., while the most common anaerobic bacteria are \textit{Bacteroides} spp., anaerobic streptococci, and \textit{Fusobacterium} spp. \textit{C albicans} and other related yeast cause the majority of fungal hepatic abscesses. Small (<1 cm), multiple abscesses should be sampled and treated with a 4- to 6-week course of antibiotics. Larger abscesses are generally amenable to percutaneous drainage, with parameters for antibiotic therapy and drain removal similar to those mentioned previously. Splenic abscesses are extremely rare and are treated in a similar fashion. Recurrent hepatic or splenic abscesses may require operative intervention—unroofing and marsupialization or splenectomy, respectively.

Secondary pancreatic infections (e.g., infected pancreatic necrosis or pancreatic abscess) occur in approximately 10\% to 15\% of patients who develop severe pancreatitis with necrosis. The surgical treatment of this disorder was pioneered by Bradley and Allen, who noted significant improvements in outcome for patients undergoing repeated pancreatic debridement of infected pancreatic necrosis.\textsuperscript{66} Care of patients with severe acute pancreatitis includes staging with dynamic, contrast-enhanced helical CT scan to evaluate the extent of pancreatitis (unless significant renal dysfunction exists, in which case one should forego the use of contrast material) coupled with the use of one of several prognostic scoring systems. Patients who exhibit clinical signs of instability (e.g., oliguria, hypoxemia, large-volume fluid resuscitation) should be carefully monitored in the ICU and undergo follow-up contrast CT examination when renal function has stabilized to evaluate for development of local pancreatic complications (Fig. 6-2). Routine use of prophylactic antibiotics to prevent infected pancreatic necrosis is not indicated. Early enteral feeding using nasojejunal feeding tubes placed past the ligament of Treitz has been associated with decreased development of infected pancreatic necrosis, possibly due to a decrease in gut translocation of bacteria.\textsuperscript{67,68}

The presence of secondary pancreatic infection should be suspected in patients whose systemic inflammatory response (fever, elevated WBC count, or organ dysfunction) fails to resolve, or in those individuals who initially recuperate, only to develop sepsis syndrome 2 to 3 weeks later. CT-guided aspiration of fluid from the pancreatic bed for performance of Gram stain and culture analysis can be useful. A positive Gram stain or culture from CT-guided aspiration, or identification of gas within the pancreas on CT scan, mandate surgical intervention.

The approach of open necrosectomy with repeated debridements, although life-saving, is associated with significant morbidity and prolonged hospitalization. Efforts to reduce the amount of surgical injury, while still preserving the improved outcomes associated with debridement of the infected sequestrum, have led to a variety of less invasive approaches, including endoscopic and laparoscopic techniques.\textsuperscript{69} There are a limited number of randomized trials reporting the use of these new techniques. An important concept common to all of these approaches, however, is the attempt to delay surgical intervention, since a number of trials have identified increased mortality when intervention occurs during the first 2 weeks of illness.

Data supporting the use of endoscopic approaches to infected pancreatic necrosis include nearly a dozen case series and a randomized trial.\textsuperscript{70,71} The reported mortality rate was 5\%, with a 30\% complication rate. Most authors noted the common requirement for multiple endoscopic debridements (similar to the open approach), with a median of four sessions required. Fewer series report experience with the laparoscopic approach, either transgastric or transperitoneal, entering the necrosis through the transverse mesocolon or gastrocolic ligament. Laparoscopic intervention is limited by the difficulty in achieving

*Figure 6-2. Contrast-enhanced CT scan of pancreas 1.5 weeks after presentation showing large central peripancreatic fluid collection (arrow).*
Infections of the Skin and Soft Tissue
These infections can be classified according to whether surgical intervention is required. For example, superficial skin and skin structure infections such as cellulitis, erysipelas, and lymphangitis invariably are effectively treated with antibiotics alone, although a search for a local underlying source of infection should be undertaken. Generally, drugs that possess activity against the causative gram-positive skin microflora are selected. Furuncles or boils may drain spontaneously or require surgical incision and drainage. Antibiotics are prescribed if significant cellulitis is present or if cellulitis does not rapidly resolve after surgical drainage. Community-acquired methicillin-resistant S aureus (MRSA) infection should be suspected if infection persists after treatment with adequate drainage and administration of first-line antibiotics. These infections may require more aggressive drainage and altered antimicrobial therapy.

Aggressive soft tissue infections are rare, difficult to diagnose, and require immediate surgical intervention plus administration of antimicrobial agents. Failure to rapidly recognize and treat these infections results in an extremely high mortality rate (~80–100%), and even with expedient therapy mortality rates are high (16–24%). Eponyms and differing classifications in the past have led to a hodgepodge of terminology—such as Meleney’s synergistic gangrene, Fournier’s gangrene, rapidly spreading cellulitis, gas gangrene, and necrotizing fasciitis—regarding these serious infections. Today it seems best to delineate them based on the soft tissue layer(s) of involvement.
BASIC CONSIDERATIONS

PART I

SKIN AND SOFT TISSUE INFECTIONS

Skin and soft tissue infections (e.g., skin and superficial soft tissue, deep soft tissue, and muscle) and the pathogen(s) that cause them.

Patients at risk for these types of infections include those who are elderly, immunosuppressed, or diabetic, and/or who suffer from peripheral vascular disease, though extremely aggressive necrotizing soft tissue infections (often caused by streptococci) have been described among healthy individuals as well. The common thread among these host factors appears to be compromise of the fascial blood supply, and if this is coupled with the introduction of exogenous microbes, the result can be devastating.

Initially, the diagnosis is established solely upon a constellation of clinical findings, not all of which are present in every patient. Not surprisingly, patients often develop sepsis syndrome or septic shock without an obvious cause. The extremities, perineum, trunk, and torso are most commonly affected, in that order. Careful examination should be undertaken for an entry site such as a small break or sinus in the skin from which grayish, turbid semipurulent material (“dishwater pus”) can be expressed, as well as for the presence of skin changes (bronze hue or brawny induration), blebs, or crepitus. The patient often develops pain at the site of infection that appears to be out of proportion to any of the physical manifestations. Any of these findings mandates immediate surgical intervention, which should consist of incision and direct visualization of potentially infected tissue (including deep soft tissue, fascia, and underlying muscle) and radical resection of affected areas. Radiologic studies should not be undertaken in patients in whom the diagnosis seriously is considered, as they delay surgical intervention and frequently provide confusing information. Unfortunately, surgical extirpation of infected tissue frequently entails amputation and/or disfiguring procedures; the surgeon must bear in mind that incomplete procedures are associated with higher rates of morbidity and mortality and debride all nonviable tissue (Fig. 6-4).

During the procedure, a Gram stain should be performed on tissue fluid. Antimicrobial agents directed against gram-positive and gram-negative aerobes and anaerobes (e.g., vancomycin plus a carbapenem), as well as high-dose aqueous penicillin G (16,000,000 to 20,000,000 U/d), the latter to treat clostridial pathogens, should be administered. Approximately 50% of such infections are polymicrobial, the remainder being caused by a single organism such as S pyogenes, P aeruginosa, or C perfringens. The microbiology of these polymicrobial infections is similar to that of secondary microbial peritonitis, with the exception that gram-positive cocci are more commonly encountered. Most patients should be returned to the operating room on a scheduled basis to determine if disease progression has occurred. If so, additional resection of infected tissue and debridement should take place. Antibiotic therapy can be refined based on culture and sensitivity results, particularly in the case of monomicrobial soft tissue infections. Hyperbaric oxygen therapy may be of use in patients with infection caused by gas-forming organisms (e.g., C perfringens), although the evidence to support efficacy is limited to underpowered studies and case reports. In the absence of such infection, hyperbaric oxygen therapy has not been shown to be effective.

Postoperative Nosocomial Infections

Surgical patients are prone to develop a wide variety of nosocomial infections during the postoperative period, which include SSIs, UTIs, pneumonia, and bacteremia. SSIs are discussed earlier, and the latter types of nosocomial infections are related to prolonged use of indwelling tubes and catheters for the purpose of urinary drainage, ventilation, and venous and arterial access, respectively.

The presence of a postoperative UTI should be considered based on urinalysis demonstrating WBCs or bacteria, a positive test for leukocyte esterase, or a combination of these elements. The diagnosis is established after >10^4 CFU/mL of microbes are identified by culture techniques in symptomatic patients, or >10^5 CFU/mL in asymptomatic individuals. Treatment for 3 to 5 days with a single antibiotic directed against the most common organisms (e.g., E Coli, K pneumoniae) that achieves high levels in the urine is appropriate. Initial therapy is directed by Gram stain results and is refined as culture results become available. Postoperative surgical patients should have indwelling urinary catheters removed as quickly as possible to avoid the development of a UTI.

Prolonged mechanical ventilation is associated with nosocomial pneumonia. These patients present with more severe disease, are more likely to be infected with drug-resistant pathogens, and suffer increased mortality compared to patients who develop community-acquired pneumonia. The diagnosis of pneumonia is established by presence of purulent sputum, elevated leukocyte count, fever, and new chest X-ray abnormalities, such as consolidation. The presence of two of the clinical findings, plus chest X-ray findings, significantly increases the likelihood of pneumonia. Consideration should be given to performing bronchoalveolar lavage to obtain samples for Gram stain and culture. Some authors advocate quantitative cultures as a means to identify a threshold for diagnosis. Surgical patients should be weaned from mechanical ventilation as soon as feasible, based on oxygenation and inspiratory effort, as risk of pneumonia increases with increased time on mechanical ventilation.

Infection associated with indwelling intravascular catheters is a common problem among hospitalized patients. Because of the complexity of many surgical procedures, these devices are increasingly used for physiologic monitoring, vascular access, drug delivery, and hyperalimentation. Among the several million catheters inserted each year in the United States, approximately 25% will become colonized, and approximately 5% will be associated with bacteremia. Duration of catheterization, insertion or manipulation under emergency or nonsterile conditions, use for hyperalimentation, and the use of multilumen catheters increase the risk of infection. Use of a central line insertion protocol that includes full barrier precautions and chlorhexidine skin prep has been shown to decrease the incidence of infection. Although no randomized trials have been performed, peripherally inserted central venous catheters have a catheter-related infection rate similar to those inserted in the subclavian or jugular veins.

Many patients who develop intravascular catheter infections are asymptomatic, often exhibiting solely an elevation in the blood WBC count. Blood cultures obtained from a peripheral site and drawn through the catheter that reveals the presence of the same organism increase the index of suspicion for the presence of a catheter infection. Obvious purulence at the exit site of the skin tunnel, severe sepsis syndrome due to any type of organism when other potential causes have been excluded, or bacteremia due to gram-negative aerobes or fungi should lead to catheter removal. Selected catheter infections due to low-virulence microbes such as S epidermidis can be effectively treated in approximately 50% to 60% of patients with a 14- to 21-day course of an antibiotic, which should be considered when no other vascular access site exists. The use of antibiotic-bonded catheters and chlorhexidine sponges at the insertion
site has been associated with lower rates of colonization. Use of ethanol or antimicrobial catheter “locks” have shown promise in reducing incidence of infection in dialysis catheters. The surgeon should carefully consider the need for any type of vascular access devices, rigorously attend to their maintenance to prevent infection, and remove them as quickly as possible. Use of systemic antibacterial or antifungal agents to prevent catheter infection is of no utility and is contraindicated.

**Sepsis**

As previously discussed, sepsis is increasing in incidence, with more than 1.1 million cases estimated per year in the United States with an annual cost of $24 billion. This rate is expected to increase as the population of aged in the United States increases. One third of sepsis cases occur in surgical populations and sepsis is a major cause of morbidity and mortality. The treatment of sepsis has improved over the last decade, with mortality rates dropping to under 30%. Factors contributing to this improvement relate both to recent randomized prospective trials demonstrating improved outcomes with new therapies, and to improvements in the process of care delivery to the sepsis patient. The “Surviving Sepsis Campaign,” a multidisciplinary group that develops treatment recommendations, published guidelines incorporating evidence-based sepsis treatment strategies most recently in 2016. These guidelines are summarized in Table 6-10.

**FIGURE 6-4.** Necrotizing soft tissue infection. (A) This patient presented with hypotension due to severe late necrotizing fasciitis and myositis due to β-hemolytic streptococcal infection. The patient succumbed to his disease after 16 hours despite aggressive debridement. (B) This patient presented with spreading cellulites and pain on motion of his right hip 2 weeks after total colectomy. Cellulitis on right anterior thigh is outlined. (C) Classic dishwater edema of tissues with necrotic fascia. (D) Right lower extremity after debridement of fascia to viable muscle.
Patients presenting with sepsis should receive resuscitation fluids early in the course of therapy. While former guidelines advocated fluids until the patient’s central venous pressure was 8 to 12 mmHg, newer guidelines recommend using dynamic monitoring systems (such as ultrasound) as well as assessment of physiological response to fluids by evaluating variables such as heart rate, blood pressure, and urine output to determine adequate resuscitation volumes. Resuscitation endpoints include achieving a goal mean arterial pressure of ≥65 mmHg, urine output of ≥0.5 mL/kg per hour, and normalization of serum lactate. Delaying this resuscitative step for as little as 3 hours has been shown to result in worse outcomes.\(^7\) Resuscitation may necessitate placement of a central venous catheter.

A number of studies have demonstrated the importance of early empiric antibiotic therapy in patients who develop sepsis or nosocomial infection; the Surviving Sepsis guidelines advocate for initiation of treatment within the first hour of the patient’s care. This therapy should be initiated as soon as possible with broad-spectrum antibiotics directed against the most likely organisms. Use of institution- and unit-specific sensitivity patterns are critical in selecting an appropriate agent for patients with nosocomial infection. Obtain appropriate cultures before

---

### Summary of Surviving Sepsis Campaign guidelines

**Initial Evaluation and Infection Issues**

**Initial resuscitation:** Begin resuscitation immediately in patients with hypotension or elevated serum lactate with resuscitation goal of at least 30 mL/kg IV crystalloid given in the first 3 hours. Ongoing fluid administration should be guided by physiologic response as measured by clinical variables (e.g., heart rate, blood pressure, urine output) and/or other invasive or noninvasive monitoring. Resuscitation goals include mean arterial pressure >65 mmHg, urine output >0.5 mL/kg per h, and mixed venous oxygen saturation >65%.

Target resuscitation to normalize lactate in patients with elevated lactate levels.

**Diagnosis:** Obtain appropriate cultures prior to antibiotics, but do not delay antibiotic therapy. Imaging studies should be performed promptly to confirm a source of infection.

**Antibiotic therapy:** Begin IV antibiotic therapy as early as possible and within the first hour after recognition of severe sepsis/septic shock. Use broad spectrum antibiotic regimen with penetration into presumed source, reassess regimen daily with de-escalation as appropriate, discontinue antibiotics in 7 to 10 days for most infections, stop antibiotics for noninfectious issues. Consider the use of serial procalcitonin levels, which may allow earlier cessation of antibiotic therapy.

**Source control:** Establish anatomic site of infection as rapidly as possible; implement source control measures as soon as possible after initial resuscitation. Remove intravascular access devices if potentially infected.

### Hemodynamic Support and Adjunctive Therapy

**Fluid therapy:** Fluid resuscitate using crystalloid, with continued fluid challenges so long as hemodynamic parameters continue to improve (i.e., for so long as the patient remains fluid-responsive). Albumin may be used as an adjunct if large volumes of crystalloid are required, but hydroxyethyl starch and gelatin-based fluids should not be used.

**Vasopressors/Inotropics Therapy:** Maintain MAP of >65 mmHg. Centrally-administered norepinephrine is the first-line choice. Add vasopressin if needed to raise MAP or to reduce norepinephrine requirement. Epinephrine is an alternative to vasopressin but has greater risk of reduced splanchnic blood flow. Dopamine is an appropriate alternative only in select patients (bradycardia, low risk of arrhythmia), and there is no role for low-dose “renal protection” dopamine. Phenylephrine is not recommended. Insert arterial catheters for patients requiring vasopressors. Consider dobutamine infusion for persistent hypotension after appropriate resuscitation and use of vasopressor agents.

**Steroids:** Consider intravenous hydrocortisone (dose <300 mg/day) for adult septic shock when hypotension responds poorly to fluids and vasopressors.

### Other Supportive Therapy

**Blood product administration:** Transfuse red blood cells when hemoglobin decreases to <7.0 g/dL in the absence of extenuating circumstances (e.g., myocardial ischemia, hemorrhage). It is not necessary to use fresh frozen plasma to correct INR abnormalities in the absence of bleeding. Consider prophylactic transfusion of platelets when counts are less than 10,000/mL in the absence of bleeding, <20,000/mL if there is a risk of bleeding, and <50,000 in the setting of active bleeding or need for procedure.

**Mechanical ventilation:** Target an initial tidal volume of 6 mL/kg body weight and plateau pressure of <30 cm H\(_2\)O in patients with acute lung injury. Use PEEP to avoid lung collapse. Adopt a conservative fluid strategy. In the setting of sepsis-induced ARDS with PaO\(_2/\)FiO\(_2\) ratio <150, use prone ventilation over continued supine position or high-frequency oscillatory ventilation. Use a weaning protocol to evaluate the potential for discontinuing mechanical ventilation. Pulmonary artery catheter placement is not indicated for routine monitoring.

**Sedation:** Minimize sedation using specific titration endpoints.

**Glucose control:** Use protocolized approach to blood glucose management targeting upper blood glucose target of 180 mg/dL.

**Prophylaxis:** Use stress ulcer (proton pump inhibitor or H\(_2\) blocker) and deep venous thrombosis (low-dose unfractionated or fractionated heparin) prophylaxis.

**Limitation of support:** Discuss advance care planning with patients and families and set realistic expectations.

starting antibiotics so that appropriate de-escalation of therapy can take place when results return, but only if doing so does not delay the initiation of treatment.

In patients who require vasopressor therapy, the first-line agent should be norepinephrine. This can be augmented with vasopressin, if needed, to achieve MAP ≥65 mmHg. It is important to titrate therapy based on other parameters such as mixed venous oxygen saturation and plasma lactate levels to reduce the risk of vasopressor-induced perfusion deficits. Patients who have persistently poor perfusion despite adequate fluid resuscitation may require addition of inotropic agents (epinephrine, dobutamine) or adjunctive therapy with low-dose corticosteroids (hydrocortisone 200 mg/day).86

Patients with acute lung injury associated with sepsis should receive mechanical ventilation with tidal volumes of 6 mL/kg and pulmonary airway plateau pressures of ≤30 cm H₂O. Finally, red blood cell transfusion should be reserved for patients with hemoglobin of <7 g/dL, with a more liberal transfusion strategy reserved for those patients with severe coronary artery disease, ongoing blood loss, or severe hypoxemia.86

**Resistant Organisms**

Penicillin was first available for widespread clinical use in the 1940s, and within a year resistant strains of *S aureus* had emerged. There are two major factors responsible for antibiotic resistance. First, there may be a genetic component innate to an organism that prevents the effect of a particular antibiotic. For instance, if an organism does not have a target receptor specific to the mechanism of action of a particular antibiotic, the antibiotic will not be effective against this organism. A good example is penicillin and gram-negative organisms, as these microbes lack penicillin-binding proteins. The second component driving resistance is inducible and related to natural selection. Over generations of exposure to a particular antibiotic, selection pressure will drive proliferation of more organisms resistant to that antibiotic. This acquired antibiotic resistance can be mutational, leading to changes in the chromosomal makeup of the microbe, or it can be extrachromosomal, induced by transfer of exogenous genetic material in the form of a plasmid or transposon. In either case, cellular mechanisms of resistance that develop include target site modification, changes in bacterial permeability or antibiotic uptake, activation of drug efflux systems, and drug deactivation. Given that millions of kilograms of antibiotics are used annually in people, in agriculture, and for animal use, environmental selection pressures are high, and antibiotic resistance has now been described in all classes of antibiotics in common use. Antibiotic resistance comes at a high cost, with a significant increase in mortality associated with infection from resistant organisms, and an economic cost of billions of dollars per year.

There are several drug-resistant organisms of interest to the surgeon. MRSA most commonly occurs as a hospital-associated infection in chronically ill patients who have received multiple courses of antibiotics. However, strains of MRSA have emerged in the community among patients without preexisting risk factors for disease.73 These strains, which produce a toxin known as Panton-Valentine leukocidin, make up an increasingly high percentage of surgical site infections since they are resistant to commonly employed prophylactic antimicrobial agents.88 Extended spectrum β-lactamase (ESBL)-producing strains of *Enterobacteriaceae*, originally geographically localized and infrequent, have become much more widespread and common in the last decade.86 These strains, typically *Klebsiella* species or *E coli*, produce a plasmid-mediated inducible β-lactamase. Commonly encountered plasmids also confer resistance to many other antibiotic classes. A common laboratory finding with ESBL is sensitivity to first-, second-, or third-generation cephalosporins, with resistance to other agents. Unfortunately, use of this seemingly active agent leads to rapid induction of resistance and failure of antibiotic therapy. The appropriate antibiotic choice in this setting is a carbapenem.

While *Enterococcus* was considered a low-virulence organism in the past, infections caused by *E faecium* and *faecalis* have been found to be increasingly severe, especially in the immunocompromised host. The last decade has seen increased isolation of a vancomycin-resistant strain of *Enterococcus*. This resistance is transposon-mediated via the vanA gene and is typically seen in *E faecium* strains. A real infection control concern is potential for transfer of genetic material to *S aureus* in a host coinfected with both organisms. This is thought to be the mechanism behind emerging cases of vancomycin resistance in *S aureus*.86

**Blood-Borne Pathogens**

The risk of human immunodeficiency virus (HIV) transmission from patient to surgeon is low. As of May 2011, there had been six cases of surgeons with HIV seroconversion from a possible occupational exposure, with no new cases reported since 1999. Of the numbers of healthcare workers with likely occupationally acquired HIV infection (*n* = 200), surgeons were one of the lower risk groups (compared to nurses at 60 cases and nonsurgeon physicians at 19 cases).90 The estimated risk of transmission from a needlestick from a source with HIV-infected blood is estimated at 0.3%. Transmission of HIV (and other infections spread by blood and body fluid) from patient to healthcare worker can be minimized by observation of universal precautions, including: (a) routine use of barriers (gloves, gown, mask, eye protection) when anticipating contact with blood or body fluids, (b) washing hands and other skin surfaces immediately after contact with blood or body fluids, and (c) careful handling and disposal of sharp instruments during and after use.

Postexposure prophylaxis for HIV has significantly decreased the risk of seroconversion for healthcare workers with occupational exposure to HIV. Steps to initiate postexposure prophylaxis should be initiated within hours for the most effective preventive therapy. Postexposure prophylaxis with a three-drug regimen should be initiated for healthcare workers with significant exposure to patients with an HIV-positive status. If a patient’s HIV status is unknown, it may be advisable to begin postexposure prophylaxis while testing is carried out, particularly if the patient is at high risk for infection due to HIV (e.g., has had a history of intravenous drug use). Generally, postexposure prophylaxis is not warranted for exposure to sources with unknown status, such as deceased persons or needles from a sharps container.92

The risks of acquiring HIV infection for surgeons are related to the prevalence of HIV infection in the patient population, the probability of transmission from a percutaneous injury suffered while caring for an infected patient, the number of such injuries sustained, and the use of postexposure prophylaxis. Average risk of HIV seroconversion is 0.3% from a percutaneous exposure, and 0.09% from a mucous membrane exposure. The overall risk is influenced by the degree of viral inoculum.
transmitted from patient to surgeon, with greater risk of seroconversion associated with hollow-bore needle injury, with larger-volume blood transmission, with direct introduction of infected blood into an artery or vein, and in exposure to blood with higher viral load. One study in Glasgow, Scotland, calculated annual risks and found a range in seroconversion rates from 1 in 200,000 for general surgeons not utilizing postexposure prophylaxis to as low as 1 in 10,000,000 with use of routine postexposure prophylaxis after significant exposures.92,93

Hepatitis B virus (HBV) is a DNA virus that affects only humans. Primary infection with HBV generally is self-limited, but it can cause fulminant hepatitis or progress to a chronic carrier state. Death from chronic liver disease or hepatocellular cancer occurs in roughly 30% of chronically infected persons. Surgeons and other healthcare workers are at high risk for this blood-borne infection and should receive the HBV vaccine; children are routinely vaccinated in the United States.94 This vaccine has contributed to a significant decline in the number of new cases of HBV per year in the United States, from approximately 250,000 annually in the 1980s to 3350 in 2010.95,96

Hepatitis C virus (HCV), previously known as non-A, non-B hepatitis, is a RNA flavivirus first identified in the late 1980s. This virus is confined to humans and chimpanzees. A chronic carrier state develops in 75% to 80% of patients with the infection, with chronic liver disease occurring in three-fourths of this subgroup. The number of new infections per year has declined since the 1980s due to routine testing of blood donors of this subgroup. The number of new infections per year has declined since the 1980s due to routine testing of blood donors.

Fortunately, HCV is not transmitted efficiently through occupational exposures to blood, with the seroconversion rate after accidental needlestick approximately 1.8%.97 To date, a vaccine to prevent HCV infection has not been developed. Experimental studies in chimpanzees with HCV immunoglobulin using a model of needlestick injury have failed to demonstrate a protective effect, and no effective antiviral agents for postexposure prophylaxis are available. Treatment of patients with HCV infection historically included ribavirin and pegylated gamma interferon; the development of novel direct-acting antiviral agents such as sofosbuvir, boceprevir, and telaprevir has led to changes in this strategy.98,99

**BIOLOGIC WARFARE AGENTS**

Several infectious organisms have been studied by the United States and the former Soviet Union and presumably other entities for potential use as biologic weapons. Programs involving biologic agents in the United States were halted by presidential decree in 1971. However, concern remains that these agents could be used by rogue states or terrorist organizations as weapons of mass destruction, as they are relatively inexpensive to make in terms of infrastructure development.

Given these concerns, physicians, including surgeons, should familiarize themselves with the manifestations of infection due to these pathogens. The typical agent is selected for the ability to be spread via the inhalational route, as this is the most efficient mode of mass exposure. Several potential agents are discussed in the following sections.

**Bacillus anthracis (Anthrax)**

Anthrax is a zoonotic disease occurring in domesticated and wild herbivores. The first identification of inhalational anthrax as a disease occurred among woolsorters in England in the late 1800s. The largest recent epidemic of inhalational anthrax occurred in 1979 in Sverdlovsk, Russia, after accidental release of anthrax spores from a military facility. Inhalational anthrax develops after a 1- to 6-day incubation period, with nonspecific symptoms, including malaise, myalgia, and fever. Over a short period of time these symptoms worsen, with development of respiratory distress, chest pain, and diaphoresis. Characteristic chest roentgenographic findings include a widened mediastinum and pleural effusions. Rapid antigen tests are under development for identification of this gram-positive rod, so a key element of establishing the diagnosis is eliciting an exposure history. Postexposure prophylaxis consists of administration of either ciprofloxacin or doxycycline.100 If an isolate is demonstrated to be penicillin-sensitive, the patient should be switched to amoxicillin. Inhalational exposure followed by the development of symptoms is associated with a high mortality rate. Treatment options include combination therapy with ciprofloxacin, clindamycin, and rifampin. Clindamycin is added to block toxin production, while rifampin penetrates into the central nervous system and intracellular locations.

**Yersinia pestis (Plague)**

Plague is caused by the gram-negative organism *Y. pestis*. The naturally occurring disease in humans is transmitted via flea bites from rodents. It was the first biologic warfare agent, and was used in the Crimean city of Caffa by the Tartar army, whose soldiers catapulted bodies of plague victims at the Genoese. When plague is used as a biologic warfare agent, clinical manifestations include epidemic pneumonia with blood-tinged sputum if aerosolized bacteria are used, or bubonic plague if fleas are used as carriers. Individuals who develop a painful enlarged lymph node lesion, termed a “bubo,” associated with fever, severe malaise, and exposure to fleas should be suspected to have plague. Diagnosis is confirmed via aspirate of the bubo and a direct antibody stain to detect plague bacillus, whose morphology is a bipolar, safety-pin-shaped gram-negative rod. Postexposure prophylaxis for patients exposed to plague consists of doxycycline. Treatment of the pneumonic or bubonic/septicemic form includes administration of either streptomycin, an aminoglycoside, doxycycline, a fluoroquinolone, or chloramphenicol.101

**Smallpox**

Variola, the causative agent of smallpox, was a major cause of infectious morbidity and mortality until its eradication in the late 1970s. Even in the absence of laboratory-preserved virus, the prolonged viability of variola virus has been demonstrated in scabs up to 13 years after collection. The potential for reverse genetic engineering using the known sequence of smallpox also makes it a potential biologic weapon. This has resulted in the United States undertaking a vaccination program for key healthcare workers.102 Variola virus is highly infectious in the aerosolized form; after an incubation period of 10 to 12 days, clinical manifestations of malaise, fever, vomiting, and headache appear, followed by development of a characteristic centripetal rash (which is found to predominate on the face and extremities). The fatality rate may reach 30%. Postexposure prophylaxis with smallpox vaccine has been noted to be effective for up to 4 days postexposure. Cidofovir, an acyclic nucleoside phosphonate analogue, has demonstrated activity in animal models of poxvirus infections and may offer promise for the treatment of smallpox.103
Francisella tularensis (Tularemia)

The principal reservoir of this gram-negative aerobic organism is the tick. After inoculation, this organism proliferates within macrophages. Tularemia is considered a potential bioterrorist threat due to a very high infectivity rate after aerosolization. Patients with tularemia pneumonia develop a cough and demonstrate pneumonia on chest roentgenogram. Enlarged lymph nodes occur in approximately 85% of patients. The organism can be cultured from tissue samples, but this is difficult, and the diagnosis is based on acute-phase agglutination tests. Treatment of inhalational tularemia consists of administration of an aminoglycoside or second-line agents such as doxycycline and ciprofloxacin.

REFERENCES

Entries highlighted in bright blue are key references.


This page intentionally left blank
Trauma
Clay Cothren Burlew and Ernest E. Moore

INTRODUCTION
Trauma, or injury, is defined as cellular disruption caused by environmental energy that is beyond the body’s resilience, which is compounded by cell death due to ischemia/reperfusion. Trauma is the most common cause of death for all individuals between the ages of 1 and 44 years, and is the third most common cause of death regardless of age. It is also the leading cause of years of productive life lost. Unintentional injuries account for over 135,000 deaths per year, with homicides, suicides, and other causes are responsible for another 60,000 deaths each year. However, death rate underestimates the magnitude of the societal toll. For example, in 2014 there were almost 200,000 injury-related deaths, but 37.2 million injured patients treated in emergency departments (EDs). Injury-related medical expenditures are estimated to be $117 billion each year in the United States. The aggregate lifetime cost for all injured patients is estimated to be in excess of $260 trillion. For these reasons, trauma must be considered a major public health issue. The American College of Surgeons Committee on Trauma addresses this issue by assisting in the development of trauma centers and systems. The organization of trauma systems has had a significant favorable impact on patient outcomes, although system integration and maldistribution of trauma centers remain challenges.

INITIAL EVALUATION AND RESUSCITATION OF THE INJURED PATIENT
Primary Survey
The Advanced Trauma Life Support (ATLS) course of the American College of Surgeons Committee on Trauma was developed in the late 1970s, based on the premise that appropriate and timely care can improve the outcome for the injured patient. ATLS provides a structured approach to the trauma patient with standard algorithms of care; it emphasizes the “golden hour” concept that timely, prioritized interventions are necessary to prevent death and disability. The ATLS format and basic tenets are followed throughout this chapter, with some modifications. The initial management of seriously injured patients consists of phases that include the primary survey/concurrent resuscitation, the secondary survey/diagnostic evaluation, definitive care, and the tertiary survey. The first step in patient management is performing the primary survey, the goal of which is to identify and treat conditions that constitute an immediate threat to life. The ATLS course refers to the primary survey as assessment of the “ABCs” (Airway, Breathing, and Circulation). The timing of emergent intubation in the hypovolemic patient remains controversial because of the risk of further compromising cardiac function. Although the concepts within the primary survey are presented in a sequential fashion, in reality they are pursued simultaneously in coordinated team resuscitation. Life-threatening injuries must be identified (Table 7-1) and treated before progressing to the secondary survey.

Airway Management With Cervical Spine Protection. Ensuring a patent airway is the first priority in the primary survey. This is essential because efforts to restore cardiovascular integrity will be futile unless the oxygen content of the blood is adequate. Simultaneously, all patients with blunt trauma require cervical spine immobilization until injury is excluded. This is typically accomplished by applying a hard cervical collar or placing sandbags on both sides of the head with the patient’s forehead taped across the bags. Soft collars do not effectively immobilize the cervical spine. For penetrating neck wounds, however, cervical collars are not recommended because they provide no benefit and may interfere with assessment and treatment.
Key Points

1. Trauma is the most common cause of death for all individuals between the ages of 1 and 44 years and is the third most common cause of death regardless of age.

2. The initial management of seriously injured patients usually follows the primary survey (the “ABCs”—Airway, Breathing, and Circulation), although at times restoring Circulatory volume may proceed active Airway intervention; the goals of the primary survey are to identify and treat conditions that constitute an immediate threat to life.

3. All patients with blunt injury should be assumed to have unstable cervical spine injuries until proven otherwise; one must maintain cervical spine precautions and in-line stabilization.

4. Patients with ongoing hemodynamic instability, whether “nonresponders” or “transient responders,” require prompt intervention; one must consider the dominant causes of acute shock, i.e., hemorrhagic, cardiogenic, and neurogenic shock.

5. Patients with trauma-induced coagulopathy (TIC) are at risk for massive transfusion and need to be identified early.

6. Indications for immediate operative intervention for penetrating cervical injury include hemodynamic instability and significant external arterial hemorrhage; the management algorithm for hemodynamically stable patients is based on the presenting symptoms and anatomic location of injury, with the neck being divided into three distinct zones.

7. The gold standard for determining if there is a blunt descending aortic injury is computed tomography angiography (CTA) scanning; indications are primarily based on injury mechanism.

8. The abdomen is a diagnostic black box. Physical examination and FAST ultrasound can identify patients requiring emergent laparotomy. Computed tomography (CT) scanning is the mainstay of evaluation in the remaining patients to more precisely identify the site and magnitude of injury.

9. Manifestation of the “bloody vicious cycle” (the lethal combination of coagulopathy, hypothermia, and metabolic acidosis) is the most common indication for damage control surgery. The primary objectives of damage control laparotomy are to control bleeding and limit gastrointestinal spillage.

10. Blunt injuries to the carotid and vertebral arteries are usually managed with systemic antithrombotic therapy.

11. The abdominal compartment syndrome may be primary (i.e., due to the injury of abdominal organs, bleeding, and packing) or secondary (i.e., due to reperfusion visceral edema, retroperitoneal edema, and ascites).

Patients who are conscious, without tachypnea, and have a normal voice are unlikely to require early airway intervention. Exceptions are penetrating injuries to the neck with an expanding hematoma; evidence of chemical or thermal injury to the mouth, nares, or hypopharynx; extensive subcutaneous air in the neck; complex maxillofacial trauma; or airway bleeding. Although these patients may initially have an adequate airway, it may become compromised if soft tissue swelling, hematoma formation, or edema progresses. In these cases, preemptive intubation should be performed before airway access becomes challenging.

Patients who have an abnormal voice, abnormal breathing sounds, tachypnea, or altered mental status require further airway evaluation. Blood, vomit, the tongue, teeth, foreign objects, and soft tissue swelling can cause airway obstruction; suctioning affords immediate relief in many patients. In the comatose patient, the tongue may fall backward and obstruct the hypopharynx; this can be relieved by either a chin lift or jaw thrust. An oral airway or a nasal trumpet is also helpful in maintaining airway patency, although the former is usually not tolerated by an awake patient. Establishing a definitive airway (i.e., endotracheal intubation) is indicated in patients with apnea; inability to protect the airway due to altered mental status; impending airway compromise due to inhalation injury, hematoma, facial bleeding, soft tissue swelling, or aspiration; and inability to maintain oxygenation. Altered mental status is the most common indication for intubation. Agitation or obtundation, often attributed to intoxication or drug use, may actually be due to hypoxia. But the timing of endotracheal intubation may be critical in the hypovolemic patient because positive airway pressure may further compromise cardiac function and precipitate cardiac arrest; thus, Circulation may take priority over Airway.

Options for endotracheal intubation include nasotracheal, orotracheal, or operative routes. Nasotracheal intubation can be accomplished only in patients who are breathing spontaneously. Although nasotracheal intubation is frequently used by prehospital providers, the technique is limited in the ED to those patients requiring emergent airway support in whom chemical paralysis

---

Table 7-1

<table>
<thead>
<tr>
<th>Airway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airway obstruction</td>
</tr>
<tr>
<td>Airway injury</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Breathing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tension pneumothorax</td>
</tr>
<tr>
<td>Open pneumothorax</td>
</tr>
<tr>
<td>Massive air leak from tracheobronchial injury</td>
</tr>
<tr>
<td>Flail chest with underlying pulmonary contusion</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Circulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemorrhagic shock</td>
</tr>
<tr>
<td>Massive hemorthorax</td>
</tr>
<tr>
<td>Massive hemoperitoneum</td>
</tr>
<tr>
<td>Mechanically unstable pelvis fracture with bleeding</td>
</tr>
<tr>
<td>Extremity blood loss</td>
</tr>
<tr>
<td>Cardiogenic shock</td>
</tr>
<tr>
<td>Cardiac tamponade</td>
</tr>
<tr>
<td>Neurogenic shock</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Disability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intracranial hemorrhage</td>
</tr>
<tr>
<td>Cervical spine injury</td>
</tr>
</tbody>
</table>

---

Immediately life-threatening injuries to be identified during the primary survey

- Airway
  - Airway obstruction
  - Airway injury
- Breathing
  - Tension pneumothorax
  - Open pneumothorax
  - Massive air leak from tracheobronchial injury
  - Flail chest with underlying pulmonary contusion
- Circulation
  - Hemorrhagic shock
  - Massive hemorthorax
  - Massive hemoperitoneum
  - Mechanically unstable pelvis fracture with bleeding
  - Extremity blood loss
  - Cardiogenic shock
  - Cardiac tamponade
  - Neurogenic shock
- Disability
  - Intracranial hemorrhage
  - Cervical spine injury
cannot be used. Orotracheal intubation is the preferred technique to establish a definitive airway. Because all patients are presumed to have cervical spine injuries, manual in-line cervical immobilization is essential. Correct endotracheal placement is verified with direct laryngoscopy, capnography, audible bilateral breath sounds, and finally a chest film. The GlideScope®, a video laryngoscope that uses fiberoptics to visualize the vocal cords, is being employed more frequently. Advantages of orotracheal intubation include the direct visualization of the vocal cords, ability to use larger-diameter endotracheal tubes, and applicability to apneic patients. The disadvantage of orotracheal intubation is that conscious patients usually require neuromuscular blockade, which may result in the inability to intubate, aspiration, or medication complications. Those who attempt rapid-sequence induction must be thoroughly familiar with the procedure (see Chapter 13).

Patients in whom attempts at intubation have failed or who are precluded from intubation due to extensive facial injuries require operative establishment of an airway. Cricothyroidotomy (Fig. 7-1) is performed through a generous vertical incision, with sharp division of the subcutaneous tissues. Visualization may be improved by having an assistant retract laterally on the neck incision using retractors. The cricothyroid membrane is verified by digital palpation and opened in a horizontal direction. The airway may be stabilized before incision of the membrane using a tracheostomy hook; the hook should be placed under the thyroid cartilage to elevate the airway. A 6.0 endotracheal tube (maximum diameter in adults) is then advanced through the cricothyroid opening and sutured into place. In patients under the age of 11, cricothyroidotomy is relatively contraindicated due to the risk of subglottic stenosis, and tracheostomy should be performed.

Emergent tracheostomy is indicated in patients with laryngotracheal separation or laryngeal fractures, in whom cricothyroidotomy may cause further damage or result in complete loss of the airway. This procedure is best performed in the operating room (OR) where there is optimal lighting and availability of advanced equipment (e.g., sternal saw). In these cases, often after a “clothesline” injury, direct visualization and instrumentation of the trachea usually is done through the traumatic anterior neck defect or after a generous collar skin incision (Fig. 7-2). If the trachea is completely transected, a nonpenetrating clamp should be placed on the distal aspect to prevent tracheal retraction into the mediastinum; this is particularly important before placement of the endotracheal tube.

Breathing and Ventilation. Once a secure airway is obtained, adequate oxygenation and ventilation must be ensured. All injured patients should receive supplemental oxygen and be monitored by pulse oximetry. The following conditions constitute an immediate threat to life due to inadequate ventilation and should be recognized during the primary survey: tension pneumothorax, open pneumothorax, flail chest with underlying pulmonary contusion, massive hemothorax, and major air leak due to a tracheobronchial injury.

Tension pneumothorax is presumed in any patient manifesting respiratory distress and hypotension in combination with any of the following physical signs: tracheal deviation away from the affected side, lack of or decreased breath sounds on

Figure 7-1. Cricothyroidotomy is recommended for emergent surgical establishment of a patent airway. A vertical skin incision avoids injury to the anterior jugular veins, which are located just lateral to the midline. Hemorrhage from these vessels obscures vision and prolongs the procedure. When a transverse incision is made in the cricothyroid membrane, the blade of the knife should be angled inferiorly to avoid injury to the vocal cords. A. Use of a tracheostomy hook stabilizes the thyroid cartilage and facilitates tube insertion. B. A 6.0 endotracheal tube is inserted after digital confirmation of airway access.

Figure 7-2. A “clothesline” injury can partially or completely transect the anterior neck structures, including the trachea. With complete tracheal transection, the endotracheal tube is placed directly into the distal aperture, with care taken not to push the trachea into the mediastinum.
the affected side, and subcutaneous emphysema on the affected side. Patients may have distended neck veins due to impedance of venous return, but the neck veins may be flat due to concurrent systemic hypovolemia. Tension pneumothorax and simple pneumothorax have similar signs, symptoms, and examination findings, but hypotension qualifies the pneumothorax as a tension pneumothorax. Although immediate needle thoracostomy decompression with a 14-gauge angiocatheter may be indicated in the field, tube thoracostomy in the midaxillary line should be performed immediately in the ED before a chest radiograph is obtained (Fig. 7-3). Recent studies suggest that preferred location for needle decompression may be the fifth intercostal space in the anterior axillary line due to body habitus. In cases of tension pneumothorax, the parenchymal tear in the lung acts as a one-way valve, with each inhalation allowing additional air to accumulate in the pleural space. The normally negative intrapleural pressure becomes positive, which depresses the ipsilateral hemidiaphragm and shifts the mediastinal structures into the contralateral chest. Subsequently, the contralateral lung is compressed and the heart rotates about the superior and inferior vena cava; this decreases venous return and ultimately cardiac output, which culminates in cardiovascular collapse.

An open pneumothorax or “sucking chest wound” occurs with full-thickness loss of the chest wall, permitting free communication between the pleural space and the atmosphere (Fig. 7-4). This compromises ventilation due to equilibration of atmospheric and pleural pressures, which prevents lung inflation and alveolar ventilation, and results in hypoxia and hypercarbia. Complete occlusion of the chest wall defect without a tube thoracostomy may convert an open pneumothorax to a tension pneumothorax. Temporary management of this injury includes covering the wound with an occlusive dressing that is taped on three sides, which allows accumulated air to escape from the pleural space and thus prevents a tension pneumothorax. Repair of the chest wall defect and tube thoracostomy remote from the wound is definitive treatment.

Flail chest occurs when three or more contiguous ribs are fractured in at least two locations. Paradoxical movement of this free-floating segment of chest wall is usually evident in patients with spontaneous ventilation, due to the negative intrapleural pressure of inspiration. The additional work of breathing and chest wall pain caused by the flail segment is rarely sufficient to compromise ventilation. Instead, it is the decreased compliance and increased shunt fraction caused by the associated pulmonary contusion that is the source of acute respiratory failure. Pulmonary contusions often progress during the first 12 hours. Resultant hypoventilation and hypoxemia may require intubation and mechanical ventilation. The patient’s initial chest...
radiograph often underestimates the extent of the pulmonary parenchymal damage (Fig. 7-5); close monitoring and frequent clinical reevaluation are warranted.

Major air leak occurs from tracheobronchial injuries. Type I injuries are those occurring within 2 cm of the carina.11,12 These may not be associated with a pneumothorax due to the envelopment in the mediastinal pleura. Type II injuries are more distal injuries within the tracheobronchial tree and hence manifest with a pneumothorax. Bronchoscopy confirms the extent of the injury and its location, and directs management.

Circulation With Hemorrhage Control. With a secure airway and adequate ventilation established, circulatory status is the next focus. An initial approximation of the patient’s cardiovascular status can be obtained by palpating peripheral pulses. In general, systolic blood pressure (SBP) must be 60 mmHg for the carotid pulse to be palpable, 70 mmHg for the femoral pulse, and 80 mmHg for the radial pulse. Any episode of hypotension (defined as a SBP <90 mmHg) is assumed to be caused by hemorrhage until proven otherwise. Patients with rapid massive blood loss may have paradoxical bradycardia.13 Blood pressure and pulse should be measured at least every 5 minutes in patients with significant blood loss until normal vital sign values are restored.

Intravenous (IV) access for fluid resuscitation and medication administration is obtained with two peripheral catheters, 16-gauge or larger in adults. For patients in whom peripheral angiocatheter access is difficult, intraosseous (IO) needles should be rapidly placed in the proximal humerus or tibia (Fig. 7-6).14,15 All medications administered IV may be administered in a similar dosage intraosseously. Although safe for emergent use, the needle should be removed once alternative access is established to prevent potential osteomyelitis. Blood should be drawn simultaneously for a bedside hemoglobin level and routine trauma laboratory tests. In the seriously injured patient arriving in shock, an arterial blood gas for base deficit (BD), cross-matching for possible blood component (RBC and plasma) transfusion, and a coagulation panel/viscoelastic hemostasis assay (e.g., TEG, ROTEM) should be obtained. In these patients, secondary large bore (7 to 9 Fr) cannulae should be obtained via the femoral or subclavian veins; Cordis introducer catheters are preferred over triple-lumen catheters. In general, initial access in trauma patients is best secured in the groin so that placement of the catheter will not interfere with the performance of other diagnostic and therapeutic thoracic procedures. A rule of thumb to consider for secondary access is placement of femoral access for thoracic trauma and jugular or subclavian access for abdominal trauma. Internal jugular or subclavian catheters provide a more reliable measurement of central venous pressure (CVP), which may be helpful in determining the volume status of the patient and in excluding cardiac tamponade. Saphenous vein cutdowns at the ankle can also provide excellent access (Fig. 7-7). The saphenous vein is reliably found 1 cm anterior and 1 cm superior to the medial malleolus. Standard 14-gauge catheters can be quickly placed, even in an exsanguinating patient with collapsed veins. In severely injured children younger than 6 years of age, the preferred venous access is peripheral intravenous catheters followed by an IO needle. Central venous catheter placement or saphenous vein cutdown may be considered as the third choice of access based upon provider experience. Inadvertent femoral artery cannulation, however, may result in limb-threatening arterial spasm.

External control of any visible hemorrhage should be achieved promptly while circulating volume is restored. For
open wounds with ongoing bleeding, manual compression should be done with a single 4 × 4 gauze and a gloved hand. Covering the wound with excessive dressings may permit ongoing unrecognized blood loss that is hidden underneath the dressing. Blind clamping of bleeding vessels should be avoided because of the risk to adjacent structures, including nerves. This is particularly true for penetrating injuries of the neck, thoracic outlet, and groin, where bleeding may be torrential and arising deep within the wound. In these situations, a gloved finger placed through the wound directly onto the bleeding vessel can apply enough pressure to control active bleeding. The surgeon performing this maneuver must then walk with the patient to the OR for definitive treatment. For bleeding of the extremities, it is tempting to apply tourniquets for hemorrhage control, but digital occlusion will usually control the bleeding; complete vascular occlusion with a tourniquet risks permanent neuromuscular impairment. Patients in shock have a lower tolerance to warm ischemia, and an occluded extremity is prone to small vessel thrombosis. For patients with open fractures, fracture reduction with stabilization via splints will limit bleeding both externally and into the subcutaneous tissues. Scalp lacerations through the galea aponeurotica tend to bleed profusely; these can be temporarily controlled with skin staples, Raney clips, or a full-thickness continuous running nylon stitch.

During the circulation section of the primary survey, four life-threatening injuries must be identified promptly: (a) massive hemothorax, (b) cardiac tamponade, (c) massive hemoperitoneum, and (d) mechanically unstable pelvic fractures with bleeding. Massive hemoperitoneum and mechanically unstable pelvic fractures are discussed in “Emergent Abdominal Exploration” and “Pelvic Fractures and Emergent Hemorrhage Control,” respectively. Critical tools used to differentiate these in the multisystem trauma patient are the chest and pelvis radiographs, and extended focused abdominal sonography for trauma (eFAST) (see “Regional Assessment and Special Diagnostic Tests”). A massive hemothorax (life-threatening injury number one) is defined as >1500 mL of blood or, in the pediatric population, >25% of the patient’s blood volume in the pleural space (Fig. 7-8). Although it may be estimated on chest radiograph, tube thoracostomy is the only reliable means to quantify the amount of hemothorax. After blunt trauma, a hemothorax is usually due to multiple rib fractures with severed intercostal vessels, but occasionally bleeding is from lacerated lung parenchyma, which is usually associated with an air leak. After penetrating trauma, a great vessel or pulmonary hilar vessel injury should be presumed. In either scenario, a massive hemothorax is an indication for operative intervention, but tube thoracostomy is critical to facilitate lung reexpansion, which may improve oxygenation and cardiac performance as well as tamponade venous bleeding. In patients arriving in shock with a high risk of pelvic fracture (e.g., autopedestrian accident), the pelvis should be presumptively stabilized with a sheet or binder.

![Figure 7-7](image-url) Saphenous vein cutdowns are excellent sites for fluid resuscitation access. A. The vein is consistently found 1 cm anterior and 1 cm superior to the medial malleolus. B. Proximal and distal traction sutures are placed with the distal suture ligated. C. A 14-gauge IV catheter is introduced and secured with sutures and tape to prevent dislodgment.

![Figure 7-8](image-url) More than 1500 mL of blood in the pleural space is considered a massive hemothorax. Chest film findings reflect the positioning of the patient. A. In the supine position, blood tracks along the entire posterior section of the chest and is most notable pushing the lung away from the chest wall. B. In the upright position, blood is visible dependently in the right pleural space.
Cardiac tamponade occurs most commonly after penetrating thoracic wounds, although occasionally blunt rupture of the heart, particularly the atrial appendage, is seen. Acutely, <100 mL of pericardial blood may cause pericardial tamponade. The classic Beck’s triad—dilated neck veins, muffled heart tones, and a decline in arterial pressure—is usually not appreciated in the trauma bay because of the noisy environment and associated hypovolemia. Because the pericardium is not acutely distensible, the pressure in the pericardial sac will rise to match that of the injured chamber. When this pressure exceeds that of the right atrium, right atrial filling is impaired and right ventricular preload is reduced. This ultimately leads to decreased right ventricular output. Additionally, increased intrapericardial pressure impedes myocardial blood flow, which leads to subendocardial ischemia and a further reduction in cardiac output.

Diagnosis of hemopericardium is best achieved by ultrasound of the pericardium (Fig. 7-9). Early in the course of tamponade, blood pressure and cardiac output will transiently improve with fluid administration due to increased central venous pressure. In patients with any hemodynamic disturbance, a pericardial drain can be placed using ultrasound guidance (Fig. 7-10). Removing as little as 15 to 20 mL of blood will often temporarily stabilize the patient’s hemodynamic status and alleviate the subendocardial ischemia that can be associated with lethal arrhythmias; this allows safe transport to the OR for sternotomy. Pericardiocentesis is successful in decompressing tamponade in approximately 80% of cases; the majority of failures are due to the presence of clotted blood within the pericardium. Patients with a persistent SBP <60 mmHg warrant resuscitative thoracotomy (RT) with opening of the pericardium for rapid decompression and control of bleeding.

The utility of RT has been debated for decades. Current indications are based on 40 years of prospective data (Table 7-2). RT is associated with the highest survival rate after isolated cardiac injury; 35% of patients presenting in shock and 20% without vital signs (i.e., no pulse or obtainable blood pressure) are salvaged after
isolated penetrating injury to the heart. For all penetrating wounds, survival rate is 15%. Conversely, patient outcome is limited when RT is done for blunt trauma, with 2% survival among patients in shock and <1% survival among those with no vital signs. Thus, patients undergoing cardiopulmonary resuscitation (CPR) upon arrival to the ED should undergo RT selectively based on injury and duration of CPR (Fig. 7-11). RT is best accomplished using a generous left anterolateral thoracotomy, with the skin incision started to the right of the sternum (Fig. 7-12). A longitudinal pericardiotomy anterior to the phrenic nerve is used to release cardiac tamponade and permits access to the heart for cardiac repair and open cardiac massage. The cardiac wound should be repaired prior to vigorous efforts of myocardial resuscitation (e.g., epinephrine, calcium). Cross-clamping of the aorta improves central circulation, augments cerebral and coronary blood flow, and limits further abdominal blood loss (Fig. 7-13). The patient must sustain a SBP of 70 mmHg after RT and associated interventions to be considered resuscitable, and hence transported to the OR.18,19

Disability and Exposure. The Glasgow Coma Scale (GCS) score should be determined for all injured patients (Table 7-3). It is calculated by adding the scores of the best motor response, best verbal response, and the best eye response. Scores range from 3 (the lowest) to 15 (normal). Scores of 13 to 15 indicate mild head injury, 9 to 12 moderate injury, and ≤8 severe injury. The GCS is a quantifiable determination of neurologic function that is useful for triage, treatment, and prognosis.

Neurologic evaluation, including spinal cord integrity, is critical before administration of neuromuscular blockade for intubation. Subtle changes in mental status can be caused by hypoxia, hypercarbia, or hypovolemia, or may be an early sign of increasing intracranial pressure. An abnormal mental status should prompt an immediate reevaluation of the patient’s ABCs and consideration of central nervous system injury. Deterioration in mental status may be subtle and may not progress in a predictable fashion. For example, previously calm, cooperative patients may become anxious and combative as they become hypoxic. However, a patient who is agitated and combative from drugs or alcohol may become somnolent if hypovolemic shock develops. Patients with neurogenic shock are typified by hypotension with relative bradycardia, and are often first recognized due to paralysis, decreased rectal tone, or priapism. Patients with high spinal cord disruption are at greatest risk for neurogenic shock due to physiologic disruption of sympathetic fibers; treatment consists of volume loading and a dopamine infusion, which is both inotropic and chronotropic, as well as a vasoconstrictor.

Seriously injured patients must have all of their clothing removed to avoid overlooking limb- or life-threatening injuries, but warmed blankets should be placed immediately to avoid hypothermia.

Shock Classification and Initial Fluid Resuscitation. Classic signs and symptoms of shock are tachycardia, hypotension, tachypnea, altered mental status, diaphoresis, and pallor

Figure 7-11. Algorithm directing the use of resuscitative thoracotomy (RT) in the injured patient undergoing cardiopulmonary resuscitation (CPR). OR = operating room; SBP = systolic blood pressure.
Figure 7-12. A. Resuscitative thoracotomy (RT) is performed through the fifth intercostal space using the anterolateral approach. B and C. The pericardium is opened anterior to the phrenic nerve, and the heart is rotated out for evaluation. D. Open cardiac massage should be performed with a hinged, clapping motion of the hands, with sequential closing from palms to fingers. The two-handed technique is strongly recommended because the one-handed massage technique poses the risk of myocardial perforation with the thumb.

Figure 7-13. A. Aortic cross-clamp is applied with the left lung retracted superiorly, below the inferior pulmonary ligament, just above the diaphragm. B. The flaccid aorta is identified as the first structure encountered on top of the spine when approached from the left chest.

(Table 7-4). In general, the quantity of acute blood loss correlates with physiologic abnormalities. For example, patients in class II shock are tachycardic, but they do not exhibit a reduction in blood pressure until over 1500 mL of blood loss, or class III shock. Physical findings should be used as an aid in the evaluation of the patient’s response to treatment. The goal of fluid resuscitation is to re-establish tissue perfusion. Fluid resuscitation usually begins with isotonic crystalloid, typically Ringer’s lactate. However, for patients arriving in shock (persistent SBP <90 mmHg in an adult), instead of crystalloid the current practice is to activate a massive transfusion protocol (MTP) in which red blood cells (RBC) and fresh frozen plasma (FFP) are administered. The details of a MTP are discussed later. Patients who have a good response to
fluid infusion (i.e., normalization of vital signs, clearing of the sensorium) and evidence of good peripheral perfusion (warm extremities with normal capillary refill) are presumed to have adequate overall perfusion. Urine output is a reliable indicator of organ perfusion but requires time to quantitate. Adequate urine output is 0.5 mL/kg per hour in an adult, 1 mL/kg per hour in a child, and 2 mL/kg per hour in an infant <1 year of age. Because measurement of this resuscitation-related variable is time dependent, it is generally more useful in the OR and intensive care unit (ICU) setting, than in initial evaluation in the trauma bay.

There are several caveats to be considered when evaluating the injured patient for shock. Tachycardia (HR >110 bpm) is often the earliest sign of ongoing blood loss, but the critical issue is change in HR over time. Individuals in good physical condition with a resting pulse rate in the 50s may manifest a relative tachycardia in the 90s; although clinically significant, this does not meet the standard definition of tachycardia. Conversely, patients on cardiac medications such as β-blockers may not be capable of increasing their heart rate to compensate for hypovolemia. Bradycardia can occur with rapid severe blood loss; this is an ominous sign, often heralding impending cardiovascular collapse. Other physiologic stresses, aside from hypovolemia, may produce tachycardia, such as hypoxia, pain, anxiety, and stimulant drugs (cocaine, amphetamines). As noted previously, decreased SBP is not a reliable early sign of hypovolemia because blood loss must exceed 30% before hypotension is evident. Additionally, younger patients may maintain their SBP due to sympathetic tone despite severe intravascular deficits until they are on the verge of cardiac arrest. Pregnant

<table>
<thead>
<tr>
<th>Eye opening</th>
<th>ADULTS</th>
<th>INFANTS/CHILDREN</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Spontaneous</td>
<td>Spontaneous</td>
</tr>
<tr>
<td>3</td>
<td>To voice</td>
<td>To voice</td>
</tr>
<tr>
<td>2</td>
<td>To pain</td>
<td>To pain</td>
</tr>
<tr>
<td>1</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Verbal</th>
<th>ADULTS</th>
<th>INFANTS/CHILDREN</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Oriented</td>
<td>Alert, normal vocalization</td>
</tr>
<tr>
<td>4</td>
<td>Confused</td>
<td>Cries, but consolable</td>
</tr>
<tr>
<td>3</td>
<td>Inappropriate words</td>
<td>Persistently irritable</td>
</tr>
<tr>
<td>2</td>
<td>Incomprehensible words</td>
<td>Restless, agitated, moaning</td>
</tr>
<tr>
<td>1</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Motor response</th>
<th>ADULTS</th>
<th>INFANTS/CHILDREN</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Obeys commands</td>
<td>Spontaneous, purposeful</td>
</tr>
<tr>
<td>5</td>
<td>Localizes pain</td>
<td>Localizes pain</td>
</tr>
<tr>
<td>4</td>
<td>Withdraws</td>
<td>Withdraws</td>
</tr>
<tr>
<td>3</td>
<td>Abnormal flexion</td>
<td>Abnormal flexion</td>
</tr>
<tr>
<td>2</td>
<td>Abnormal extension</td>
<td>Abnormal extension</td>
</tr>
<tr>
<td>1</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>

\[\text{Score is calculated by adding the scores of the best motor response, best verbal response, and eye opening. Scores range from 3 (the lowest) to 15 (normal).}\]

<table>
<thead>
<tr>
<th>Signs and symptoms of advancing stages of hemorrhagic shock</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood loss (mL)</td>
</tr>
<tr>
<td>Blood loss (% BV)</td>
</tr>
<tr>
<td>Pulse rate</td>
</tr>
<tr>
<td>Blood pressure</td>
</tr>
<tr>
<td>Pulse pressure (mmHg)</td>
</tr>
<tr>
<td>Respiratory rate</td>
</tr>
<tr>
<td>Urine output (mL/h)</td>
</tr>
<tr>
<td>CNS/mental status</td>
</tr>
</tbody>
</table>

BV = blood volume; CNS = central nervous system.
patients have a progressive increase in circulating blood volume over gestation; therefore, they must lose a relatively larger volume of blood before manifesting signs and symptoms of hypovolemia (see “Special Populations”).

Based on the initial response to fluid resuscitation, hypovolemic injured patients can be separated into three broad categories: responders, transient responders, and nonresponders. Individuals who are stable or have a good response to initial fluid therapy as evidenced by normalization of vital signs, mental status, and urine output are unlikely to have significant ongoing hemorrhage, and further diagnostic evaluation for occult injuries can proceed in an orderly fashion (see “Secondary Survey”). At the other end of the spectrum are patients classified as “nonresponders” who have persistent hypotension despite aggressive resuscitation. These patients mandate immediate identification of the source of hypotension with appropriate intervention to prevent a fatal outcome. Transient responders are those who respond initially to volume loading with improvement in vital signs, but subsequently deteriorate hemodynamically. This group of patients can be challenging to triage for definitive management.

**Persistent Hypotension.** Patients with ongoing hemodynamic instability, whether “nonresponders” or “transient responders,” require systematic evaluation and prompt intervention. The spectrum of disease in patients with persistent hypotension ranges from easily reversible problems such as a tension pneumothorax to multisystem injury with a number of sites of ongoing active hemorrhage. One must first consider the dominant categories of postinjury shock that may be the underlying cause: hemorrhagic, cardiogenic, and neurogenic. In patients with persistent hypotension and tachycardia, cardiogenic or hemorrhagic shock are the likely causes. Ultrasound evaluation of the pericardium, pleural cavities, and abdomen in combination with plain radiographs of the chest and pelvis will usually identify the source of shock. In patients with persistent hypotension following blunt trauma, the pelvis should be wrapped with a sheet for stabilization until radiography can be done; external blood loss should be controlled, and extremity fractures should be splinted to minimize further blood loss. Evaluation of the CVP or ultrasound of the IVC may further assist in distinguishing between cardiogenic and hypovolemic shock. Base deficit measurement is critical; a base deficit of >8 mmol/L implies ongoing cellular shock. Serum lactate also is used to monitor the patient’s physiologic response to resuscitation. Evolving technology, such as measurement of the critical reserve index, may provide noninvasive monitoring.

The differential diagnosis of cardiogenic shock in the trauma patient is: (a) tension pneumothorax, (b) pericardial tamponade, (c) blunt cardiac injury, (d) myocardial infarction, and (e) bronchovenous air embolism. Tension pneumothorax, the most frequent cause of cardiac failure, and pericardial tamponade have been discussed earlier. Although as many as one-third of patients sustaining significant blunt chest trauma experience some degree of blunt cardiac injury (BCI), few such injuries result in hemodynamic embarrassment. Patients with electrocardiographic (ECG) abnormalities or dysrhythmias require continuous ECG monitoring and antidyssrhythmic treatment as indicated. Unless myocardial infarction is suspected, there is no role for routine serial measurement of cardiac enzyme levels—they lack specificity and do not predict significant dysrhythmias. In patients with no identified injuries who are being considered for discharge from the ED, the combination of a normal EKG and troponin level rules out significant BCI. The patient with hemodynamic instability from BCI requires appropriate resuscitation and may benefit from hemodynamic monitoring to optimize preload and guide inotropic support. Echocardiography (ECHO) is performed to exclude valvular or septal injuries; the most common ECHO finding in BCI is right ventricular dyskinesia due to the anterior orientation of the right ventricle. Transsthoracic and transesophageal ECHO are now becoming routine in the trauma bay and surgical intensive care unit (SICU). Rarely, patients with refractory cardiogenic shock may require placement of an intra-aortic balloon pump to enhance coronary perfusion and decrease myocardial work. Acute myocardial infarction may be the cause of a motor vehicle collision or other trauma in older patients. Although optimal initial management includes treatment for the evolving infarction, such as thrombolytic therapy, anticoagulation, and emergent angioplasty, these decisions must be individualized in accordance with the patient’s other injuries.

Bronchovenous air embolism is a frequently overlooked lethal complication of pulmonary injury. Air emboli can occur after blunt or penetrating trauma, when air from an injured bronchus enters an adjacent injured pulmonary vein and returns air to the left heart. Air accumulation in the left ventricle impedes diastolic filling, and during systole air is pumped into the coronary arteries, disrupting coronary perfusion. The typical case is a patient with a penetrating thoracic injury who is hemodynamically stable but experiences sudden cardiac arrest after being intubated and placed on positive pressure ventilation. The patient should be placed immediately in Trendelenburg’s position to trap the air in the apex of the left ventricle. Emergency thoracotomy is followed by cross-clamping of the pulmonary hilum on the side of the injury to prevent further introduction of air (Fig. 7-14). Air is aspirated from the apex of the left ventricle and then the aortic root with an 18-gauge needle and 50-mL syringe. Vigorous massage is used to force the air bubbles through the coronary arteries; if this is unsuccessful, a tuberculin syringe is used to aspirate air bubbles from the coronary artery (most common is the right). Once circulation is restored, the patient should be kept in Trendelenburg’s position with the pulmonary hilum clamped until the pulmonary venous injury is controlled operatively.

Persistent hypotension due to uncontrolled hemorrhage is associated with high mortality. A rapid search for the source or sources of hemorrhage includes visual inspection with knowledge of the injury mechanism, eFAST, and chest and pelvic radiographs. During diagnostic evaluation, type O RBCs (O-negative for women of childbearing age) and thawed AB plasma should be administered at a ratio of 2:1. Type-specific RBCs should be administered as soon as available. The acute coagulopathy of trauma is now well recognized and underscores the importance of preemptive blood component administration. The resurgent interest in viscoelastic hemostatic assays (thrombelastography [TEG] and thrombelastometry [ROTEM]) has facilitated the appropriate and timely use of clotting adjuncts, including the prompt recognition of fibrinolysis. In patients with clear indications for operation, essential radiographs should be taken, and the patient should be transported to the OR immediately. Such patients include those with blunt trauma and massive hemorrhage, those with penetrating trauma and an initial chest tube output of >1 L, and those with abdominal trauma and ultrasound evidence of extensive hemoperitoneum. In patients with gunshot wounds to the chest or abdomen, a chest and
abdominal film, with radiopaque markers at the wound sites, should be obtained to determine the trajectory of the bullet or location of a retained fragment prior to transport to the OR. For example, a patient with a gunshot wound to the upper abdomen should have a chest radiograph to ensure that the bullet did not traverse the diaphragm causing intrathoracic injury. Similarly, a chest radiograph is important in a patient with a gunshot wound to the right chest to evaluate the left hemithorax. If a patient arrives with a penetrating weapon remaining in place, the weapon should not be removed in the ED because it could be tamponading a lacerated blood vessel (Fig. 7-15). The surgeon should extract the offending instrument in the controlled environment of the OR, ideally once an incision has been made with adequate exposure for vascular control. In situations where knives are embedded in the head or neck, preoperative imaging may be useful to anticipate arterial injuries.

In patients with persistent hypotension and no clear operative indications, one should systematically evaluate the five potential sources of blood loss: scalp, chest, abdomen, pelvis, and extremities. Significant bleeding at the scene may be reported by paramedics, but its quantification is unreliable. Examination should exclude active bleeding from a scalp laceration that may be readily controlled with clips or staples. Thoracoabdominal trauma should be evaluated with a combination of eFAST, chest radiograph, and pelvic radiograph. If the FAST results are negative and no other source of hypotension is obvious, diagnostic peritoneal aspiration should be entertained. Extremity examination and radiographs should be used to identify associated fractures. Fracture-related blood loss, when additive, may be a potential source of the patient’s hemodynamic instability. Each rib fracture can produce 100 to 200 mL of blood loss; for tibial fractures, 300 to 500 mL; for

Figure 7-14. A Satinsky clamp is used to clamp the pulmonary hilum to prevent further bronchovenous air embolism. B Sequential sites of aspiration include the left ventricle, the aortic root, and the right coronary artery.

Figure 7-15. If a weapon is still in place, it should be removed in the operating room because it could be tamponading a lacerated blood vessel.
femur fractures, 800 to 1000 mL; and for pelvic fractures >2000 mL. Although no single injury can account for the patient’s hemodynamic instability, the sum of the injuries may result in life-threatening blood loss. The diagnostic measures advocated earlier are those that can be easily performed in the trauma bay. Transport of a hypotensive patient out of the ED for CT scanning is hazardous; monitoring is compromised, and the environment is suboptimal for dealing with acute problems. Fast track CT scanning should be used in all patients manifesting evidence of shock. The surgeon must accompany the patient and be prepared to abort the CT scan with diversion to the OR. This dilemma is becoming less common in many trauma centers where CT scanning is available in the ED.

The concept of hypotensive resuscitation remains controversial, and it is primarily relevant for patients with penetrating vascular injuries. Experimental work suggests that an endogenous sealing clot of an injured artery may be disrupted at a SBP of >90 mmHg; thus, many believe that this should be the preoperative blood pressure target for patients with potential noncompressible arterial injuries. On the other hand, optimal management of traumatic brain injury (TBI) includes maintaining the SBP >100 mmHg and thus, hypotensive resuscitation is not appropriate for most blunt trauma patients.

Secondary Survey

Once the immediate threats to life have been addressed, a thorough history is obtained, and the patient is examined in a systematic fashion. The patient and surrogates should be queried to obtain an AMPLE history (Allergies, Medications, Past illnesses or Pregnancy, Last meal, and Events related to the injury). The physical examination should be literally head to toe, with special attention to the patient’s back, axillae, and perineum, because injuries here are easily overlooked. All seriously injured patients should undergo digital rectal examination to evaluate for sphincter tone, presence of blood, rectal perforation, or a high-riding prostate; this is particularly critical in patients with suspected spinal cord injury, pelvic fracture, or transpelvic gunshot wounds. Vaginal examination with a speculum should be performed in women with pelvic fractures to exclude an open fracture. Specific injuries, their associated signs and symptoms, diagnostic options, and treatments are discussed in detail later in this chapter.

Adjuncts to the physical examination include vital sign and ECG monitoring, nasogastric tube placement, Foley catheter placement, radiographs, hemoglobin, base deficit measurements, urinalysis, and repeat FAST exam. A nasogastric tube should be inserted in all intubated patients to decrease the risk of gastric aspiration, but it may not be necessary in the awake patient. Nasogastric tube placement in patients with complex mid-facial fractures is contraindicated; rather, a tube should be placed orally if required. Nasogastric tube evaluation of stomach contents for blood may suggest occult gastroduodenal injury, or the errant path of the nasogastric tube on a chest film may indicate a left diaphragm injury. A Foley catheter should be inserted in patients unable to void to decompress the bladder, obtain a urine specimen, and monitor urine output. Gross hematuria demands evaluation of the genitourinary system for injury. Foley catheter placement may be deferred until urologic evaluation in patients with signs of urethral injury: blood at the meatus, perineal or scrotal hematomas, or a high-riding prostate. Although policies vary at individual institutions, most agree that patients in extremis should undergo one attempt at Foley catheter placement; if the catheter does not pass easily, a suprapubic cystostomy should be considered.

Selective radiography and laboratory tests are done early in the evaluation after the primary survey. For patients with severe blunt trauma, chest and pelvic radiographs should be obtained. Historically, a lateral cervical spine radiograph was also obtained, hence the reference to the big three films, but currently patients preferentially undergo CT scanning of the spine rather than plain film radiography. For patients with truncal gunshot wounds, anteroposterior and at times lateral radiographs of the chest and abdomen are warranted. It is important to mark the entrance and exit sites of penetrating wounds with ECG pads, metallic clips, or staples so that the trajectory of the missile can be estimated. Limited one-shot extremity radiographs may also be appropriate to assist in application of a splint. In critically injured patients, blood samples for a routine trauma panel (type and crossmatch, complete blood count, blood chemistries, coagulation studies, and arterial blood gas analysis) should be sent to the laboratory. For less severely injured patients, only a complete blood count and urinalysis may be required. Because older patients may present in subclinical shock, even with minor injuries, routine analysis of an arterial blood gas in patients over the age of 55 should be considered. Repeat FAST is mandatory if there are any signs of abdominal injury or unexplained blood loss.

Many trauma patients cannot provide specific information about the mechanism of their injury. Emergency medical service personnel and police are trained to evaluate an injury scene and should be questioned while they are present in the ED. For automobile collisions, the speed of the vehicles involved, angle of impact, location of the patient within the vehicle, use of restraints, airbag deployment, condition of the steering wheel and windshield, amount of intrusion, ejection of the patient from the vehicle, and fate of other passengers should be ascertained. For other injury mechanisms, critical information includes such things as height of a fall, surface impact, helmet use, and weight of an object by which the patient was crushed. In patients sustaining gunshot wounds, bullet characteristics, distance, and presumed path of the bullet are important, if known. For patients with stab wounds, the length and type of object is helpful. Finally, some patients experience a combination of blunt and penetrating trauma. Do not assume that someone who was stabbed was not also assaulted; the patient may have a multitude of injuries and cannot be presumed to have only injuries associated with the more obvious penetrating wound. In short, these details of information are critical to the clinician to determine overall mechanism of injury and anticipate associated injury patterns.

Mechanisms and Patterns of Injury

In general, more energy is transferred over a wider area during blunt trauma than from a penetrating wound. As a result, blunt trauma is associated with multiple widely distributed injuries, whereas in penetrating wounds the damage is localized to the path of the bullet or knife. In blunt trauma, organs that cannot yield to impact by elastic deformation are most likely to be injured, namely, the solid organs (liver, spleen, and kidneys). For penetrating trauma, organs with the largest surface area are most prone to injury (small bowel, liver, and colon). Additionally, because bullets and knives usually follow straight lines, adjacent structures are commonly injured.

Patients who have sustained blunt trauma separate them into categories according to their risk for multiple injuries: those
sustaining high energy transfer injuries and those sustaining low energy transfer injuries. Injuries involving high energy transfer include auto-pedestrian accidents, motor vehicle collisions in which the car’s change of velocity (ΔV) exceeds 20 mph or in which the patient has been ejected, motorcycle collisions, and falls from heights >20 ft. In fact, for motor vehicle collisions the variables strongly associated with life-threatening injuries, and hence reflective of the magnitude of the mechanism, are death of another occupant in the vehicle, extrication time of >20 minutes, ΔV >20 mph, lack of restraint use, and lateral impact. Low-energy trauma, such as being struck with a club or falling from a bicycle, usually does not result in widely distributed injuries. However, potentially lethal injuries of internal organs can occur because the net energy transfer to any given location may be substantial.

In blunt trauma, particular constellations of injury or injury patterns are associated with specific injury mechanisms. For example, when an unrestrained driver sustains a frontal impact, the head strikes the windshield, the chest and upper abdomen hit the steering column, and the legs or knees contact the dashboard. The resultant injuries can include facial fractures, cervical spine fractures, injury of the descending thoracic aorta, myocardial contusion, injury to the spleen and liver, and fractures of the pelvis and lower extremities. When such patients are evaluated, the discovery of one of these injuries should prompt a search for the others. Collisions with side impact also carry the risk of cervical spine and thoracic trauma, diaphragm rupture, and crush injuries of the pelvic ring, but solid organ injury usually is limited to either the liver or spleen based on the direction of impact. However, patients ejected from a vehicle or thrown a significant distance from a motorcycle have the risk of any injury pattern.

Penetrating injuries are classified according to the wounding agent (i.e., stab wound, gunshot wound, or shotgun wound). Gunshot wounds are subdivided further into high- and low-velocity injuries because the speed of the bullet is much more important than its weight in determining potential kinetic energy transfer. High-velocity gunshot wounds (bullet speed >2000 ft/s) are infrequent in the civilian setting. Shotgun injuries are divided into close-range (<20 feet) and long-range wounds. Close-range shotgun wounds are tantamount to high-velocity wounds because the entire energy of the load is delivered to a small area, often with devastating results. In contrast, long-range shotgun blasts result in a diffuse pellet pattern in which many pellets miss the victim, and those that do strike are dispersed and are of comparatively low energy.

Regional Assessment and Special Diagnostic Tests

Based on mechanism, location of injuries identified on physical examination, screening radiographs, and the patient’s overall condition, additional diagnostic studies often are indicated. However, the seriously injured patient is in constant jeopardy when undergoing special diagnostic testing; therefore, the surgeon must be in attendance and must be prepared to alter plans as circumstances demand. Hemodynamic, respiratory, and mental status will determine the most appropriate course of action. With these issues in mind, additional diagnostic tests are discussed on an anatomic basis.

Head. Evaluation of the head includes examination for injuries to the scalp, eyes, ears, nose, mouth, facial bones, and intracranial structures. Palpation of the head is done to identify scalp lacerations, which should be evaluated for depth, and presence of associated depressed or open skull fractures. The eye examination includes not only pupillary size and reactivity, but also examination for visual acuity and for hemorrhage within the globe. Ocular entrapment, caused by orbital fractures with impingement of the ocular muscles, is evident when the patient cannot move his or her eyes through an entire range of motion. It is important to perform the eye examination early because significant orbital swelling may prevent later evaluation. A lateral canthotomy may be needed to relieve periorbital pressure. The tympanic membrane is examined to identify hematotympanum, otorrhea, or rupture, which may signal an underlying head injury. Otorrhea, rhinorrhea, raccoon eyes, and Battle’s sign (ecchymosis behind the ear) suggest a basilar skull fracture. Although such fractures may not require treatment, there is an association with blunt cerebrovascular injuries, cranial nerve injuries, and risk of meningitis.

Anterior facial structures should be examined to rule out fractures. This entails palpating for bony step-off of the facial bones and instability of the midface (by grasping the upper palate and seeing if this moves separately from the patient’s head). A good question to ask awake patients is whether their bite feels normal to them; abnormal dental closure suggests malalignment of facial bones and the possibility for a mandible or maxillary fracture. Nasal fractures, which may be evident on direct inspection or palpation, typically bleed vigorously. This may result in the patient having airway compromise due to blood running down the posterior pharynx, or there may be vomiting provoked by swallowed blood. Nasal packing or balloon tamponade may be necessary to control bleeding. Examination of the oral cavity includes inspection for open fractures, loose or fractured teeth, and sublingual hematomas.

All patients with a significant closed head injury (GCS score <14) should undergo CT scanning of the head. Additionally, elderly patients or those patients on antiplatelet agents or anticoagulation should be imaged despite a GCS of 15. For penetrating injuries, plain skull films may be helpful in the trauma bay to determine the trajectory of the bullet. The presence of lateralizing findings (e.g., a unilateral dilated pupil unreactive to light, asymmetric movement of the extremities either spontaneously or in response to noxious stimuli, or unilateral Babinski’s reflex) suggests an intracranial mass lesion or major structural damage.

Such intracranial lesions following trauma include hematomas, contusions, hemorrhage into ventricular and subarachnoid spaces, and diffuse axonal injury (DAI). Epidural hematomas occur when blood accumulates between the skull and dura, and are caused by disruption of the middle meningeal artery or other small arteries in that potential space, typically after a skull fracture (Fig. 7-16). Subdural hematomas occur between the dura and cortex and are caused by venous disruption or laceration of the parenchyma of the brain. Due to associated parenchymal injury, subdural hematomas have a much worse prognosis than epidural collections. Hemorrhage into the subarachnoid space may cause vasospasm and further reduce cerebral blood flow. Intraparenchymal hematomas and contusions can occur anywhere within the brain. DAI results from high-speed deceleration injury and represents direct axonal damage from shear effects. CT scan may demonstrate blurring of the gray and white matter interface and multiple small punctate hemorrhages, but magnetic resonance imaging is a more accurate test. Although prognosis for these injuries is extremely variable, early evidence of DAI is associated with a poor outcome. Stroke syndromes should prompt a search for carotid or vertebral artery injury using multislice CTA (Fig. 7-17).
Significant intracranial penetrating injuries usually are produced by bullets from handguns, but an array of other weapons or instruments can injure the cerebrum via the orbit or through the thinner temporal region of the skull. Although the diagnosis usually is obvious, in some instances wounds in the auditory canal, mouth, and nose can be elusive. Prognosis is variable, but virtually all supratentorial wounds that injure both hemispheres are fatal.

**Neck.** All blunt trauma patients should be assumed to have cervical spine injuries until proven otherwise. During the physical examination, one must maintain cervical spine precautions and in-line stabilization. Due to the devastating consequences of quadriplegia, a diligent evaluation for occult cervical spine injuries is mandatory. In the awake patient, the presence of posterior midline pain or tenderness should provoke a thorough radiologic evaluation. Additionally, intubated patients, patients with distracting injuries, significant mechanism, or another identified spine fracture should undergo CT imaging. A ligamentous injury may not be visible with standard imaging techniques. Flexion and extension views or magnetic resonance imaging (MRI) are obtained to further evaluate patients at risk or those with persistent symptoms.

Spinal cord injuries can vary in severity. Complete injuries cause either quadriplegia or paraplegia, depending on the level of injury. These patients have a complete loss of motor function and sensation two or more levels below the bony injury. Patients with high spinal cord disruption are at risk for shock due to physiologic disruption of sympathetic fibers. Significant neurologic recovery is rare. However, there are several partial or incomplete spinal cord injury syndromes where the prognosis is better. Central cord syndrome typically occurs in older persons who experience hyperextension injuries. Motor function, pain, and temperature sensation are preserved in the lower extremities but diminished in the upper extremities. Some functional

---

**Figure 7-16.** Epidural hematomas (A) have a distinctive convex shape on computed tomographic scan, whereas subdural hematomas (B) are concave along the surface of the brain.

**Figure 7-17.** A. A right middle cerebral infarct noted on a computed tomographic scan of the head. Such a finding should prompt imaging to rule out an associated extracranial cerebrovascular injury. B. An internal carotid artery pseudoaneurysm documented by angiography.
recovery usually occurs, but is often not a return to normal. Anterior cord syndrome is characterized by diminished motor function, pain, and temperature sensation below the level of the injury, but position sensing, vibratory sensation, and crude touch are maintained. Prognosis for recovery is poor. Brown-Séquard syndrome is usually the result of a penetrating injury in which one-half of the spinal cord is transected. This lesion is characterized by the ipsilateral loss of motor function, proprioception, and vibratory sensation, whereas pain and temperature sensation are lost on the contralateral side.

During the primary survey, identification of injuries to the neck with exsanguination, expanding hematomas, airway obstruction, or aerodigestive injuries is a priority. A more subtle injury that may not be identified is a fracture of the larynx due to blunt trauma. Signs and symptoms include hoarseness, subcutaneous emphysema (Fig. 7-18), or a palpable fracture. Penetrating injuries of the anterior neck that violate the platysma are potentially life-threatening because of the density of critical structures in this region. Although operative exploration is appropriate for overt injuries, selective nonoperative management has been proven safe (Fig. 7-19).

Indications for immediate operative intervention for penetrating cervical injury include hemodynamic instability, significant external hemorrhage, or evidence of aerodigestive injury. The management algorithm for hemodynamically stable patients is based on the presenting symptoms and anatomic location of injury, with the neck being divided into three distinct zones (Fig. 7-20). Zone I is inferior to the clavicles encompassing the thoracic outlet structures, zone II is between the thoracic outlet and the angle of the mandible, and zone III is above the angle of the mandible. Due to technical difficulties of injury exposure and varying operative approaches, a precise preoperative diagnosis is desirable for symptomatic zone I and III injuries. Therefore, these patients should ideally undergo diagnostic imaging before operation if they remain hemodynamically stable. Management of patients is further divided into those who are symptomatic and those who are not (see Fig 7-19). Specific symptoms or signs that should be identified include dysphagia, hoarseness, hematoma, venous bleeding, minor hemoptysis, and

---

**Figure 7-18.** A laryngeal fracture results in air tracking around the trachea along the prevertebral space (arrows).

**Figure 7-19.** Algorithm for the management of penetrating neck injuries. CTA = computed tomographic angiography; GSW = gunshot wound.
subcutaneous emphysema. Symptomatic patients, without overt injuries, should undergo CTA with further evaluation or operation based upon the imaging findings. Overall, less than 15% of penetrating cervical trauma requires neck exploration. Asymptomatic patients are typically observed for 6 to 12 hours. The one caveat is asymptomatic patients with a transcervical gunshot wound; these patients should undergo CTA to determine the trajectory of the bullet; further studies are performed based on proximity to major structures. Such additional imaging includes angiography, soluble contrast esophagram followed by barium esophagogram, esophagoscopy, or bronchoscopy. Angiographic diagnosis, particularly of zone III injuries, can then be managed by selective angioembolization.

Chest. Blunt trauma to the chest may involve the chest wall, thoracic spine, heart, lungs, thoracic aorta and great vessels, and rarely the esophagus. Most of these injuries can be evaluated by physical examination and chest radiography, with supplemental CT scanning to exclude vascular injury. Any patient who undergoes an intervention in the ED—endotracheal intubation, central line placement, tube thoracostomy—needs a repeat chest radiograph to document the adequacy of the procedure. This is particularly true in patients undergoing tube thoracostomy for a pneumothorax or hemothorax. Patients with persistent pneumothorax, large air leaks after tube thoracostomy, or difficulty ventilating should undergo fiber-optic bronchoscopy to exclude a tracheobronchial injury or presence of a foreign body. Patients with hemothorax must have a chest radiograph documenting complete evacuation of the chest; a persistent hemothorax that is not drained by two chest tubes is termed a caked hemothorax and mandates immediate thoracotomy.

Figure 7-20. For the purpose of evaluating penetrating injuries, the neck is divided into three zones. Zone I is to the level of the clavicular heads and is also known as the thoracic outlet. Zone II is located between the clavicles and the angle of the mandible. Zone III is above the angle of the mandible.

Occult thoracic vascular injury must be diligently sought due to the high mortality of a missed lesion. Widening of the mediastinum on initial anteroposterior chest radiograph, caused by a hematoma around an injured vessel that is contained by the mediastinal pleura, suggests an injury of the great vessels. The mediastinal abnormality may suggest the location of the arterial injury (i.e., left-sided hematomas are associated with descending blunt aortic injuries [BAI], whereas right-sided hematomas are seen with innominate injuries) (Fig. 7-22). Posterior rib fractures, sternal fractures with laceration of small vessels, and mediastinal venous bleeding also can produce similar hematomas. Other chest radiographic findings suggestive of a BAI are summarized in Table 7-5 (Fig. 7-23). However, at least 7% of patients with a descending BAI have a normal chest radiograph. Therefore, screening CTA is performed based on the mechanism of injury: high-energy deceleration motor vehicle collision with frontal or lateral impact (>30 mph frontal impact and >23 mph lateral impact), motor vehicle collision with ejection, falls of >25 ft, or direct impact (horse kick to chest, snowmobile, or ski collision with tree). In >95% of patients who survive to reach the ED, the BAI occurs just distal to the left subclavian artery, where it is tethered by the ligamentum arteriosum (Fig. 7-24). In 2% to 5% of patients the injury occurs in the ascending aorta, in the transverse arch, or at the diaphragm. Reconstructions with multislice CTA obviate the need for invasive arteriography.

For penetrating thoracic trauma, physical examination, plain posteroanterior and lateral chest radiographs with metallic markings of wounds, and pericardial ultrasound will identify the majority of injuries. Injuries of the esophagus and trachea are the exceptions. Bronchoscopy should be performed to evaluate the trachea in patients with a persistent air leak from the chest tube or mediastinal air. Patients at risk for an esophageal injury should undergo bedside esophagoscopy or soluble contrast esophagography followed by barium examination to look for extravasation of contrast. As with neck injuries,
hemodynamically stable patients with transmediastinal gunshot wounds should undergo CT scanning to determine the path of the bullet; trajectory in proximity to vascular or visceral structures dictates the need for angiography, endoscopy, or operative plan. If there is a suspicion of a subclavian artery injury, brachial-brachial indices should be measured, but >60% of patients with an injury may not have a pulse deficit. Therefore, CTA should be performed based on injury proximity to intrathoracic vasculature. Finally, with GSWs identified on the chest, penetrating trauma should not be presumed to be isolated to the thorax. Injury to contiguous body cavities (i.e., the abdomen and neck) must be excluded; plain radiographs are a rapid, effective screening modality to identify retained bullet fragments.

**Abdomen.** The abdomen is a diagnostic black box. Fortunately, with few exceptions, it is not necessary to determine in the ED which intra-abdominal organs are injured, only whether an exploratory laparotomy is necessary. However, physical examination of the abdomen can be unreliable in making this determination, and drugs, alcohol, and head and spinal cord injuries can complicate the clinical evaluation. The presence of abdominal rigidity and hemodynamic compromise is an undisputed indication

---

**Table 7-5**

<table>
<thead>
<tr>
<th>Findings on chest radiograph suggestive of a descending thoracic aortic tear</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Widened mediastinum</td>
</tr>
<tr>
<td>2. Abnormal aortic contour</td>
</tr>
<tr>
<td>3. Tracheal shift</td>
</tr>
<tr>
<td>4. Nasogastric tube shift</td>
</tr>
<tr>
<td>5. Left apical cap</td>
</tr>
<tr>
<td>6. Left or right paraspinal stripe thickening</td>
</tr>
<tr>
<td>7. Depression of the left main bronchus</td>
</tr>
<tr>
<td>8. Obliteration of the aorticopulmonary window</td>
</tr>
<tr>
<td>9. Left pulmonary hilar hematoma</td>
</tr>
</tbody>
</table>

---

**Figure 7-22.** Location of the hematoma within the mediastinal silhouette suggests the type of great vessel injury. A predominant hematoma on the left suggests the far more common descending torn aorta (**A; arrows**), whereas a hematoma on the right indicates a relatively unusual but life-threatening innominate artery injury (**B; arrows**).
for prompt surgical exploration. For the remainder of patients, a variety of diagnostic adjuncts are used to identify abdominal injury.

The diagnostic approach differs for penetrating trauma and blunt abdominal trauma. As a rule, laparotomy is warranted for gunshot or shotgun wounds that penetrate the peritoneal cavity because most have significant internal injuries. The standard has been that anterior truncal gunshot wounds between the fourth intercostal space and the pubic symphysis whose trajectory as determined by radiograph or wound location indicates peritoneal penetration should undergo laparotomy (Fig. 7-25). However, there has been increased use of CT scanning to facilitate nonoperative management of abdominal GSWs. The exception is penetrating trauma isolated to the right upper quadrant; in hemodynamically stable patients with trajectory confined to the liver by CT scan, nonoperative observation may be reasonable. In obese patients, if the gunshot wound is thought to be tangential through the subcutaneous tissues, CT scan can delineate the track and exclude peritoneal violation. Laparoscopy is another option to assess peritoneal penetration for tangential wounds it should not be done in unstable patients. In the scenario of tangential high energy GSWs, however, it is possible to sustain a transmitted intraperitoneal hollow visceral injury due to a blast insult. Gunshot wounds to the back or flank are more difficult to evaluate because of the retroperitoneal location of the injured abdominal organs. Triple-contrast CT scan can delineate the trajectory of the bullet and identify peritoneal violation or retroperitoneal entry, and associated injuries.

In contrast to gunshot wounds, stab wounds that penetrate the peritoneal cavity are less likely to injure intra-abdominal organs. Anterior abdominal stab wounds (from costal margin to inguinal ligament and bilateral midaxillary lines) should be explored under local anesthesia in the ED to determine if the fascia has been violated. Injuries that do not penetrate the peritoneal cavity do not require further evaluation, and the patient...
may be discharged from the ED. Patients with fascial penetration must be further evaluated for intra-abdominal injury because there is up to a 50% chance of requiring laparotomy. Debate remains over whether the optimal diagnostic approach is serial examination, diagnostic peritoneal lavage (DPL), or CT scanning. The most recent evidence supports serial examination and laboratory evaluation. Patients with stab wounds to the right upper quadrant can undergo CT scanning to determine trajectory and confinement to the liver for potential nonoperative care. Those with stab wounds to the flank and back should undergo contrasted CT to assess for the potential risk of retroperitoneal injuries of the colon, duodenum, and urinary tract.

Penetrating thoracoabdominal wounds may cause occult injury to the diaphragm. Patients with gunshot or stab wounds to the left lower chest should be evaluated with diagnostic laparoscopy or DPL to exclude diaphragmatic injury. In general, penetrating right diaphragm injury is ignored unless there is a major underlying liver injury with a risk of bilio-pleural fistula. Diagnostic laparoscopy may be preferred in patients with a positive chest radiograph (hemothorax or pneumothorax) or in those who would not tolerate a DPL. For patients undergoing DPL evaluation, laboratory value cutoffs to rule out diaphragm injury are different from traditional values formerly used for abdominal stab wounds (Table 7-6). An RBC count of >10,000/μL is considered a positive finding and an indication for abdominal evaluation; patients with a DPL RBC count between 1000/μL and 10,000/μL should undergo laparoscopy or thoracoscopy.

Blunt abdominal trauma is now evaluated initially by FAST examination, and this has supplanted DPL (Fig. 7-26). FAST is not 100% sensitive, however, so diagnostic peritoneal aspiration is warranted in hemodynamically unstable patients without a defined source of blood loss to rule out abdominal hemorrhage. FAST is used to identify free intraperitoneal fluid (Fig. 7-27) in Morrison’s pouch, the left upper quadrant, and the pelvis. Although this method is sensitive for detecting

---

**Figure 7-25.** Algorithm for the evaluation of penetrating abdominal injuries. AASW = anterior abdominal stab wound; CT = computed tomography; DPL = diagnostic peritoneal lavage; GSW = gunshot wound; LWE = local wound exploration; RUQ = right upper quadrant; SW = stab wound.

---

**Table 7-6**

Criteria for “positive” finding on diagnostic peritoneal lavage

<table>
<thead>
<tr>
<th></th>
<th>ABDOMINAL TRAUMA</th>
<th>THORACO-ABDOMINAL STAB WOUNDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red blood cell count</td>
<td>&gt;100,000/mL</td>
<td>&gt;10,000/mL</td>
</tr>
<tr>
<td>White blood cell count</td>
<td>&gt;500/mL</td>
<td>&gt;500/mL</td>
</tr>
<tr>
<td>Amylase level</td>
<td>&gt;19 IU/L</td>
<td>&gt;19 IU/L</td>
</tr>
<tr>
<td>Alkaline phosphatase level</td>
<td>&gt;2 IU/L</td>
<td>&gt;2 IU/L</td>
</tr>
<tr>
<td>Bilirubin level</td>
<td>&gt;0.01 mg/dL</td>
<td>&gt;0.01 mg/dL</td>
</tr>
</tbody>
</table>

---

*Tangential GSWs may also be evaluated with diagnostic laparoscopy.

**A positive local wound exploration is defined as violation of the posterior fascia.*
Figure 7-26. Algorithm for the initial evaluation of a patient with suspected blunt abdominal trauma. CT = computed tomography; DPA = diagnostic peritoneal aspiration; FAST = focused abdominal sonography for trauma; Hct = hematocrit.

Figure 7-27. Focused abdominal sonography for trauma imaging detects intra-abdominal hemorrhage. Hemorrhage is presumed when a fluid stripe is visible between the right kidney and liver (A), between the left kidney and spleen (B), or in the pelvis (C).
intrapertioneal fluid of >250 mL, it does not reliably determine the source of hemorrhage nor grade solid organ injuries.\textsuperscript{51} Patients with fluid on FAST examination, considered a “positive FAST,” who do not have immediate indications for laparotomy (hemodynamically stable, no evidence of peritonitis) undergo CT scanning to quantify their injuries. Injury grading using the American Association for the Surgery of Trauma (AAST) grading scale (Table 7-7) is an important component of nonoperative management of solid organ injuries. Additional findings that should be noted on CT scan in patients with solid organ injury include contrast extravasation (i.e., a “blush”), the amount of intra-abdominal hemorrhage, and presence of pseudoaneurysms (Fig. 7-28). CT also is indicated for hemodynamically stable patients for whom the physical examination is unreliable. Despite the increasing diagnostic accuracy of multidetector CT scanners, identification of intestinal injuries remains a limitation. Bowel injury is suggested by findings of thickened bowel wall, “streaking” in the mesentery, free fluid without associated solid organ injury, or free intrapertioneal air.\textsuperscript{52,53} Patients with free intra-abdominal fluid without solid organ injury are closely monitored for evolving signs of peritonitis; if patients have a significant closed head injury or cannot be serially examined, DPL should be performed to exclude bowel injury. If DPL is pursued, an intraumbilical approach is used (Fig. 7-29). After placement of the catheter, a 10-mL syringe is connected and the abdominal contents aspirated (termed a diagnostic peritoneal aspiration). The aspirate is considered positive if >10 mL of blood is aspirated. If <10 mL is withdrawn, a liter of normal saline is instilled. The effluent is withdrawn via siphoning and sent to the laboratory for RBC count, white blood cell (WBC) count, and determination of amylase, bilirubin, and alkaline phosphatase levels. Values representing positive findings are summarized in Table 7-6.

### Table 7-7

| American Association for the Surgery of Trauma grading scales for solid organ injuries |
|---------------------------------|---------------------------------|
| **Liver Injury Grade** | **Subcapsular Hematoma** | **Laceration** |
| Grade I | <10% of surface area | <1 cm in depth |
| Grade II | 10%-50% of surface area | 1–3 cm |
| Grade III | >50% of surface area or >10 cm in depth | >3 cm |
| Grade IV | 25%-75% of a hepatic lobe | |
| Grade V | >75% of a hepatic lobe | |
| Grade VI | Hepatic avulsion | |

<table>
<thead>
<tr>
<th><strong>Splenic Injury Grade</strong></th>
<th><strong>Hilum</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade I</td>
<td>&lt;10% of surface area</td>
</tr>
<tr>
<td>Grade II</td>
<td>10%-50% of surface area</td>
</tr>
<tr>
<td>Grade III</td>
<td>&gt;50% of surface area or &gt;10 cm in depth</td>
</tr>
<tr>
<td>Grade IV</td>
<td>&gt;25% devascularization</td>
</tr>
<tr>
<td>Grade V</td>
<td>Shattered spleen Complete devascularization</td>
</tr>
</tbody>
</table>

**Pelvis.** Blunt injury to the pelvis may produce mechanically unstable fractures with major hemorrhage (Fig. 7-30). Plain radiographs will reveal gross abnormalities, but CT scanning is necessary to determine the precise geometry. Sharp spicules of bone can lacerate the bladder, rectum, or vagina. Alternatively, bladder rupture may result from a direct blow to the torso if the bladder is full. CT cystography is performed if the urinalysis demonstrates RBCs. Urethral injuries are suspected if examination reveals blood at the meatus, scrotal or perineal hematomas, or a high-riding prostate on rectal examination. Urethrograms should be obtained for stable patients before placing a Foley catheter to avoid false passage and subsequent stricture. Major vascular injuries of the external iliacs causing bleeding are uncommon in blunt pelvic trauma; however, thrombosis of either the arteries or veins in the iliofemoral system may occur, and CTA should be performed for evaluation if there is a pulse differential. Life-threatening hemorrhage can be associated with pelvic fractures and may initially
Figure 7-29. Diagnostic peritoneal lavage is performed through an infraumbilical incision unless the patient has a pelvic fracture or is pregnant. A. The linea alba is sharply incised, and the catheter is directed into the pelvis. B. The abdominal contents should initially be aspirated using a 10-mL syringe.

preclude definitive imaging. Treatment algorithms for patients with complex pelvic fractures and hemodynamic instability are presented later in the chapter.

**Extremities.** Blunt or penetrating trauma to the extremities requires an evaluation for fractures, ligamentous disruption, and neurovascular injury. Plain radiographs are used to evaluate fractures, whereas ligamentous injuries, particularly those of the knee and shoulder, can be imaged with magnetic resonance imaging. Physical examination identifies the majority of arterial injuries, and findings are classified as either hard signs or soft signs of vascular injury (Table 7-8). In general, hard signs constitute indications for operative exploration, whereas soft signs are indications for further testing or observation. Bony fractures or knee dislocations should be realigned before definitive vascular examination. In management of vascular trauma, controversy exists regarding the treatment of patients with soft signs of injury, particularly those with injuries in proximity to major vessels. It is known that some of these patients will have arterial injuries that require repair. The most common approach has been to measure SBP using Doppler ultrasonography and compare the value for the injured side with that for the uninjured side, termed the A-A index. If the pressures are within 10% of each other, a significant injury is unlikely, and no further evaluation is performed. If the difference is >10%, CTA or arteriography is indicated. Others argue that there are occult injuries, such as pseudoaneurysms or injuries of the profunda femoris or peroneal arteries, which may not be detected with this technique. If hemorrhage occurs from these injuries, compartment syndrome and limb loss may occur. Although busy trauma centers continue to debate this issue, the surgeon who is obliged to treat the occasional injured patient may be better served by performing CTA in selected patients with soft signs. In patients with hard signs of vascular injury, on-table angiography may be useful to localize the arterial injury and thus, limit tissue dissection. For example, a patient with an absent popliteal pulse and femoral shaft fracture due to a bullet that entered the lateral hip and exited below the medial knee could have injured either the femoral or popliteal artery anywhere along its course (Fig. 7-31).

**GENERAL PRINCIPLES OF MANAGEMENT**

Over the past 25 years there has been a remarkable change in management practices and operative approach for the injured patient. With the advent of CT scanning, nonoperative management of solid organ injuries has replaced routine operative exploration. Those patients who do require operation may be treated with less radical resection techniques, such as splenorrhaphy or partial nephrectomy. Colonic injuries, previously mandating colostomy, are now repaired primarily in virtually all cases. Additionally, the type of anastomosis has shifted from a double-layer closure to a continuous running single-layer closure; this method is technically equivalent to and faster than the interrupted multilayer techniques. Adoption of damage control surgical techniques in physiologically deranged patients has resulted in limited initial operative time, with definitive injury repair delayed until after resuscitation in the surgical intensive care unit (SICU) with physiologic restoration. Abdominal drains, once considered mandatory for parenchymal injuries and some anastomoses, have disappeared; fluid collections are managed by percutaneous techniques. Newer endovascular
techniques such as stenting of arterial injuries and angioembolization are routine adjuncts. Blunt cerebrovascular injuries have been recognized as a significant, preventable source of neurologic morbidity and mortality. The use of preperitoneal pelvic packing for unstable pelvic fractures as well as early fracture immobilization with external fixators are paradigm shifts in management. Recently resuscitative endovascular balloon occlusion (REBOA) has been added to the armamentarium for life-threatening pelvic fracture bleeding\textsuperscript{57,58} (Fig. 7-32). Finally, the institution of massive transfusion protocols balances the benefit of blood component therapy against immunologic risk. Viscoelastic hemostatic assays (TEG and ROTEM) have been shown to be superior to traditional laboratory tests and have been central to the evolving concept of goal-directed hemoostasis.\textsuperscript{59} These many conceptual changes have significantly improved survival of critically injured patients; they have been promoted and critically reviewed by academic trauma centers via forums such as the American College of Surgeons Committee on Trauma, the American Association for the Surgery of Trauma, the International Association of Trauma Surgery and Intensive Care, the Pan-American Trauma Congress, and other surgical organizations.

**Transfusion Practices**

Injured patients with life-threatening hemorrhage develop acute coagulopathy of trauma (ACOT).\textsuperscript{60,61} The mechanism for inadequate clot formation remains uncertain, but it is believed to involve activation of protein C, which impairs Va and VIIa, glycocalyx breakdown, which releases heparin sulfate, immune

---

**Table 7-8**

<table>
<thead>
<tr>
<th>Signs and symptoms of peripheral arterial injury</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HARD SIGNS (OPERATION MANDATORY)</strong></td>
</tr>
<tr>
<td>Pulsatile hemorrhage</td>
</tr>
<tr>
<td>Absent pulses</td>
</tr>
<tr>
<td>Acute ischemia</td>
</tr>
<tr>
<td>A-A index of &lt;0.9</td>
</tr>
</tbody>
</table>

A-A index = systolic blood pressure on the injured side compared with that on the uninjured side.

---

**Figure 7-30.** The three types of mechanically unstable pelvis fractures are lateral compression (A), anteroposterior compression (B), and vertical shear (C).
activation with the releases of DAMPs, DNA, histone, polyphosphate, and PMN elastase, and complement activation. Fibrinolysis is an important component of the ACOT; hyperfibrinolysis and fibrinolysis shutdown are both associated with increased mortality. Fresh whole blood, arguably the optimal replacement, has not been available in the United States since the early 1980s. Rather, its component parts, packed red blood cells (PRBCs), fresh frozen plasma, platelets, and cryoprecipitate, are administered. Specific transfusion triggers for individual blood components remain debated. Although current critical care guidelines indicate that PRBC transfusion should occur once the patient’s hemoglobin level is <7 g/dL, in the acute phase of resuscitation a hemoglobin of 10 g/dL is suggested to facilitate hemostasis via platelet margination. The traditional thresholds for blood component replacement in the patient manifesting a coagulopathy have been INR >1.5, PTT >1.5 normal, platelet count >50,000/μL, and fibrinogen >100 mg/dL. However, these guidelines have been replaced by TEG and ROTEM criteria in many trauma centers. Such guidelines are designed to limit the transfusion of immunologically active blood components and decrease the risk of transfusion-associated lung injury and secondary multiple organ failure.

In the critically injured patient requiring large amounts of blood component therapy, a massive transfusion protocol should be followed (Fig. 7-33). This approach calls for administration of various components in a specific ratio during transfusion to achieve restoration of blood volume to reverse shock and correct coagulopathy. Although the optimal ratio is unknown, current evidence suggests a presumptive 1:2 red cell:plasma ratio in patients at risk for massive transfusion (10 units of PRBCs in 6 hours). Because complete typing and cross-matching takes up to 45 minutes, patients requiring emergent transfusions are given type O-negative RBCs. Similarly, without time for blood typing, AB plasma is the universal donor, although A plasma appears to be a safe option. Blood typing, and to a lesser extent cross-matching, is essential to avoid life-threatening intravascular hemolytic transfusion reactions. Trauma centers and their associated blood banks must have the capability of transfusing tremendous quantities of blood components because it is not unusual to have >50 component units transfused during one procedure and have the patient survive. Massive transfusion protocols, established preemptively, permit coordination of the activities of surgeons, anesthesiologists, and blood bankers to facilitate transfusion of the appropriate blood products.

Postinjury coagulopathy due to shock is aggravated by core hypothermia and metabolic acidosis, termed the bloody vicious cycle, and now commonly referred to as the lethal triad. The pathophysiology is multifactorial and includes inhibition of temperature-dependent enzyme-activated coagulation cascades, platelet dysfunction, endothelial abnormalities, and fibrinolytic activity. Such coagulopathy may be insidious, so the surgeon must be cognizant of subtle signs such as excessive bleeding from the cut edges of skin. Point-of-care viscoelastic assays (TEG and ROTEM), which provide a comprehensive assessment of clot capacity and fibrinolysis, can provide useful information within 15 minutes. In contrast, traditional laboratory tests of coagulation capability (i.e., INR, PTT, fibrinogen levels, and platelet count) requires at least 45 minutes. Using damage control techniques to limit operative time and provide physiologic restoration in the SICU can be lifesaving (see “Damage Control Surgery”).
Prophylactic Measures

All injured patients undergoing an operation should receive preoperative antibiotics. The type of antibiotic is determined by the anticipated source of contamination in the abdomen or other operative region; additional doses should be administered during the procedure based on blood loss and the half-life of the antibiotic. Extended postoperative antibiotic therapy is administered only for contaminated open fractures. Tetanus prophylaxis is administered to all patients according to published guidelines.

Trauma patients are at risk for venous thromboembolism and its associated morbidity and mortality. In fact, pulmonary embolus can occur much earlier in the patient’s hospital course than previously believed. Patients at higher risk for venous thromboembolism are those with multiple fractures of the pelvis and lower extremities, those with TBI or spinal cord injury, and those requiring ligation of large veins in the abdomen and lower extremities. Morbidly obese patients and those over 55 years of age are at additional risk. Administration of low molecular weight heparin (LMWH) is initiated as soon as possible.
as bleeding has been controlled and there is stable intracranial pathology. In high-risk patients, antiplatelet therapy should be added.\textsuperscript{74} Removable inferior vena caval filters should be considered if there are prolonged contraindications to administration of LMWH. Additionally, pulsatile compression stockings (also termed \textit{sequential compression devices}) are used routinely unless there is a fractured lower extremity or vascular injury.

A final prophylactic measure that is usually not considered is thermal protection. Hemorrhagic shock impairs perfusion and metabolic activity throughout the body, with resultant decrease in heat production and body temperature. Removing the patient’s clothes causes a second thermal insult, and infusion of cold PRBCs or room temperature crystalloid exacerbates the problem. As a result, injured patients can become hypothermic, with temperatures below 34°C (93.2°F) upon arrival in the OR. Hypothermia aggravates coagulopathy and provokes myocardial irritability. Therefore, prevention must begin in the ED by maintaining a comfortable ambient temperature, covering patients with warm blankets, and administering warmed IV fluids and blood products. Additionally, in the OR a Bair Hugger\textsuperscript{®} warmer (the upper body or lower body blanket) and heated inhalation via the ventilatory circuit is instituted. For cases of severe hypothermia (temperature <30°C [86°F]), arteriovenous rewarming should be considered.

**Operative Approaches and Exposure**

**Cervical Exposure.** Operative exposure for midline structures of the neck (e.g., trachea, thyroid, bilateral carotid sheaths) is obtained through a collar incision; this is typically performed two finger breadths above the sternal notch, but can be varied based on the level of anticipated injury. After subplatysmal flap elevation, the strap muscles are divided in the midline to gain access to the central neck compartment. More superior and lateral structures are accessed by extending the collar incision upward along the sternocleidomastoid muscle; this may be done bilaterally if necessary. For unilateral injuries, neck exploration is done through an incision extending from the mastoid down to the clavicle, along the anterior border of the sternocleidomastoid muscle (Fig. 7-34).

---

**Figure 7-34.** A. Unilateral neck exploration is performed through an incision along the anterior border of the sternocleidomastoid muscle; exposure of the carotid artery requires early division of the facial vein. B. The distal internal carotid artery is exposed by dividing the ansa cervicalis, which permits mobilization of the hypoglossal nerve. C. Further exposure is facilitated by resection of the posterior belly of the digastric muscle.
BASIC CONSIDERATIONS

PART I

**Figure 7-36.** A. A “clamshell” thoracotomy provides exposure to bilateral thoracic cavities. B. Sternal transection requires individual ligation of both the proximal and distal internal mammary arteries on the undersurface of the sternum.

The carotid sheath, containing the carotid artery, jugular vein, and vagus nerve, is opened widely to examine these structures. The facial vein, which marks the carotid bifurcation, is usually ligated for exposure of the internal carotid artery and the hypoglossal nerve is the next structure encountered.

Exposure of the distal carotid artery in zone III is difficult (see Fig. 7-34). The first step is division of the ansa cervicalis to facilitate mobilization of the hypoglossal nerve. Next, the posterior portion of the digastric muscle, which overlies the internal carotid, is transected. The glosopharyngeal and vagus nerves are also mobilized and retracted as necessary. If accessible, the styloid process and attached muscles are removed. In desperate situations, anterior displacement of the mandible (subluxation) may be helpful or the vertical ramus of the mandible may be divided. However, the latter maneuver often entails resection of the parotid gland, and the facial nerve is at risk of injury.

**Thoracic Incisions.** An anterolateral thoracotomy, with the patient placed supine, is the most versatile incision for emergent thoracic exploration. The location of the incision is in the fifth interspace, in the inframammary line (Fig. 7-35). If access is needed to both pleural cavities, the original incision can be extended across the sternum with a Lebsche knife, into a “clamshell” thoracotomy (Fig. 7-36). If the sternum is divided, the internal mammary arteries should be ligated to prevent blood loss. The heart, lungs, descending aorta, pulmonary hilum, and esophagus are accessible with this approach. For control of the great vessels, the superior portion of the sternum may be divided vertically with extension of the incision into the neck considered. A method advocated for access to the proximal left subclavian artery is through a fourth interspace anterolateral thoracotomy, superior sternal extension, and left supraclavicular incision (“trap door” thoracotomy). Although the trap door procedure is appropriate after resuscitative thoracotomy, the proximal left subclavian artery can be accessed more easily via a sternotomy with a supraclavicular extension. If the left subclavian artery is injured outside the thoracic outlet, vascular control can be obtained via the sternotomy and definitive repair done through the supraclavicular incision. Emergent median sternotomy is optimal for anterior stab wounds to the heart. Typically, these patients have pericardial tamponade and may undergo placement of a pericardial drain before a semiurgent median sternotomy is performed. Patients in extremis, however, should undergo anterolateral thoracotomy.

Median sternotomy with cervical extension is used for rapid exposure in patients with presumed proximal subclavian, innominate, or proximal carotid artery injuries. Care must be taken to avoid injury to the phrenic and vagus nerves that pass over the subclavian artery and to the recurrent laryngeal nerve passing posteriorly. Posterolateral thoracotomies are used for exposure of injuries to the trachea or main stem bronchi near the carina or the upper esophagus (right posterolateral thoracotomy) and tears of the descending thoracic aorta or lower esophagus (left posterolateral thoracotomy).

**Emergent Abdominal Exploration.** Abdominal exploration in adults is performed using a midline incision because of its versatility. For children under the age of 6, a transverse incision may be advantageous. Making the incision is faster with a scalpel than with an electrosurgical unit; incisional abdominal wall bleeding should be ignored until intra-abdominal sources of hemorrhage are controlled. Liquid and clotted blood are
evacuated with multiple laparotomy pads to identify the major source(s) of active bleeding. After blunt trauma, the spleen and liver should be palpated first and packed if fractured, and the infracolic mesentery should be inspected for zone I vascular injury. In contrast, after a penetrating wound the search for bleeding should pursue the trajectory of the penetrating device. If the patient has an SBP of <70 mmHg when the abdomen is opened, digital pressure or a clamp should be placed on the aorta at the diaphragmatic hiatus. After the source of hemorrhage is localized, direct digital occlusion (vascular injury) or laparotomy pad packing (solid organ injury) is used to control bleeding (Fig. 7-37). If the liver is the source in a hemodynamically unstable patient, additional control of bleeding is obtained by clamping the hepatic pedicle with a vascular clamp or Rummel tourniquet (termed the Pringle maneuver) (Fig. 7-38). Similarly, clamping the splenic hilum may be required for hilar bleeding. When the spleen is mobilized, it should be gently rotated medially to expose the lateral peritoneum; this peritoneum and endoabdominal fascia are incised, which allows blunt dissection of the spleen and pancreas as a composite from the retroperitoneum anterior to Gerota’s fascia (Fig. 7-39).

Rapid exposure of the intra-abdominal vasculature can prove challenging in the face of exsanguinating hemorrhage. Proximal control of the aorta is obtained at the diaphragmatic hiatus; if an aortic injury is supraceliac, transecting the left crus of diaphragm or extending the laparotomy via a left thoracotomy may be necessary.

Figure 7-37. A sagittal view of packs placed to control hepatic hemorrhage. LAP = laparotomy.

Figure 7-38. The Pringle maneuver, performed with a vascular clamp, occludes the hepatic pedicle containing the portal vein, hepatic artery, and common bile duct.

Figure 7-39. To mobilize the spleen, an incision is made into the endoabdominal fascia 1 cm lateral to the reflection of the peritoneum onto the spleen (A). While the spleen is gently rotated medially, a plane is developed between the pancreas and left kidney (B). With complete mobilization, the spleen can reach the level of the abdominal incision.
An alternative for a contained hematoma is placement of a trans-femoral REBOA into zone I. The first decision is whether the patient has a supracolic or an infracolic vascular injury. Supracolic injuries (aorta, celiac axis, proximal superior mesenteric artery [SMA], and left renal arteries) are best approached via a left medial visceral rotation (Fig. 7-40). This is done by incising the lateral peritoneal reflection (white line of Toldt) beginning at the distal descending colon and extending the incision along the colonic splenic flexure, around the posterior aspect of the spleen, and behind the gastric fundus, ending at the esophagus. The left colon, spleen, pancreas, and stomach are then rotated toward the midline. The authors prefer to leave the kidney in situ when mobilizing the viscera because this exaggerates the separation of the renal vessels from the SMA. The operative approach for SMA injuries is based on the level of injury. Fullen zone I SMA injuries, located posterior to the pancreas, are best exposed by a left medial visceral rotation. Fullen zone II SMA injuries, extending from the pancreatic edge to the middle colic branch, on the other hand, are approached via the lesser sac along the inferior edge of the pancreas at the base of the transverse mesocolon; the pancreatic body may be divided to gain proximal vascular access. More distal SMA injuries, Fullen zones III and IV, are approached directly within the mesentery. A venous injury behind the pancreas, from the junction of the superior mesenteric, splenic, and portal veins, is accessed by dividing the neck of the pancreas. Inferior vena cava injuries are approached by a right medial visceral rotation (Fig. 7-41). Proximal control is obtained just above the iliac bifurcation with direct pressure via a sponge stick; the injury is identified by cephalad dissection along the anterior surface of the inferior vena cava. A Satinsky clamp can be used to control anterior caval wounds.

Injuries of the iliac vessels pose a unique problem for emergent vascular control due to the number of vessels, their close proximity, and cross circulation. Proximal control at the infrarenal aorta arrests the arterial bleeding and avoids splanchnic and renal ischemia; however, venous injuries are not controlled with aortic clamping. Tamponade with digital pressure or with a folded laparotomy pad held directly over the bleeding site usually will establish hemostasis sufficient to prevent exsanguination. If hemostasis is not adequate to expose the vessel proximal and distal to the injury, sponge sticks can be strategically placed on either side of the injury and carefully adjusted to improve hemostasis. Alternatively, complete pelvic vascular isolation (Fig. 7-42) may be required to control hemorrhage for adequate visualization of the injuries. The right common iliac artery obscures the bifurcation of the vena cava and the right iliac vein; the iliac artery may require division to expose venous injuries in this area (Fig. 7-43). The artery must be repaired after the venous injury is treated, however, because of limb-threatening ischemia.

Once overt hemorrhage is controlled, sources of enteric contamination are identified by serially running along the small and large bowel, looking at all surfaces. Associated hematomas should be unroofed to rule out adjacent bowel injury. The anterior and posterior aspects of the stomach should be inspected, which requires opening the lesser sac for complete visualization. Duodenal injuries should be evaluated with a wide Kocher maneuver. During exploration of the lesser sac, visualization and palpation of the pancreas is done to exclude injury. Palpating the anterior surface is not sufficient because the investing

Figure 7-40. A left medial visceral rotation is used to expose the abdominal aorta.

Figure 7-41. A right medial visceral rotation is used to expose the infrahepatic vena cava.
fascia may mask a pancreatic injury; mobilization, including evaluation of the posterior aspect, is critical. After injuries are identified, whether to use damage control techniques or perform primary repair of injuries is based on the patient’s intraoperative physiologic status (see “Damage Control Surgery” and “Treatment of Specific Injuries”). In a patient with multisystem trauma, enteral access via gastrostomy or jejunostomy tube should be considered. If abdominal closure is indicated after the patient’s injuries are addressed, the abdomen is irrigated with warm saline and the midline fascia is closed with a running heavy absorbable suture. The skin is closed selectively based on the amount of intra-abdominal contamination.

Vascular Repair Techniques. Initial control of vascular injuries is accomplished digitally by applying enough direct pressure to stop the hemorrhage. Sharp dissection is used to define the injury and mobilize sufficient length for proximal and distal control. Fogarty thromboembolectomy should be done proximally and distally to optimize collateral blood flow. Heparinized saline (50 units/mL) is then injected into the proximal and distal ends of the injured vessel to prevent small clot formation on the exposed intima and media. Ragged edges of the injury site should be debrided using sharp dissection. Intravascular shunts are used when there are multiple life-threatening injuries or the arterial injury is anticipated to require saphenous vein interposition reconstruction.

Options for the treatment of vascular injuries are listed in Table 7-9. Arterial repair should always be done for the aorta, carotid, innominate, brachial, superior mesenteric, proper hepatic, renal, iliac, femoral, and popliteal arteries. Named arteries that usually tolerate ligation include the right or left hepatic artery and the celiac artery. In the lower extremities, at least one artery with distal runoff should be salvaged. Arterial injuries that may be treated nonoperatively include small pseudoaneurysms, intimal dissections, small intimal flaps, and small arteriovenous

<table>
<thead>
<tr>
<th>Table 7-9</th>
<th>Options for the treatment of vascular injuries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observation</td>
<td></td>
</tr>
<tr>
<td>Ligation</td>
<td></td>
</tr>
<tr>
<td>Lateral suture repair</td>
<td></td>
</tr>
<tr>
<td>End-to-end primary anastomosis</td>
<td></td>
</tr>
<tr>
<td>Interposition grafts</td>
<td></td>
</tr>
<tr>
<td>Autogenous vein</td>
<td></td>
</tr>
<tr>
<td>Polytetrafluoroethylene graft</td>
<td></td>
</tr>
<tr>
<td>Dacron graft</td>
<td></td>
</tr>
<tr>
<td>Transpositions</td>
<td></td>
</tr>
<tr>
<td>Extra-anatomic bypass</td>
<td></td>
</tr>
<tr>
<td>Endovascular</td>
<td></td>
</tr>
<tr>
<td>Stents</td>
<td></td>
</tr>
<tr>
<td>Embolization</td>
<td></td>
</tr>
</tbody>
</table>

Figure 7-42. Pelvic vascular isolation. A. Initially, clamps are placed on the aorta, inferior vena cava, and bilateral external iliac vessels. B. With continued dissection, the clamps can be moved progressively closer to the vascular injury to limit unwarranted ischemia.

Figure 7-43. The right common iliac artery can be divided to expose the bifurcation of the inferior vena cava and the right common iliac vein.
fistulas in the extremities. Follow-up imaging is performed 1 to 2 weeks after injury to confirm healing. Venous repair should be performed for injuries of the superior vena cava, the inferior vena cava proximal to the renal veins, and the portal vein, although the portal vein may be ligated in extreme cases. The SMV should be repaired optimally, but >80% of patients will survive following ligation. Similarly, the left renal vein can usually be ligated adjacent to the IVC due to collateral decompression.

The type of operative repair for a vascular injury is based on the extent and location of injury. Lateral suture repair is preferred for arterial injuries with minimal loss of tissue. End-to-end primary anastomosis is performed if the vessel can be repaired without tension. Arterial defects of 1 to 2 cm often can be bridged by mobilizing the severed ends of the vessel after ligating small branches. The aorta, subclavian artery, brachial artery, and popliteal artery however, are difficult to mobilize for additional length. To avoid postoperative stenosis, particularly in smaller arteries, beveling or spatulation should be used so that the completed anastomosis is slightly larger in diameter than the native artery (Fig. 7-44). The authors emphasize the parachute technique to ensure precision placement of the posterior suture line (Fig. 7-45). If this technique is used, traction must be maintained on both ends of the suture, or leakage from the posterior aspect of the suture line may occur.

Interposition grafts are used when end-to-end anastomosis cannot be accomplished without tension despite mobilization. For vessels <6 mm in diameter (e.g., internal carotid, brachial, superficial femoral, and popliteal arteries), autogenous greater saphenous vein (GSV) from the contralateral groin should be used because polytetrafluoroethylene (PTFE) grafts of <6 mm have a prohibitive rate of thrombosis. When GSV is not available, autologous options include the cephalic and basilic veins. Larger arteries (e.g., subclavian, innominate, aorta, common iliac) are bridged by PTFE grafts. PTFE is preferred over Dacron because of the reported decreased risk of infection. Aortic or iliac arterial injuries may be complicated by enteric contamination from colon or small bowel injuries. There is a natural reluctance to place artificial grafts in such circumstances, but graft infections are rare, and the time required to perform an axillofemoral bypass is excessive. Therefore, after the control of hemorrhage, bowel contamination is contained and the abdomen irrigated before placing PTFE grafts. After placement of the graft, it is covered with peritoneum or omentum before definitive treatment of the enteric injuries.

Transposition procedures can be used when an artery has a bifurcation and one vessel can be ligated safely. Injuries of the proximal internal carotid can be treated by mobilizing the adjacent external carotid, dividing it distal to the internal injury, and performing an end-to-end anastomosis between it and the distal internal carotid (Fig. 7-46). The proximal stump of the internal carotid is oversewn, with care taken to avoid a blind pocket where a clot may form. Injuries of the common and external iliac arteries can be handled in a similar fashion (Fig. 7-47), while maintaining flow in at least one internal iliac artery.

Venous injuries should be repaired when technically feasible. Small injuries without loss of tissue can be treated with lateral suture repair. More complex repairs with interposition grafts may thrombose, but this typically occurs gradually over 1 to 2 weeks. During this time adequate collateral circulation develops, which is sufficient to avoid acute venous hypertension. Therefore, it is reasonable to use ringed PTFE for venous interposition grafting and accept a gradual, but eventual, thrombosis while allowing time for collateral circulation to develop. Such an approach is reasonable for venous injuries of the superior vena cava, suprarenal vena cava, SMV, and popliteal vein because ligation of these is associated with significant morbidity. In the remainder of venous injuries, the vein may be ligated. In such patients, chronic venous hypertensive complications in the lower extremities often can be avoided by (a) temporary use of elastic bandages (Ace wraps) applied from the toes to the hips at the end of the procedure, and (b) judicious elevation of the lower extremities. These measures should be maintained for

---

**Figure 7-44.** Small arteries repaired with an end-to-end anastomosis are prone to stricture. Enlarging the anastomosis by beveling the cut ends of the injured vessel can minimize this problem. A curved hemostat is a useful adjunct to create the curve.

**Figure 7-45.** The parachute technique is helpful for accurate placement of the posterior sutures of an anastomosis when the arterial end is fixed and an interposition graft is necessary. Traction must be maintained on both ends of the suture to prevent loosening and leakage of blood. Six stitches should be placed before the graft is pulled down to the artery.
Carotid transposition is an effective approach for treating injuries of the proximal internal carotid artery.

Figure 7-46. Carotid transposition is an effective approach for treating injuries of the proximal internal carotid artery.

Transposition procedures can be used for iliac artery injuries to eliminate the dilemma of placing an interposition polytetrafluoroethylene graft in the presence of enteric contamination. A. Right common iliac artery transposed to left common iliac artery. B. Left internal iliac artery transposed to the distal right common iliac artery. C. Right internal iliac artery transposed to the right external iliac artery.

Figure 7-47. Transposition procedures can be used for iliac artery injuries to eliminate the dilemma of placing an interposition polytetrafluoroethylene graft in the presence of enteric contamination. A. Right common iliac artery transposed to left common iliac artery. B. Left internal iliac artery transposed to the distal right common iliac artery. C. Right internal iliac artery transposed to the right external iliac artery.

1 week; if the patient has no peripheral edema with ambulation, these maneuvers are no longer required.

**Damage Control Surgery**

The recognition of the bloody vicious cycle and the introduction of damage control surgery (DCS) have improved the survival of critically injured patients. Conceptually, the bloody vicious cycle, first described in 1981, is the lethal combination of coagulopathy, hypothermia, and metabolic acidosis (Fig. 7-48). Hypothermia from evaporative and conductive heat loss and diminished heat production occurs despite the use of warming blankets and blood warmers. The metabolic acidosis of shock is exacerbated by aortic clamping, administration of vasopressors, massive RBC transfusions, and impaired myocardial performance. The ACOT, described previously, is compounded by hemodilution, hypothermia, and acidosis. Once the cycle starts, each component magnifies the other, which leads to a downward spiral and ultimately a fatal arrhythmia. The purpose of DCS is to limit operative time so that the patient can be returned to the SICU for physiologic restoration and the cycle thereby broken. Indications to limit the initial operation and institute DCS techniques include a combination of refractory hypothermia (temperature <35°C), profound acidosis (arterial pH <7.2, base deficit >15 mmol/L), and refractory coagulopathy. The decision to abbreviate a trauma laparotomy is made intraoperatively as the patient’s clinical course becomes clearer and laboratory values become available.

The goal of DCS is to control surgical bleeding and limit GI spillage. The operative techniques used are temporary measures, with definitive repair of injuries delayed until the patient is physiologically replete. Controlling surgical bleeding while preventing ischemia is of utmost importance during DCS. Aortic injuries must be repaired using an interposition PTFE
graft. Although celiac artery injuries may be ligated, the SMA must maintain flow, and the early insertion of an intravascular shunt is advocated. Similarly, perfusion of the iliac system and infrainguinal vessels can be restored with a vascular shunt, with interposition graft placement delayed. Arterial reconstruction following shunt placement should be done optimally within 6 hours. Venous injuries are preferentially treated with ligation in damage control situations, except for the suprarenal inferior vena cava and popliteal vein. For extensive solid organ injuries to the spleen or one kidney, excision is indicated rather than an attempt at operative repair. For hepatic injuries, perihepatic packing of the liver will usually tamponade bleeding (see Fig. 7-37). Translobar gunshot wounds of the liver are best controlled with balloon catheter tamponade, whereas deep lacerations can be controlled with Foley catheter inflation deep within the injury track (Fig. 7-49). For thoracic injuries requiring DCS several options exist. For bleeding peripheral pulmonary injuries, wedge resection using a stapler is performed. In penetrating injuries, pulmonary tractotomy is used to divide the parenchyma (Fig. 7-50); individual vessels and bronchi are then ligated using a 3-0 polydioxanone suture (PDS), and the track is left open. Patients who sustain more proximal injuries may require formal pulmonary resection, but pneumonectomy is poorly tolerated. Recent experimental work suggests inhaled nitric oxide (NO) will reduce right heart failure following pneumonectomy. Cardiac injuries may be temporarily controlled using a running 3-0 nonabsorbable polypropylene suture or skin staples. Pledged repair should be performed for the relatively thin right ventricle.

The second key component of DCS is limiting enteric content spillage. Small GI injuries (stomach, duodenum, small intestine, and colon) may be controlled using a rapid whipstitch of 3-0 PDS. Complete transection of the bowel or segmental damage is controlled using a GIA stapler, often with resection of the injured segment. Alternatively, open ends of the bowel may be ligated using umbilical tapes to limit spillage. Pancreatic injuries, regardless of location, are packed and the evaluation of ductal integrity postponed. Urologic injuries may require catheter diversion. Before the patient is returned to the SICU, the abdomen must be closed temporarily. Temporary closure of the abdomen is accomplished using an antimicrobial surgical incise drape (Ioban, 3M Health Care, St Paul, MN) (Fig. 7-51). In this technique, the bowel is covered with a fenestrated subfascial sterile drape (45 × 60 cm Steri-Drape 3M Health Care), and two Jackson-Pratt drains are placed along the fascial edges; this is

**Figure 7-48.** The bloody vicious cycle. FFP = fresh frozen plasma; RBC = red blood cell.
then covered using an Ioban drape, which allows closed suction to control reperfusion-related ascitic fluid egress while providing adequate space for bowel expansion to prevent abdominal compartment syndrome. During the initial DCS stage, the subfascial sterile drape is not covered by a blue towel so that the status of the bowel and hemorrhage control can be assessed. The use of direct peritoneal resuscitation following DCS should be considered; 18 Fr round Blake drains may be placed intraoperatively for the instillation of the dialysate solution. Return to the OR within 24 hours is planned once the patient clinically improves, as evidenced by normothermia, normalization of coagulation test results, and correction of acidosis.

**TREATMENT OF SPECIFIC INJURIES**

**Head Injuries**

**Intracranial Injuries.** CT scanning, performed on all patients with a significant closed head injury (GCS score <14), identifies and quantitates intracranial lesions as well as intracranial hypertension. Patients with intracranial hemorrhage, including epidural hematoma, subdural hematoma, subarachnoid hemorrhage, intracerebral hematoma or contusion, and diffuse axonal injury, are admitted to the SICU. In patients with abnormal findings on CT scans and GCS scores of ≤8, intracranial pressure (ICP) should be monitored using fiber-optic intraparenchymal devices or intraventricular catheters. Although an ICP of 10 mmHg is the upper limit of normal, therapy is targeted to maintain an ICP of <20 mmHg. The newest neurosurgery guidelines additionally recommend maintaining the cerebral perfusion pressure (CPP) at >60 mmHg. Indications for operative intervention to remove space-occupying hematomas are based on the clot volume, amount of midline shift, location of the clot, GCS score, and ICP. A shift of >5 mm typically is considered an indication for evacuation, but this is not an absolute rule. Smaller hematomas that are in treacherous locations, such as the posterior fossa, may require drainage due to brain stem compression or impending herniation. Removal of small hematomas may also improve ICP and cerebral perfusion in patients with elevated ICP that is refractory to medical therapy. Patients with diffuse cerebral edema resulting in excessive ICP may require a decompressive craniectomy, although a recent AAST multicenter trial questioned the benefits. Patients with open or depressed skull fractures, with or without sinus involvement, may require operative intervention. Penetrating injuries to the head may require operative intervention for hemorrhage control, evacuation of blood, skull fracture fixation, or debridement.

General surgeons in communities without emergency neurosurgical coverage should have a working knowledge of burr hole placement in the event that emergent evacuation is required for a life-threatening epidural hematoma (Fig. 7-52). The typical clinical course of an epidural hematoma is an initial loss of consciousness, a lucid interval, and recurrent loss of consciousness with an ipsilateral fixed and dilated pupil. While decompression of subdural hematomas may be delayed, epidural hematomas require evacuation within 70 minutes. The final stages of this sequence are caused by blood accumulation that forces the temporal lobe medially, with resultant compression of the third cranial nerve and eventually the brain stem. The burr hole is made on the side of the dilated pupil to decompress the intracranial space. After stabilization, the patient is transferred to a facility with neurosurgical capability for formal craniotomy.

In addition to operative intervention, postinjury care directed at limiting secondary injury to the brain is critical. The goal of resuscitation and management in patients with head injuries is to avoid hypotension (SBP of <100 mmHg) and hypoxia (partial pressure of arterial oxygen of <60 or arterial oxygen saturation of <90%). Attention, therefore, is focused on maintaining cerebral perfusion rather than merely lowering
ICP. Resuscitation efforts aim for a euvolemic state and a SBP of >100 mmHg. Cerebral perfusion pressure (CPP) is equal to the mean arterial pressure minus the ICP, with a target range of >60 mmHg. CPP can be increased by either lowering ICP or raising mean arterial pressure. Sedation, osmotic diuresis, paralysis, ventricular drainage, and barbiturate coma are used in sequence, with coma induction being the last resort. The role of decompressive craniectomy for refractory ICH remains controversial. The partial pressure of carbon dioxide (PCO₂) should be maintained in a normal range (35–40 mmHg), but for temporary management of acute intracranial hypertension, inducing cerebral vasoconstriction by hyperventilation to a PCO₂ of

Figure 7-51. Temporary closure of the abdomen entails covering the bowel with a fenestrated subfascial 45 × 60 cm sterile drape (A), placing Jackson-Pratt drains along the fascial edge (B), and then occluding with an Ioban drape (C, D).
TRAUMA

CHAPTER 7

<30 mmHg is occasionally warranted. Moderate hypothermia (32°–33°C [89.6°–91.4°F]) has been shown experimentally to improve neurologic outcomes, but clinical studies have not validated this concept.33,85,86 Patients with intracranial hemorrhage should be monitored for postinjury seizures, and prophylactic anticonvulsant therapy is indicated for 7 days after injury.33

Maxillofacial Injuries. Maxillofacial injuries are common with multisystem trauma and require coordinated management by the trauma surgeon and the specialists in otolaryngology, plastic surgery, ophthalmology, and oral and maxillofacial surgery. Delay in addressing these systems that control vision, hearing, smelling, breathing, eating, and phonation may produce dysfunction and disfigurement with serious psychological impact. The maxillofacial complex is divided into three regions; the upper face containing the frontal sinus and brain; the midface containing the orbits, nose, and zygomaticomaxillary complex; and the lower face containing the mandible. High-impact kinetic energy is required to fracture the frontal sinus, orbital rims, and mandible, whereas low-impact forces will injure the nasal bones and zygoma.

The most common scenario, which at times may be life-threatening, is bleeding from facial fractures.87 Temporizing measures include nasal packing, Foley catheter tamponade of posterior nasal bleeding, and oropharyngeal packing. Prompt angioembolization will halt exsanguinating hemorrhage. Fractures of tooth-bearing bone are considered open fractures and require antibiotic therapy and semiurgent repair to preserve the airway as well as the functional integrity of the occlusion (bite) and the aesthetics of the face. Orbital fractures may compromise vision, produce muscle injury causing diplopia, or change orbital volume to produce a sunken appearance to the orbit. Nose and nasoethmoidal fractures should be assessed carefully to identify damage to the lacrimal drainage system or to the cribiform plate producing cerebrospinal fluid rhinorrhea. After initial stabilization, a systematic physical examination of the head and neck should be performed that also includes cranial nerve examination and three-dimensional CT scanning of the maxillofacial complex (Fig. 7-53).

Cervical Injuries

Spine. Treatment of injuries to the cervical spine is based on the level of injury, the stability of the spine, the presence of subluxation, the extent of angulation, the level of neurologic deficit, and the overall condition of the patient. In general, physician-supervised axial traction, via cervical tongs or the more commonly used halo vest, is used to reduce subluxations and stabilize the injury. Immobilization of injuries also is achieved with spinal orthoses (braces), particularly in those with associated thoracolumbar injuries. Surgical fusion typically is performed in patients with neurologic deficit, those with angulation of >11° or translation of >3.5 mm, and those who remain unstable after halo placement. Indications for immediate operative intervention are deterioration in neurologic function and fractures or dislocations with incomplete deficit. Historically, methylprednisolone was administered to patients with acute spinal cord injury after blunt injury, with clinical data suggesting a small benefit to initiating a 24-hour infusion if started within 3 hours and a 48-hour infusion if started within 3 to 8 hours.88 Current guidelines, however, no longer recommend steroids for acute injuries.89 The role and timing of operative surgical decompression after acute spinal cord injury is debated, and
the concept of damage control has been suggested. However, evidence supports urgent decompression of bilateral locked facets in patients with incomplete tetraplegia or with neurologic deterioration. Performing surgery within 24 hours may decrease length of stay and complications. Complete injuries of the spinal cord remain essentially untreatable. Yet, approximately 3% of patients who present with flaccid quadriplegia have concussive injuries, and these patients represent the very few who seem to have miraculous recoveries.

**Vascular.** Cervical vascular injuries due to either blunt or penetrating trauma can result in devastating neurologic sequelae or exsanguination. Penetrating injuries to the carotid artery and internal jugular vein usually are obvious on operative neck exploration. The principles of vascular repair techniques (discussed previously) apply to carotid injuries, and options for repair include end-to-end primary repair (often possible with mobilization of the common carotid), graft interposition, and transposition procedures. All carotid injuries should be repaired except in patients who present in coma with a delay in transport. Prompt revascularization of the internal carotid artery, using a temporary Pruitt-Inahara shunt, should be considered in patients arriving in profound shock. Otherwise, carotid shunting should be done selectively as in elective carotid endarterectomy, but the patient should be systemically anticoagulated. Currently, we administer heparin with an ACT target of 250 sec. Tangential wounds of the internal jugular vein should be repaired by lateral venorrhaphy, but extensive wounds are efficiently addressed by ligation. However, it is not advisable to ligate both jugular veins due to potential intracranial hypertension. Vertebral artery injuries due to penetrating trauma are difficult to control operatively because of the artery’s protected location within the foramen transversarium. Although exposure from an anterior approach can be accomplished by removing the anterior elements of the bony canal and the tough fascia covering the artery between the elements, typically the most efficacious control of such injuries is angioembolization. Fogarty catheter balloon occlusion, however, is useful for controlling acute bleeding if encountered during neck exploration.

Blunt injury to the carotid or vertebral arteries may cause dissection, thrombosis, or pseudoaneurysm, typically in the surgically inaccessible distal internal carotid (Fig. 7-54). Early recognition and management of these injuries is paramount because patients treated with antithrombotics have a stroke rate of <1% compared with stroke rates of 20% in untreated patients. Because treatment must be instituted during the latent period between injury and onset of neurologic sequelae, diagnostic imaging is performed based on identified risk factors (Fig. 7-55). After identification of an injury, antithrombotics are administered if the patient does not have contraindications (intracranial hemorrhage, falling hemoglobin level with solid organ injury or complex pelvic fractures). Heparin, started without a loading dose at 15 units/kg per hour, is titrated to achieve a PTT between 40 and 50 seconds or antiplatelet agents are initiated (aspirin 325 mg/d or clopidogrel 75 mg/d). The types of

Figure 7-54. The Denver grading scale for blunt cerebrovascular injuries. Grade I: irregularity of the vessel wall, dissection/intramural hematoma with <25% luminal stenosis. Grade II: visualized intraluminal thrombus or raised intimal flap, or dissection/intramural hematoma with 25% or more luminal narrowing. Grade III: pseudoaneurysm. Grade IV: vessel occlusion. CAI = carotid artery injury; VAI = vertebral artery injury.
antithrombotic treatment appear equivalent in published studies to date, and the duration of treatment is empirically recommended to be 6 months.93,94 The role of carotid stenting for grade II or III internal carotid artery injuries remain controversial; current literature suggests stenting be reserved for symptomatic patients or markedly enlarging pseudoaneurysms.95 Thrombosis of the internal jugular veins caused by blunt trauma can occur unilaterally or bilaterally and is often discovered incidentally because most patients are asymptomatic. Bilateral thrombosis can aggravate cerebral edema in patients with serious head injuries; stent placement should be considered in such patients if ICP remains elevated.

Figure 7-55. Screening and treatment algorithm for blunt cerebrovascular injuries (BCVIs). ASA = acetylsalicylic acid; BRB = bright red blood; CHI = closed head injury; C-spine = cervical spine; CT = computed tomography; DAI = diffuse axonal injury; GCS = Glasgow Coma Scale score; MRI = magnetic resonance imaging; MS = mental status; Neg = negative; pt = patient; PTT = partial thromboplastin time; TIA = transient ischemic attack.
Aerodigestive. Fractures of the larynx and trachea may manifest as cervical emphysema. Fractures documented by CT scan are usually repaired. Common injuries include thyroid cartilage fractures, rupture of the thyroepiglottic ligament, disruption of the arytenoids or vocal cord tears, and cricoid fractures. After debridement of devitalized tissue, tracheal injuries are repaired end-to-end using a single layer of interrupted absorbable sutures. Associated injuries of the esophagus are common in penetrating injuries due to its close proximity. After debridement and repair, vascularized tissue is interposed between the repaired esophagus and trachea, and a closed suction drain is placed. The sternocleidomastoid muscle or strap muscles are useful for interposition and help prevent postoperative fistulas.

Chest Injuries
The most common injuries from both blunt and penetrating thoracic trauma are hemothorax and pneumothorax. More than 85% of patients can be definitively treated with a chest tube. The indications for thoracotomy include significant initial or ongoing hemorrhage from the tube thoracostomy and specific imaging-identified diagnoses (Table 7-10). One caveat concerns the patient who presents after a delay. Even when the initial chest tube output is 1.5 L, if the output ceases and the lung is reexpanded, the patient may be managed nonoperatively if hemodynamically stable.

Great Vessels. Over 90% of thoracic great vessel injuries are due to penetrating trauma, although blunt injury to the innominate, subclavian, or descending aorta may cause a pseudoaneurysm or frank rupture. Simple lacerations of the ascending or transverse aortic arch can be repaired with lateral aortorrhaphy. Repair of posterior aortic injuries, complex ascending or transverse injuries, or those requiring interposition grafting of the arch, require full cardiopulmonary bypass. Innominate artery injuries are repaired using the bypass exclusion technique, which avoids the need for cardiopulmonary bypass. Bypass grafting from the proximal aorta to the distal innominate with a prosthetic tube graft is performed before the postinjury hematoma is entered. The PTFE graft is anastomosed end-to-side from the proximal undamaged aorta and anastomosed end-to-end to the innominate artery (Fig. 7-56). The origin of the innominate is then oversewn at its base to exclude the pseudoaneurysm or other injury. Subclavian artery injuries can be repaired using lateral aortorrhaphy or PTFE graft interposition; due to its multiple branches and tethering of the artery, end-to-end primary anastomosis is not advocated if there is a significant segmental loss.

Descending BAI may require urgent intervention. However, operative intervention for intracranial or intra-abdominal hemorrhage or unstable pelvic fractures takes precedence. To prevent aortic rupture, pharmacologic therapy with a selective $\beta_1$-antagonist, esmolol, should be instituted in the trauma bay, with a target SBP of <100 mmHg and heart rate of <100/min. Endovascular stenting is now the mainstay of treatment.

While endograft sizing has improved, the major question is long-term outcome in younger patients. Open repair of the descending aorta is accomplished using partial left heart bypass to prevent spinal cord and splanchnic ischemia and reduce left ventricular afterload (Fig. 7-57). Nonoperative management for grade I

---

**Table 7-10**

<table>
<thead>
<tr>
<th>Indications for operative treatment of thoracic injuries</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Initial tube thoracostomy drainage of $&gt;$1000 mL (penetrating injury) or $&gt;$1500 mL (blunt injury)</td>
</tr>
<tr>
<td>- Ongoing tube thoracostomy drainage of $&gt;$200 mL/h for 3 consecutive hours in noncoagulopathic patients</td>
</tr>
<tr>
<td>- Caked hemothorax despite placement of two chest tubes</td>
</tr>
<tr>
<td>- Great vessel injury (endovascular techniques may be used in selected patients)</td>
</tr>
<tr>
<td>- Pericardial tamponade</td>
</tr>
<tr>
<td>- Cardiac herniation</td>
</tr>
<tr>
<td>- Massive air leak from the chest tube with inadequate ventilation</td>
</tr>
<tr>
<td>- Tracheal or main stem bronchial injury diagnosed by endoscopy or imaging</td>
</tr>
<tr>
<td>- Open pneumothorax</td>
</tr>
<tr>
<td>- Esophageal perforation</td>
</tr>
<tr>
<td>- Air embolism</td>
</tr>
</tbody>
</table>

---

**Figure 7-56.**

A. Angiography reveals a 1-cm pseudoaneurysm of the innominate artery origin. B. In the first stage of the bypass exclusion technique, a 12-mm polytetrafluoroethylene graft is anastomosed end to side from the proximal undamaged aorta, tunneled under the vein, and anastomosed end to end to the innominate artery. C. The origin of the innominate is then oversewn at its base to exclude the pseudoaneurysm.
When repairing a tear of the descending thoracic aorta, perfusion of the spinal cord while the aorta is clamped is achieved by using partial left heart bypass. The venous cannula is inserted into the left superior pulmonary vein because it is less prone to tearing than the left atrium (LA).

intimal aortic injuries is accomplished with antiplatelet agents and blood pressure control.

Heart. Blunt and penetrating cardiac injuries have widely differing presentations and therefore disparate treatments. Survivable penetrating cardiac injuries consist of wounds that can be repaired operatively; most are stab wounds. Before repair of the injury is attempted, hemorrhage should be controlled; injuries to the atria can be clamped with a Satinsky vascular clamp, whereas digital pressure is used to occlude the majority of ventricular wounds. Foley catheter occlusion of larger stellate lesions is described, but even minimal traction may enlarge the original injury. Temporary control of hemorrhage, and at times definitive repair, may be accomplished with skin staples for left ventricular lacerations; the myocardial edges of the laceration must coapt in diastole for stapling to be technically feasible.

Definitive repair of cardiac injuries is performed with either running 3-0 polypropylene suture or interrupted, pledgeted 2-0 polypropylene suture (Fig. 7-58). Use of pledgets may be particularly important in the right ventricle to prevent sutures from pulling through the thinner myocardium. Injuries adjacent to coronary arteries should be repaired using horizontal mattress sutures because use of running sutures results in coronary occlusion and distal infarction. Gunshot wounds may result in stellate lesions or contused, extremely friable myocardium adjacent to the wound. When the edges of such complex wounds cannot be fully approximated and hence the repair is not hemostatic, the authors have used surgical adhesive (BioGlue) to achieve hemostasis. Occasionally, interior structures of the heart may be damaged. Intraoperative auscultation or postoperative hemodynamic assessment usually identifies such injuries. ECHO can diagnose the injury and quantitate its effect on cardiac output. Immediate repair of valvular damage or septal defects rarely is necessary and would require cardiopulmonary bypass, but structural intracardiac lesions may progress, and thus patients must have a follow-up ECHO.

Patients with blunt cardiac injury typically present with persistent tachycardia or conduction disturbances, but occasionally present with tamponade due to atrial or right ventricular rupture. There are no pathognomonic ECG findings, and cardiac enzyme levels do not correlate with the risk of cardiac complications. Therefore, patients for whom there is high clinical suspicion of cardiac contusion and who are hemodynamically stable should be monitored for dysrhythmias for 24 hours by telemetry. Patients with hemodynamic instability should undergo ECHO to evaluate for wall motion abnormalities (particularly hypokinesis of the right ventricle), pericardial fluid, valvular dysfunction, chordae rupture, or diminished ejection fraction. If such findings are noted or if vasoactive agents are required, cardiac function can be continuously monitored using a pulmonary artery catheter and serial SICU trans-thoracic or transesophageal ECHO.

Trachea, Bronchi, and Lung Parenchyma. Less than 1% of all injured patients sustain intrathoracic tracheobronchial
injuries, and only a small number require operative intervention. Although penetrating injuries may occur throughout the tracheobronchial system, blunt injuries most commonly occur within 2.5 cm of the carina. For patients with a massive air leak requiring emergent exploration, initial control of the injury to provide effective ventilation is obtained by passing an endotracheal tube either beyond the injury or into the contralateral mainstem bronchus. Principles of repair are similar to those for repair of cervical tracheal injuries. Devitalized tissue is debrided, and primary end-to-end anastomosis with 3-0 PDS suture is performed. Dissection should be limited to the area of injury to prevent disruption of surrounding bronchial vasculature and ensuing ischemia and stricture. Suture lines should be encircled with vascularized tissue, either pericardium, intercostal muscle, or pleura. Expectant management is employed for bronchial injuries that are less than one-third the circumference of the airway and have no evidence of a persistent major air leak.11 In patients with peripheral bronchial injuries, indicated by persistent air leaks from the chest tube and documented by endoscopy, bronchoscopically directed fibrin glue sealing may be useful.

The majority of pulmonary parenchymal injuries are suspected based upon identification of a pneumothorax; the vast majority can be managed with a tube thoracostomy. Identified parenchymal injuries encountered during thoracic exploration for a massive hemorthax are managed without resection as much as possible. Bronchovenuous fistula is a constant threat and should be minimized by prompt control of a major air leak. Peripheral lacerations with persistent bleeding can be managed with stapled wedge resection. For central injuries, the current treatment is pulmonary tractotomy, which permits selective ligation of individual bronchioles and bleeder, prevents the development of an intraparenchymal hematoma or air embolism, and reduces the need for formal lobar resection (see Fig. 7-50).106,107 A stapling device, preferably the longest stapler available (e.g., GIA-100), is inserted directly into the injury track and positioned along the thinnest section of overlapping parenchyma. The injury track is thus filleted open, which allows direct access to the bleeding vessels and leaking bronchi. The majority of injuries are definitively managed with selective ligation, and the defect is left open. Occasionally, tractotomy reveals a more proximal vascular on bronchial injury that must be treated with formal lobectomy. Injuries severe enough to mandate pneumonectomy usually are fatal because of right heart decompensation.108

One parenchymal injury that may be discovered during thoracic imaging is a posttraumatic pulmonary pseudocyst, colloquially termed a pneumatocele.109 Traumatic pneumatoceles typically follow a benign clinical course and are treated with aggressive pain management, pulmonary toilet, and serial chest radiography to monitor for resolution of the lesion. If the patient has persistent fever or leukocytosis, however, chest CT is done to evaluate for an evolving abscess because pneumatoceles may become infected. CT-guided catheter drainage may be required in such cases because 25% of patients do not respond to antibiotic therapy alone. Surgery, ranging from partial resection to anatomic lobectomy, is indicated for unresolved complex pneumatoceles or infected lesions refractory to antibiotic therapy and drainage.

The most common complication after thoracic injury is development of an empyema. Management is based on CT diagnostic criteria.110 Percutaneous drainage is indicated for a single loculation without appreciable rind. While fibrinolytics are often used for empyema, there is a paucity of data to support their use. Early decortication via video-assisted thoracic surgery should be done promptly in patients with multiple loculations or a pleural rind of >1 cm.111 Antibiotic treatment is based on definitive culture results, but presumptive antibiotics should cover MRSA in the SICU.

**Esophagus.** Due to the proximity of the structures, esophageal injuries often occur with tracheobronchial injuries, particularly in cases of penetrating trauma. Operative options are based on the extent and location of esophageal injury. With sufficient mobilization, a primary single-layer end-to-end anastomosis may be performed after appropriate debridement. As with cervical repairs, if there are two suture lines in close approximation (trachea or bronchi and esophagus) interposition of a vascularized pedicle is warranted to prevent fistula formation. Perforations at the gastroesophageal junction may be treated with repair and Nissan fundoplication or, for destructive injuries, segmental resection and gastric pull-up. Small esophageal injuries can be managed with stenting. With large destructive injuries or delayed presentation of injuries, esophageal exclusion with wide drainage, diverting loop esophagostomy, and placement of a gastrostomy tube should be considered.

**Chest Wall and Diaphragm.** Virtually all chest wall injuries, consisting of rib fractures and laceration of intercostal vessels, are treated nonoperatively with pain control, pulmonary toilet or ventilatory management, and drainage of the pleural space as indicated. Early institution of effective pain control is essential. The authors advocate preemptive rib blocks with 0.25% bupivacaine hydrochlo-
ride (Marcaine) in the trauma bay, followed by thoracic wall pain catheters.112 Epidural anesthesia is reserved for multiple segmental fractures. Persistent hemorrhage from a chest tube after blunt trauma most often is due to injured intercostal arteries; for unusual persistent bleeding (see Table 7-10), thoracotomy with direct ligation or angioembolization may be required to arrest hemorrhage. In cases of extensive flail chest segments, markedly displaced biconical rib fractures, or loss of 20% of the thoracic volume, open reduction and internal fixation of the fracture with plates may be warranted. Chest wall defects, particularly those seen with open pneumothorax, are repaired using local approximation of tissues or tissue transfer for coverage. Scapular and sternal fractures rarely require operative intervention but are markers for significant thoracoabdominal force during injury; significant displacement may benefit from sternal plating (Fig. 7-59). Careful examination and imaging should exclude associated injuries, including blunt cardiac injury and descending BAL. On the other hand, clavicle fractures are often isolated injuries and should be managed with pain control and immobilization. The exception is posterior dislocation of the clavicular head, which may injure the subclavian vessels.

Blunt diaphragmatic injuries usually result in a linear tear, and most injuries are large, whereas penetrating injuries are variable in size and location depending on the agent of injury. Regardless of the etiology, acute injuries are usually repaired through an abdominal approach to manage potential associated visceral injury. After delineation of the injury, the chest should be evacuated of all blood and particulate matter, and a thoracostomy tube placed if not previously done. Allis clamps are used to approximate the diaphragmatic edges, and the defect is closed with a running No. 1 polypropylene suture. Occasionally, large avulsions or shotgun wounds with extensive tissue loss will require polypropylene or biologic mesh to bridge the defect.
Alternatively, transposition of the diaphragm cephalad one to two intercostal spaces may allow repair without undue tension.73

**Abdominal Injuries**

**Liver and Extrahepatic Biliary Tract.** The liver’s large size makes it the organ most susceptible to blunt trauma, and it is frequently involved in upper torso penetrating wounds. Nonoperative management of solid organ injuries is pursued in hemodynamically stable patients who do not have overt peritonitis or other indications for laparotomy. Patients with >grade II injuries should be admitted to the SICU with frequent hemodynamic monitoring, determination of hemoglobin, and abdominal examination. The only absolute contraindication to nonoperative management is hemodynamic instability from intraperitoneal hemorrhage. Factors such as high injury grade, large hemoperitoneum, contrast extravasation, or pseudoaneurysms may predict complications or failure of nonoperative management. Angioembolization and endoscopic retrograde cholangiopancreatography (ERCP) are useful adjuncts that can improve the success rate of nonoperative management.113,114 The indication for angiography to control hepatic hemorrhage is transfusion of 4 units of RBCs in 6 hours or 6 units of RBCs in 24 hours attributable to the liver.

In the 15% of patients for whom emergent laparotomy is mandated, the primary goal is to arrest hemorrhage. Initial control of hemorrhage is best accomplished using perihpatic packing and manual compression. The edges of the liver laceration should be opposed for local pressure control of bleeding. Hemorrhage from most major hepatic injuries can be controlled with effective perihpatic packing. The right costal margin is elevated, and the pads are strategically placed over and around the bleeding site (see Fig. 7-37). Additional pads should be placed between the liver, diaphragm, and anterior chest wall until the bleeding has been controlled. Sometimes 10 to 15 pads may be required to control the hemorrhage from an extensive right lobar injury. Packing of injuries of the left lobe is not as effective because there is insufficient abdominal and thoracic wall anterior to the left lobe to provide adequate compression with the abdomen open. Fortunately, hemorrhage from the left lobe usually can be controlled by mobilizing the lobe and compressing it between the surgeon’s hands. With extensive injuries and major hemorrhage, a Pringle maneuver should be done immediately. Intermittent release of the Pringle is helpful to attenuate hepatic cellular loss.

If the patient has persistent bleeding despite packing, injuries to the hepatic artery, portal vein, and retrohepatic vasculature should be considered. A Pringle maneuver can help delineate the source of hemorrhage. Hemorrhage from hepatic artery and portal vein injuries will halt with the application of a vascular clamp across the portal triad, whereas bleeding from the hepatic veins and the retrohepatic vena cava will continue despite a Pringle maneuver.

Injuries of the portal triad vasculature should be addressed immediately. In general, ligation from the celiac axis to the level of the common hepatic artery at the gastroduodenal arterial branch is tolerated due to extensive collaterals, but the proper hepatic artery should be repaired. The right or left hepatic artery, or in urgent situations the portal vein, may be selectively ligated; occasionally, lobar necrosis will necessitate delayed anatomic resection. If the right hepatic artery is ligated, cholecystectomy also should be performed. If the vascular injury is a stab wound with clean transection of the vessels, primary end-to-end repair is done. If the injury is destructive, temporary shunting should be performed followed by interposition reversed saphenous vein graft (RSVG). Blunt avulsions of the portal structures are particularly problematic if located at the hepatic plate, flush with the liver; hemorrhage control at the liver can be attempted with directed packing or Fogarty catheters. If injury to the portal triad vasculature is more proximal, at the superior border of the pancreatic body or even retropancreatic, the pancreas must be transected to gain access for hemorrhage control and repair.

If massive venous hemorrhage is seen from behind the liver despite use of the Pringle maneuver, the patient likely has a hepatic vein or retrohepatic vena cava injury. If bleeding can be controlled with perihpatic packing, the packing should be left undisturbed and the patient observed in the SICU. Placement of a hepatic vein stent by interventional radiology may be considered. If bleeding continues despite repeated attempts at packing, then direct repair, with or without hepatic vascular isolation, should be attempted. Three techniques have been used to accomplish hepatic vascular isolation: (a) direct repair with suprahpatic and infrahepatic clamping of the vena cava and stapled assisted parenchymal resection115; (b) temporary shunting of the retrohepatic vena cava; and (c) venovenous bypass (Fig. 7-60).116

A number of methods for the definitive control of hepatic parenchymal hemorrhage have been developed. Minor lacerations may be controlled with manual compression applied directly to the injury site. Topical hemostatic techniques include

---

**Figure 7-59.** Significant sternal displacement (A; arrows) can be reduced and stabilized with sternal plating (B).
the use of an electrocautery (with the device set at 100 watts), argon beam coagulator, microcrystalline collagen, thrombin-soaked gelatin foam sponge, fibrin glue, and BioGlue. Suturing of the hepatic parenchyma with a blunt tipped 0 chronic suture (e.g., a “liver suture”) can be an effective hemostatic technique. A running suture is used to approximate the edges of shallow lacerations, whereas deeper lacerations are approximated using interrupted horizontal mattress sutures placed parallel to the edge of the laceration. When the suture is tied, tension is adequate when visible hemorrhage ceases or the liver blanches around the suture. Caution must be used to prevent hepatic necrosis. This technique of placing large liver sutures controls bleeding through reapproximation of the liver laceration rather than direct ligation of bleeding vessels. Aggressive finger fracture to identify bleeding vessels followed by individual clip or suture ligation was advocated previously but currently has a limited role in hemostasis. Hepatic lobar arterial ligation may be appropriate for patients with recalcitrant arterial hemorrhage from deep within the liver and is a reasonable alternative to a deep hepatotomy, particularly in unstable patients. Omentum can be used to fill large defects in the liver. The tongue of omentum not only obliterates potential dead space with viable tissue but also provides an excellent source of macrophages. Additionally, the omentum can provide buttressing support for parenchymal sutures.

Translobar penetrating injuries are particularly challenging because of the extent of the injury cannot be fully visualized. As discussed in “Damage Control Surgery,” options include intraparenchymal tamponade with a Foley catheter or balloon occlusion (see Fig. 7-49). If tamponade is successful with either modality, the balloon is left inflated for 24 to 48 hours followed by sequential deflation and removal at a second laparotomy. Hepatotomy with ligation of individual bleeders occasionally may be required; however, division of the overlying viable hepatic tissue may cause considerable blood loss in the coagulopathic patient. Finally, angioembolization is an effective adjunct in any of these scenarios and should be considered early in the course of treatment.

Several centers have reported patients with devastating hepatic injuries or necrosis of the entire liver who have undergone successful hepatic transplantation. Clearly this is dramatic therapy, and the patient must have all other injuries delineated, particularly those of the central nervous system, and have an excellent chance of survival excluding the hepatic injury. Because donor availability will limit such procedures, hepatic transplantation for trauma will continue to be performed only in extraordinary circumstances.

Cholecystectomy is performed for injuries of the gallbladder and after operative ligation of the right hepatic artery. Injuries of the extrahepatic bile ducts are a challenge due to their small size and thin walls. Because of the proximity of other portal structures and the vena cava, associated vascular injuries are common. These factors may preclude primary repair. Small lacerations with no accompanying loss or devitalization of adjacent tissue can be treated by the insertion of a T-tube through the wound or by lateral suturing using 6-0 monofilament absorbable suture. Virtually all transections and any injury associated with significant tissue loss will require a Roux-en-Y choledochojjunostomy. The anastomosis is performed using a single-layer interrupted technique with 5-0 monofilament absorbable suture. To reduce anastomotic tension, the jejunum should be sutured to the areolar tissue of the hepatic pedicle or porta hepatis. Injuries of the hepatic ducts are almost impossible to satisfactorily repair under emergent circumstances. One approach is to intubate the duct for external drainage and attempt a repair when the patient recovers or attempt stenting via ERC. Alternatively, the duct can be ligated if the opposite lobe is normal and uninjured.

Patients undergoing perihepatic packing for extensive liver injuries typically are returned to the OR for pack removal 24 hours after initial injury. Earlier exploration may be indicated in patients with evidence of ongoing hemorrhage. Signs of rebleeding are usually conspicuous, and include a falling hemoglobin, accumulation of blood clots under the temporary abdominal closure device, and bloody output from drains; the magnitude of hemorrhage is reflected in ongoing hemodynamic instability and metabolic monitoring. Postoperative hemorrhage should be reevaluated in the OR once the patient’s coagulopathy is corrected. Alternatively, angioembolization is appropriate for complex injuries. Patients with hepatic ischemia due to prolonged intraoperative use of the Pringle maneuver have an expected elevation but subsequent resolution of transaminase levels, whereas patients requiring hepatic artery ligation may have frank hepatic necrosis. Although febrile patients should be evaluated for infectious complications, patients with complex hepatic injuries typically have intermittent “liver fever” for the first 5 days after injury.

Aside from hemorrhage and hepatic necrosis, additional complications after significant hepatic trauma include bilomas, arterial pseudoaneurysms, and biliary fistulas (Fig. 7-61). Bilomas are loculated collections of bile, which may or may not be infected. If infected, they should be treated like an abscess via percutaneous drainage. Although small, sterile bilomas
Complications after hepatic trauma include bilomas (A; arrow), hepatic duct injuries (B), and hepatic necrosis after hepatic artery ligation or embolization (C). Eventually will be reabsorbed, larger fluid collections should be drained. Biliary ascites, due to the disruption of a major bile duct, often requires reoperation and wide drainage. Primary repair of the injured intrahepatic duct is unlikely to be successful. Resectional debridement is indicated for the removal of peripheral portions of nonviable hepatic parenchyma.

Pseudoaneurysms and biliary fistulas are rare complications in patients with hepatic injuries. Because hemorrhage from hepatic injuries often is treated without isolating individual bleeding vessels, arterial pseudoaneurysms may develop, with the potential for rupture. Rupture into a bile duct results in hemobilia, which is characterized by intermittent episodes of right upper quadrant pain, upper GI hemorrhage, and jaundice. If the aneurysm ruptures into a portal vein, portal venous hypertension with bleeding esophageal varices may occur. Either scenario is best managed with hepatic arteriography and embolization. Biliovenous fistulas, causing jaundice due to rapid increases in serum bilirubin levels, should be treated with ERCP and sphincterotomy. Rarely, a biliary fistulous communication will form with intrathoracic structures in patients with associated diaphragm injuries, resulting in a bronchobiliary or pleurobiliary fistula. Due to the pressure differential between the biliary tract (positive) and the pleural cavity (negative), the majority require operative closure. Occasionally, endoscopic sphincterotomy with stent placement will be required to address the pressure differential, and the pleurobiliary fistula will close spontaneously.

Spleen. Until the 1970s, splenectomy was considered mandatory for all splenic injuries. Recognition of the immune function of the spleen refocused efforts on operative splenic salvage in the 1980s. After demonstrated success in pediatric patients, nonoperative management has become the preferred means of splenic salvage for all patients. The identification of contrast extravasation as a risk factor for failure of nonoperative management led to liberal use of angioembolization. The role of selective angioembolization (SAE) continues to be defined, but appears warranted in high grade injuries, particularly those with contrast blush. It is clear, however, that up to 15% to 20% of patients with splenic trauma warrant early splenectomy and that failure of nonoperative management often represents inappropriate patient selection. Indications for early intervention in adults include initiation of blood transfusion within the first 12 hours and hemodynamic instability. Unlike hepatic injuries, which usually rebleed within 48 hours, delayed hemorrhage or rupture of the spleen can occur up to weeks after injury.

Splenic injuries are managed operatively by splenectomy, partial splenectomy, or splenic repair (splenorrhaphy), based on the extent of the injury and the physiologic condition of the patient. Splenectomy is indicated for significant hilar injuries, pulverized splenic parenchyma, or any >grade II injury in a patient with coagulopathy or multiple life-threatening injuries. The authors use autotransplantation of splenic implants (Fig. 7-62) to achieve partial immunocompetence in younger patients who do not have an associated enteric injury. Drains are not used. Partial splenectomy can be employed in patients in whom only the superior or inferior pole has been injured. Hemorrhage from the raw splenic edge is controlled with horizontal mattress sutures, with gentle compression of the parenchyma (Fig. 7-63). During splenorrhaphy hemostasis is achieved by topical methods (electrocautery; argon beam coagulation; application of thrombin-soaked gelatin foam sponges, fibrin glue, or BioGlue), envelopment of the injured spleen in absorbable mesh, and pledgeted suture repair.

After splenectomy or splenorrhaphy, postoperative hemorrhage may be due to an improperly ligated or unrecognized short gastric artery, or recurrent bleeding from the splenic parenchyma if splenic repair was used. An immediate postsplenectomy
increase in platelets and WBCs is normal; however, beyond postoperative day 5, a WBC count above 15,000/mm³ and a platelet/WBC ratio of <20 are associated with sepsis and should prompt a thorough search for underlying infection. A common infectious complication after splenectomy is a subphrenic abscess, which should be managed with percutaneous drainage. Additional sources of morbidity include a concurrent or unrecognized iatrogenic injury to the pancreatic tail during rapid splenectomy resulting in pancreatic ascites or fistula, or gastric perforation during short gastric vessel ligation. Enthusiasm for splenic salvage was driven by the rare, but often fatal, complication of overwhelming postsplenectomy sepsis. Overwhelming postsplenectomy sepsis is caused by encapsulated bacteria, *Streptococcus pneumoniae*, *Haemophilus influenzae*, and *Neisseria meningitidis*, which are resistant to antimicrobial treatment. In patients undergoing splenectomy, prophylaxis against these bacteria is provided via vaccines administered optimally at >14 days postinjury.

**Stomach and Small Intestine.** Little controversy exists regarding the repair of injuries to the stomach or small bowel because of a rich blood supply. Gastric wounds can be oversewn with a running single-layer suture line or closed with a stapler. If a single-layer closure is chosen, full-thickness bites should be taken to ensure hemostasis from the well-vascularized gastric wall. The most commonly missed gastric injury is the posterior wound of a through and through penetrating injury. Injuries also can be overlooked if the wound is located within the mesentry of the lesser curvature or high in the fundus. To delineate a questionable injury, the stomach can be digitally occluded at the pylorus while methylene blue-colored saline is instilled via a nasogastric (NG) tube. Alternatively, air can be introduced via the NG tube with the abdomen filled with saline. Partial gastrectomy may be required for destructive injuries, with resection of the distal antrum or pylorus reconstructed using a Billroth procedure. Patients with injuries that damage both Latarjet nerves or vagi should undergo a drainage procedure (see Chapter 26). Small intestine injuries can be repaired using a transverse running 3-0 PDS suture if the injury is less than one-third the circumference of the bowel. Destructive injuries or multiple penetrating injuries occurring close together are treated with segmental resection followed by end-to-end anastomosis using a continuous, single-layer 3-0 polypropylene suture. Mesenteric injuries may result in an ischemic segment of intestine, which mandates resection.

Following repair of GI tract injuries, patients may develop a postoperative ileus. Return of bowel function is indicated by a decrease in gastrostomy or nasogastric tube output. The topic of nutrition is well covered in other chapters (see Chapter 2), but a few issues warrant mention. Multiple studies have confirmed the importance of early total enteral nutrition (TEN) in the trauma population, particularly its impact in reducing septic complications. The route of enteral feedings (stomach vs. small bowel) tends to be less important because gut tolerance appears equivalent unless there is upper GI tract pathology. Although early enteral nutrition is the goal, evidence of bowel function should be apparent before advancing to goal tube feedings. Overzealous jejunal feeding can lead to small bowel necrosis in the patient recovering from profound shock. Patients undergoing monitoring for nonoperative management of grade III or higher solid organ injuries should receive nothing by mouth for at least 48 hours in case they require an operation. Although there is general reluctance to initiate TEN in patients with an open abdomen, a recent multicenter trial demonstrates that TEN in the postinjury open abdomen is feasible. For those patients without a bowel injury, TEN was associated with higher fascial closure rates, decreased
complications, and decreased mortality. TEN in patients with bowel injuries does not appear to alter fascial closure rates, complications, or mortality; hence, EN appears to be neither advantageous nor detrimental in these patients. Once resuscitation is complete, initiation of TEN, even at trophic levels (20 mL/h), should be considered in all injured patients with an open abdomen.

**Duodenum and Pancreas.** The spectrum of injuries to the duodenum includes hematomas, perforation (blunt blow-outs, lacerations from stab wounds, or blast injury from gunshot wounds), and combined pancreaticoduodenal injuries. The majority of duodenal hematomas are managed nonoperatively with nasogastric suction and parenteral nutrition. Patients with suspected associated perforation, suggested by clinical deterioration or imaging with retroperitoneal free air or contrast extravasation, should undergo operative exploration. A marked drop in nasogastric tube output heralds resolution of the hematoma, which typically occurs within 2 weeks; repeat imaging to confirm these clinical findings is optional. If the patient shows no clinical or radiographic improvement within 3 weeks, operative evaluation is warranted.

Small duodenal perforations or lacerations should be treated by primary repair using a running single-layer suture of 3-0 monofilament. The wound should be closed in a direction that results in the largest residual lumen. Challenges arise when there is a substantial loss of duodenal tissue. Extensive injuries of the first portion of the duodenum (proximal to the duct of Santorini) can be repaired by debridement and end-to-end anastomosis because of the mobility and rich blood supply of the distal gastric atrium and pylorus. In contrast, the second portion is tethered to the head of the pancreas by its blood supply and the ducts of Wirsung and Santorini; therefore, no more than 1 cm of duodenum can be mobilized away from the pancreas. Moreover, suture repair using an end-to-end anastomosis in the second portion often results in an unacceptably narrow lumen. Therefore, defects in the second portion of the duodenum should be “patched” with a Roux-en-Y duodenojejunostomy. Duodenal injuries with tissue loss distal to the papilla of Vater and proximal to the superior mesenteric vessels are best treated by Roux-en-Y duodenojejunostomy with the distal portion of the duodenum oversewn (Fig. 7-64). In particular, injuries in the distal third and fourth portions of the duodenum (behind the mesenteric vessels) should be resected, and a duodenojejunostomy should be performed on the D3 side of the superior mesenteric vessels.

Optimal management of pancreatic trauma is determined by where the parenchymal damage is located and whether the intrapancreatic common bile duct and main pancreatic duct remain intact. Patients with pancreatic contusions (defined as injuries that leave the ductal system intact) can be treated nonoperatively or with closed suction drainage if undergoing laparotomy for other indications. Patients with proximal pancreatic injuries, defined as those that lie to the right of the superior mesenteric vessels, are also managed with closed suction drainage. In contrast, distal pancreatic injuries are managed based upon ductal integrity. Pancreatic duct disruption can be identified through direct exploration of the parenchymal laceration, operative pancreatography, ERCP, or magnetic resonance cholangiopancreatography. Patients with distal ductal disruption undergo distal pancreatectomy, preferably with splenic preservation.

Injuries to the pancreatic head add substantial complexity because the intrapancreatic portion of the common bile duct traverses this area and often converges with the pancreatic duct. In contrast to diagnosis of pancreatic duct injuries, identification of intrapancreatic common bile duct disruption is relatively simple. The first method is to squeeze the gallbladder and look for bile leaking from the pancreatic wound. Otherwise, cholangiography, optimally via the cystic duct, is diagnostic. Definitive treatment of this injury entails division of the common bile duct superior to the first portion of the duodenum, with ligation of the distal duct and reconstruction with a Roux-en-Y choledochojejunostomy. For injuries to the head of the pancreas that involve the main pancreatic duct but not the intrapancreatic bile duct, there are few options. Distal pancreatectomy alone is rarely indicated due to the extended resection of normal gland and the resultant risk of pancreatic insufficiency. Central pancreatectomy preserves the common bile duct, and mobilization of the pancreatic body permits drainage into a posterior wall pancreaticogastrostomy or a Roux-en-Y pancreaticojejunostomy (Fig. 7-65). Although this approach avoids a pancreaticoduodenectomy (Whipple procedure), the complexity may make the pancreaticoduodenectomy more appropriate in patients with multiple injuries and is usually done in a damage control scenario. Some injuries of the pancreatic head do not involve either the pancreatic or common bile duct; if no clear ductal injury is present, drains are placed. Rarely, patients sustain destructive injuries to the head of the pancreas or combined pancreaticoduodenal injuries that require pancreaticoduodenectomy. Examples of such injuries include transection of both the intrapancreatic bile duct and the main pancreatic duct in the head of the pancreas, avulsion of the papilla of Vater from the duodenum, and destruction
of the entire second portion of the duodenum. In these cases of extensive injuries, damage control principles are often employed.

In contrast to proximal injuries, pancreatic resection continues to be advocated for major ductal disruption in the more distal pancreas. Several options exist for treating injuries of the pancreatic body and tail. In stable patients, spleen-preserving distal pancreatectomy should be performed. An alternative, which preserves both the spleen and distal transected end of the pancreas, is either a Roux-en-Y pancreaticojejunostomy or pancreaticogastrostomy. If the patient is physiologically compromised, distal pancreatectomy with splenectomy is the preferred approach. Regardless of the choice of definitive procedure, the pancreatic duct in the proximal edge of transected pancreas should be individually ligated or occluded with a TA stapler. Application of fibrin glue over the stump may be advantageous.

Pyloric exclusion may be used to divert the GI stream after high-risk, complex duodenal repairs, particularly with adjacent pancreatic injuries (Fig. 7-66). If the duodenal repair breaks down, the resultant fistula is an end fistula, which is easier to manage and more likely to close than a lateral fistula. To perform a pyloric exclusion, first a gastrostomy is made on the greater curvature near the pylorus. The pylorus is then grasped with a Babcock clamp, via the gastrostomy, and oversewn with an O polypropylene suture. A gastrojejunostomy restores GI tract continuity. Vagotomy is not necessary because a risk of marginal ulceration has not been documented. Perhaps surprisingly, the sutures maintain diversion for only 3 to 4 weeks. Alternatively, the most durable pyloric closure is a double external staple line across the pylorus using a TA stapler.

Complications should be expected after major pancreaticoduodenal injuries. Delayed hemorrhage is rare but may occur with pancreatic necrosis or abdominal infection; this usually can be managed by angioembolization. If closed suction drains have been inserted for major pancreatic trauma, these should remain in place until the patient is tolerating an oral diet or enteral nutrition. Pancreatic fistula is diagnosed after postoperative day 5 in patients with drain output of ≥30 mL/d and a drain amylase level three times the serum value. Pancreatic fistula develops in over 20% of patients with combined injuries and should be managed similar to fistulas after elective surgery (see Chapter 33). Similarly, a duodenal fistula, presumptively an end fistula if a pyloric exclusion has been done, will typically heal in 6 to 8 weeks with adequate drainage and control of intra-abdominal sepsis. Pancreatic pseudocysts in patients managed nonoperatively suggest a missed injury, and ERCP should be done to evaluate the integrity of the pancreatic duct. Late pseudocysts may be a complication of operative management and are treated much like those in patients with pancreatitis (see Chapter 33). Intra-abdominal abscesses are common and routinely managed with percutaneous drainage.

Colon and Rectum. Currently, three methods for treating colonic injuries are used: primary repair, end colostomy, and primary repair with diverting loop ileostomy. Primary repairs include lateral suture repair or resection of the damaged segment with reconstruction by ileocolostomy or colocolostomy. All suturing and anastomoses are performed using a running single-layer technique (Fig. 7-67). The advantage of definitive treatment must be balanced against the possibility of anastomotic leakage if suture lines are created under suboptimal conditions. Alternatively, although use of an end colostomy requires a second operation, an unprotected suture line with
**Figure 7-66.** A. Pyloric exclusion is used to treat combined injuries of the duodenum and the head of the pancreas as well as isolated duodenal injuries when the duodenal repair is less than optimal. B and C. The pylorus is oversewn through a gastrotomy, which is subsequently used to create a gastrojejunostomy. The authors frequently use operatively placed feeding jejunostomy tube feedings for these patients.

**Figure 7-67.** Technique for bowel repair and anastomosis. A. The running, single-layer suture is started at the mesenteric border. B. Stitches are spaced 3 to 4 mm from the edge of the bowel and advanced 3 to 4 mm, including all layers except the mucosa. C. The continuous suture is tied near the antimesenteric border.
Rectal injuries are similar to colonic injuries with respect to the ecology of the luminal contents, overall structure, and blood supply of the wall, but access to extraperitoneal injuries is limited due to the surrounding bony pelvis. Therefore, indirect treatment with intestinal diversion usually is required. The current options are loop ileostomy and sigmoid loop colostomy. The latter is preferred because it is quick and easy to perform, and provides essentially total fecal diversion. For sigmoid colostomy, technical elements include: (a) adequate mobilization of the sigmoid colon so that the loop will rest on the abdominal wall without tension, (b) maintenance of the spur of the colostomy (the common wall of the proximal and distal limbs after maturation) above the level of the skin with a one-half-inch Penrose drain or similar device, (c) longitudinal incision in the tenia coli, and (d) immediate maturation in the OR (Fig. 7-68). If the injury is accessible (e.g., in the posterior intraperitoneal portion of the rectum), repair of the injury should also be attempted. However, it is not necessary to explore the extraperitoneal rectum to repair a distal perforation. If the rectal injury is extensive, another option is to divide the rectum at the level of the injury, oversew or staple the distal rectal pouch if possible, and create an end colostomy (Hartmann’s procedure). Extensive injuries may warrant presacral drainage with Penrose drains placed along Waldeyer’s fascia via a perianal incision (see Fig. 7-68), but routine presacral drainage and rectal washout is no longer practiced. In rare instances in which destructive injuries are present, an abdominoperineal resection may be necessary to avert lethal pelvic sepsis.

Complications related to colorectal injuries include intra-abdominal abscess, fecal fistula, wound infection, and stomal complications. Intra-abdominal abscesses occur in approximately 10% of patients, and most are managed with percutaneous drainage. Fistulas occur in 1% to 3% of patients and usually present as an abscess or wound infection with subsequent continuous drainage of fecal output; the majority will heal spontaneously with routine care (see Chapter 29). Stomal complications (necrosis, stenosis, obstruction, and prolapse) occur in 5% of patients and may require either immediate or delayed reoperation. Stomal necrosis should be carefully monitored because spread beyond the mucosa may result in septic complications, including necrotizing fasciitis of the abdominal wall. Penetrating injuries that involve both the rectum and adjacent bony structures are prone to development of osteomyelitis. Bone biopsy is performed for diagnosis and bacteriologic analysis, and treatment entails long-term IV antibiotic therapy and occasionally debridement.

**Abdominal and Pelvic Vasculature.** Injury to the major arteries and veins in the abdomen can be a technical challenge. Although penetrating trauma indiscriminately affects all blood vessels, blunt trauma most commonly involves renal vasculature and occasionally the abdominal aorta. Patients with a penetrating aortic wound who survive to reach the OR frequently have a contained hematoma within the retroperitoneum. Due to lack of mobility of the abdominal aorta, few injuries are amenable to primary repair. Supraceliac aortic wounds are particularly challenging due to the need for proximal control. Small lateral perforations may be controlled with 4-0 polypropylene suture or a PTFE patch, but end-to-end interposition grafting with a PTFE tube graft is the most common repair. Blunt injuries are typically extensive intimal tears of the infrarenal aorta with resultant thrombosis, and are exposed via a direct approach; most require an interposition graft. To avoid future vascular-enteric fistulas, the vascular suture lines should be covered with omentum.

---

**Figure 7-68.** Loop colostomy will completely divert the fecal flow, allowing the low rectal injury to heal. For extensive wounds, presacral drains are inserted through a perianal incision (box) and advanced along Waldeyer’s fascia (dashed line).
Penetrating wounds to the superior mesenteric artery (SMA) are typically encountered after exploration for a gunshot wound, with “black bowel” and associated supramesocolic hematoma being pathognomonic. Blunt avulsions of the SMA are rare but should be considered in patients with a seat belt sign who have midepigastric pain or tenderness and associated hypotension. For injuries of the SMA, temporary damage control with a Pruitt-Inahara shunt can prevent extensive bowel necrosis. For definitive repair, end-to-end interposition RSVG from the proximal SMA to the SMA past the point of injury can be performed if there is no associated pancreatic injury. Alternatively, if the patient has an associated pancreatic injury, the graft should be tunnelled from the distal aorta beneath the duodenum to the distal SMA. For proximal SMV injuries, digital compression for hemorrhage control is followed by attempted venorrhaphy; ligation is an option in a life-threatening situation, but the resultant bowel edema requires aggressive fluid resuscitation. Temporary abdominal closure and a second-look operation to evaluate bowel viability should be done.

Transpelvic gunshot wounds or blunt injuries with associated pelvic fractures are the most common scenarios in patients with iliac artery injuries. As with abdominal vascular injuries, a Pruitt-Inahara shunt can be used for temporary shunting of the vessel for damage control. Definitive interposition grafting with excision of the injured segment is appropriate (see “Vascular Repair Techniques”). Careful monitoring for distal embolic events and reperfusion injury necessitating fasciotomy is imperative.

In general, outcome after pelvic vascular injuries is related to (a) the technical success of the vascular reconstruction and (b) associated soft tissue and nerve injuries. Vascular repairs rarely fail after the first 12 hours, whereas soft tissue infection is a limb threat for several weeks. Following aortic interposition grafting, the patient’s SBP should not exceed 120 mmHg for at least the first 72 hours postoperatively. Patients requiring ligation of an inferior vena cava injury often develop marked bilateral lower extremity edema. To limit the associated morbidity, the patient’s legs should be wrapped with elastic bandages from the toes to the hips and elevated. For superior mesenteric vein injuries, either ligation or thrombosis after venorrhaphy results in marked bowel edema; fluid resuscitation should be aggressive and abdominal pressure monitoring routine in these patients. Prosthetic graft infections are rare complications, but prevention of bacteremia is imperative; administration of antibiotics perioperatively and treatment of secondary infections is indicated. Long-term arterial graft complications such as stenosis or pseudoaneurysms are uncommon, and routine graft surveillance rarely is performed. Consequently, long-term administration of antiplatelet agents or antithrombotics is not routine.

**Genitourinary Tract.** Historically, when undergoing laparotomy for trauma, all penetrating wounds to the kidneys were explored; recently, routine exploration of Gerota’s fascia has been questioned. Parenchymal renal injuries are treated with hemostatic and reconstructive techniques similar to those used for injuries of the liver and spleen: topical methods (electrocautery; argon beam coagulation; application of thrombin-soaked gelatin foam sponge, fibrin glue, or BioGlue) and pledged suture repair. However, two caveats are recognized: the collecting system should be closed separately, and the renal capsule should be preserved to close over the repair of the collecting system (Fig. 7-69). Renal vascular injuries are common after penetrating trauma and may be deceptive due to tamponade by Gerota’s

---

**Figure 7-69.** When renorrhaphy is undertaken, effective repair is assisted by attention to several key points: **A.** Vascular occlusion controls bleeding and permits adequate visualization. **B.** The renal capsule is carefully preserved. **C.** The collecting system is closed separately with absorbable suture. **D.** The preserved capsule is closed over the collecting system repair.
fascia, which results in delayed hemorrhage. Arterial reconstruction using graft interposition should be attempted within 5 hours of injury for renal preservation. For destructive parenchymal or irreparable renovascular injuries, nephrectomy may be the only option; a normal contralateral kidney must be palpated because unilateral renal agenesis occurs in 0.1% of patients.

Over 90% of blunt renal injuries are treated nonoperatively. Hematuria typically resolves within a few days with bed rest, although rarely bleeding is so persistent that bladder irrigation to displace blood clots is warranted. Persistent gross hematuria may require embolization, whereas urinomas can be drained percutaneously. Operative intervention after blunt trauma is limited to renovascular injuries and destructive parenchymal injuries that result in hypotension. The renal arteries and veins are uniquely susceptible to traction injury caused by blunt trauma. As the artery is stretched, the inelastic intima and media may rupture, which causes thrombus formation and resultant stenosis or occlusion. The success rate for renal artery repair is limited, but an attempt is reasonable if the injury is <5 hours old or if the patient has a solitary kidney or bilateral injuries.140 Image-guided endoluminal stent placement is now employed for many of these injuries recognized by CT scanning. Reconstruction after blunt renal injuries may be difficult, however, because the injury is typically at the level of the aorta. If repair is not possible within this time frame, leaving the kidney in situ does not necessarily lead to the late sequelae of hypertension or abscess formation. The renal vein may be torn or completely avulsed from the vena cava due to blunt trauma. Typically, the large hematoma causes hypotension, which leads to operative intervention. During laparotomy for blunt trauma, expanding or pulsatile perinephric hematomas should be explored. If necessary, emergent vascular control can be obtained by placing a curved vascular clamp across the hilum from an inferior approach. Techniques of repair and hemostasis are similar to those described earlier.

Injuries to the ureters are uncommon but may occur in patients with pelvic fractures and penetrating trauma. An injury may not be identified until a complication (i.e., a urinoma) becomes apparent. If an injury is suspected during operative exploration but is not clearly identified, methylene blue or indigo carmine is administered IV with observation for extravasation. Injuries are repaired using 5-0 absorbable monofilament, and mobilization of the kidney may reduce tension on the anastomosis. Distal ureteral injuries can be treated by reimplantation facilitated with a psoas hitch and/or Boari flap. In damage control circumstances, the ureter can be ligated on both sides of the injury and a nephrostomy tube placed.

Bladder injuries are subdivided into those with intraperitoneal extravasation and those with extraperitoneal extravasation. Ruptures or lacerations of the intraperitoneal bladder are operatively closed with a running, single-layer, 3-0 absorbable monofilament suture. Laparoscopic repair is becoming common in patients not requiring laparotomy for other injuries. Extraperitoneal ruptures are treated nonoperatively with bladder decompression for 2 weeks. Urethral injuries are managed by bridging the defect with a Foley catheter, with or without direct suture repair. Strictures are not uncommon but can be managed electively.

**Female Reproductive Tract.** Gynecologic injuries are rare. Occasionally the vaginal wall will be lacerated by a bone fragment from a pelvic fracture. Although repair is not mandated, it should be performed if physiologically feasible. More important, however, is recognition of the open fracture, need for possible drainage, and potential for pelvic sepsis. Penetrating injuries to the vagina, uterus, fallopian tubes, and ovaries are also uncommon, and routine hemostatic techniques are used. Repair of a transected fallopian tube can be attempted but probably is unjustified because a suboptimal repair will increase the risk of tubal pregnancy. Transection at the injury site with proximal ligation and distal salpingectomy is a more prudent approach.

**Pelvic Fracture Hemorrhage Control**

Patients with pelvic fractures who are hemodynamically unstable are a diagnostic and therapeutic challenge for the trauma team. These injuries often occur in conjunction with other life-threatening injuries, and there is no universal agreement among clinicians on management. Current management algorithms in the United States incorporate variable time frames for bony stabilization and fixation, as well as hemorrhage control by preperitoneal pelvic packing and/or angiembolization. Early institution of a multidisciplinary approach with the involvement of trauma surgeons, orthopedic surgeons, interventional radiologists, the director of the blood bank, and anesthesiologists is imperative due to high associated mortality rates (Fig. 7-70).

Evaluation in the ED focuses on identification of injuries mandating operative intervention (e.g., massive hemothorax or hemoperitoneum) and injuries related to the pelvic fracture that alter management (e.g., injuries to the iliac artery, rectum, urethra, or bladder). Immediate temporary stabilization with sheeting of the pelvis or application of commercially available compression devices should be performed in hemodynamically unstable patients. In patients with profound shock who are at high risk based upon mechanism (e.g., autopedestrian accident), pelvic stabilization should be done before radiographic confirmation. If the patient’s primary source of bleeding is the fracture-related pelvic hematoma, several options exist for hemorrhage control. Because 85% of bleeding due to pelvic fractures is venous or bony in origin, the authors advocate immediate external fixation and preperitoneal pelvic packing.141 Anterior external fixation decreases pelvic volume, which promotes tamponade of venous bleeding and prevents secondary hemorrhage from the shifting of bony elements. Pelvic packing, in which six laparotomy pads (four in children) are placed directly into the paravesical space through a small anterior suprapubic incision, provides tamponade for the bleeding (Fig. 7-71). Pelvic packing also eliminates the often difficult decision by the trauma surgeon: OR vs. interventional radiology? All patients can be rapidly transported to the OR, and packing can be accomplished in under 30 minutes. In the authors’ experience, this results in hemodynamic stability and abrupt cessation of the need for ongoing blood transfusion in the majority of cases.141 Patients also can undergo additional procedures such as laparotomy, thoracotomy, external fixation of extremity fractures, fasciotomy, revascularization, or craniotomy. Following pelvic packing, angiography is reserved for patients with evidence of ongoing pelvic bleeding after admission to the SICU (>4 units of RBCs in the first 12 postoperative hours after the coagulopathy is corrected). Patients undergo standard posttrauma resuscitative SICU care, and the pelvic packs are removed within 48 hours; prior to unpacking, however, the patient’s coagulopathy should be corrected to prevent the need for repacking of the pelvic space. Repacking of the pelvic space is associated with an infection rate of 47% and should be avoided; directed hemostasis with topical agents, suture repair, or electrocauterity at pelvic pack removal should be performed.

Another clinical challenge is the open pelvic fracture. In many instances, the wounds are located in the perineum, and the risk of pelvic sepsis and osteomyelitis is high. To reduce
the risk of infection, performance of a diverting sigmoid colostomy is recommended. The pelvic wound is manually debrided and then irrigated daily with a high-pressure pulsatile irrigation system until granulation tissue covers the wound. The wound is then left to heal by secondary intention with a wound vacuum-assisted wound closure (VAC) device.

Extremity Vascular Injuries, Fractures, and Compartment Syndromes

Patients with injured extremities often require a multidisciplinary approach with involvement of trauma, orthopedic, and plastic surgeons to address vascular injuries, fractures, soft tissue injuries, and compartment syndromes. Immediate stabilization of fractures or unstable joints is done in the ED using Hare traction, knee immobilizers, or plaster splints. In patients with open fractures, the wound should be covered with povidone-iodine (Betadine)-soaked gauze and antibiotics administered. Options for fracture fixation include external fixation or open reduction and internal fixation with plates or intramedullary nails. Vascular injuries, either isolated or in combination with fractures, require emergent repair. Common combined injuries include clavicle/first rib fractures and subclavian artery injuries, dislocated shoulder/proximal humeral fractures and axillary artery injuries, supracondylar fractures/elbow dislocations and brachial artery injuries, femur fracture and superficial femoral artery injuries, and knee dislocation and popliteal vessel injuries. On-table angiography in the OR facilitates rapid intervention and is warranted in patients with evidence of limb threat on arrival. Arterial access for on-table lower extremity angiography can be obtained percutaneously at the femoral vessels with a standard arterial catheter, via femoral vessel exposure and direct cannulation, or with superficial femoral artery (SFA) exposure just above the medial knee. Controversy exists regarding which should be done first, fracture fixation or arterial repair. The authors prefer placement of temporary intravascular shunts first to reestablish arterial flow and minimize ischemia during fracture treatment, with definitive vascular repair following. Rarely, immediate amputation may be considered due to the severity of orthopedic and neurovascular injuries. This is particularly true if primary nerve transection is present in addition to fracture and arterial injury. Collaborative decision making by the trauma, orthopedic, and plastic/reconstructive team is essential.

Figure 7-70. Management algorithm for patients with pelvic fractures with hemodynamic instability. CT = computed tomography; ED = emergency department; FAST = focused abdominal sonography for trauma; HD = hemodynamic; PLT = platelets; PRBCs = packed red blood cells; SICU = surgical intensive care unit.
Operative intervention for vascular injuries should follow standard principles of repair (see “Vascular Repair Techniques”). For subclavian or axillary artery repairs, 6-mm PTFE graft or RSVG are used depending on the location. Because associated injuries of the brachial plexus are common, a thorough neurologic examination of the extremity is mandated before operative intervention. Operative approach for a brachial artery injury is via a medial upper extremity longitudinal incision; proximal control may be obtained at the axillary artery, and an S-shaped extension through the antecubital fossa provides access to the distal brachial artery. The injured vessel segment is excised, and an end-to-end interposition RSVG graft is performed. Upper extremity fasciotomy is rarely required because of the rich collateral perfusion via the profunda. For SFA injuries, external fixation of the femur typically is performed, followed by end-to-end RSVG of the injured SFA segment. Close monitoring for calf compartment syndrome is mandatory. Preferred access to the popliteal space for an acute injury is the medial, one-incision approach with detachment of the semitendinosus, semimembranosus, and gracilis muscles (Fig. 7-72). Another option is a medial approach

Figure 7-71. A. Pelvic packing is performed through a 6- to 8-cm midline incision made from the pubic symphysis cephalad, with division of the midline fascia. B. The pelvic hematoma often dissects the preperitoneal and paravesical space down to the presacral region, which facilitates packing; alternatively, blunt digital dissection opens the preperitoneal space for packing. C. Three standard surgical laparotomy pads are placed on each side of the bladder, deep within the preperitoneal space; the fascia is closed with an O polydioxanone monofilament suture and the skin with staples.

Figure 7-72. A. The popliteal space is commonly accessed using a single medial incision (the detached semitendinosus, semimembranosus, and gracilis muscles are identified by different suture types). B. Alternatively, a medial approach with two incisions may be used. Insertion of a Pruitt-Inahara shunt (arrow) provides temporary restoration of blood flow, which prevents ischemia during fracture treatment.
with two incisions using a longer RSVG, but this requires interval ligation of the popliteal artery and geniculate branches. Rarely, with open wounds a straight posterior approach with an S-shaped incision can be used. If the patient has an associated popliteal vein injury, this should be repaired first with a PTFE interposition graft while the artery is shunted. For an isolated popliteal artery injury, RSVG is performed with an end-to-end anastomosis. Compartment syndrome is common, and presumptive four-compartment fasciotomies are warranted in patients with combined arterial and venous injury. Once the vessel is repaired and restoration of arterial flow documented, completion angiography should be done in the OR if there is no palpable distal pulse. Vasoparalysis with verapamil, nitroglycerin, and papaverine may be used to treat vasoconstriction (Table 7-11).

Compartment syndromes, which can occur anywhere in the extremities, involve an acute increase in pressure inside a closed space, which impairs blood flow to the structures within. Causes of compartment syndrome include arterial hemorrhage into a compartment, venous ligation or thrombosis, crush injuries, and reperfusion injury. In conscious patients, pain is the prominent symptom, and active or passive motion of muscles in the involved compartment increases the pain. Paresthesias may also be described. In the lower extremity, numbness between the first and second toes is the hallmark of early compartment syndrome in the exquisitely sensitive anterior compartment and its enveloped deep peroneal nerve. Progression to paralysis can occur, and loss of pulses is a late sign. In comatose or obtunded patients, the diagnosis is more difficult to secure. In patients with a compatible history and a tense extremity, compartment pressures should be measured with a hand-held Stryker device. Fasciotomy is indicated in patients with a gradient of <30 mmHg (gradient = diastolic pressure – compartment pressure), an absolute compartment pressure >30 mmHg, ischemic periods of >6 hours, or combined arterial and venous injuries. The lower extremity is most frequently involved, and compartment release is performed using a two-incision, four-compartment fasciotomy (Fig. 7-73). Of note, the soleus muscle must be detached from the tibia to decompress the deep flexor compartment.

### SURGICAL INTENSIVE CARE MANAGEMENT

#### Postinjury Resuscitation

ICU management of the trauma patient, either with direct admission from the ED or after emergent operative intervention, is considered in distinct phases because there are differing goals and priorities. The period of acute resuscitation, typically lasting for the first 12 to 24 hours after injury, combines several key principles: optimizing tissue perfusion, ensuring normothermia, and restoring coagulation status. There are a multitude of management algorithms aimed at accomplishing these goals, the majority of which involve goal-directed resuscitation with initial volume loading to attain adequate preload, followed by judicious use of inotropic agents or vasopressors. Although the optimal hemoglobin level remains debated, during shock resuscitation a hemoglobin level of >10 g/dL is generally accepted to optimize hemostasis and ensure adequate oxygen delivery. After the first 24 hours of resuscitation, a more judicious transfusion trigger of a hemoglobin level of <7 g/dL in the euvolemic patient limits the adverse inflammatory effects of stored RBCs. Recent trends have focused on limiting crystalloid loading. In fact, optimizing crystalloid administration is a challenging aspect of early care (i.e., balancing cardiac performance against generation of an abdominal compartment syndrome and generalized tissue edema). Although early colloid administration is appealing, evidence to date does not support this concept.

Invasive monitoring with pulmonary artery catheters has been supplanted with specialized catheters that measure arterial pulse contour analysis; in mechanically ventilated patients, stroke volume (SV) and continuous cardiac output can be measured. A patient’s volume status may be ascertained by measuring the SV following either 10 cc/kg volume bolus or following a passive leg raise, which augments preload by 250 to 500 mL. A change in SV of ≥10% suggests preload responsiveness, and additional resuscitation fluid should be given. Although norepinephrine is the agent of choice for patients with low systemic vascular resistance who are unable to maintain a mean arterial pressure of >60 mmHg, patients may have an element of myocardial dysfunction requiring inotropic support. The role of relative adrenal insufficiency is another controversial area.

Optimal early resuscitation is mandatory and determines when the patient can (a) undergo additional necessary imaging, and (b) be returned to the OR after initial damage control surgery for definitive repair of injuries. Specific goals of resuscitation before repeated “semi-elective” transport include a core temperature of >35°C (95°F), base deficit of <6 mmol/L, and normal coagulation indices. Although correction of metabolic acidosis is desirable, how quickly this should be accomplished requires careful consideration. Adverse sequelae of excessive crystalloid resuscitation include increased intracranial pressure, worsening pulmonary edema, and intra-abdominal visceral and retroperitoneal edema resulting in secondary abdominal compartment syndrome. Therefore, it should be the overall trend of the resuscitation rather than a rapid reduction of the base deficit that is the goal. The goal is to normalize lactate within 24 hours.

In general, wounds sustained from trauma should be examined daily for progression of healing and signs of infection. Complex soft tissue wounds of the abdomen, such as degloving injuries after blunt trauma (termed Morel-Lavallée lesions), shotgun wounds, and other destructive blast injuries, are particularly difficult to manage. Following initial debridement of devitalized tissue, wound care includes wet-to-dry dressing changes twice daily or application of a VAC device. Repeated operative debridement may be necessary, and early involvement of the reconstructive surgery service for possible flap coverage is advised. Midline laparotomy wounds are inspected 48 hours postoperatively by removing the sterile surgical dressing. If an ileostomy or colostomy is required, one should inspect it...
daily to ensure that it is viable. If the patient develops a high-grade fever, the wound should be inspected sooner to exclude an early necrotizing infection. If a wound infection is identified— as evidenced by erythema, pain along the wound, or purulent drainage—the wound should be widely opened by removing skin staples. After ensuring that the midline fascia is intact with digital palpation, the wound is initially managed with wet-to-dry dressing changes. The most common intra-abdominal complications are anastomotic failure and abscess. The choice between percutaneous and operative therapy is based on the location, timing, and extent of the collection.

**Abdominal Compartment Syndrome**

The abdominal compartment syndrome is classified as pathologic intra-abdominal hypertension due to intra-abdominal injury (primary) or splanchnic reperfusion after massive resuscitation (secondary). Secondary abdominal compartment syndrome may result from any condition requiring extensive crystalloid resuscitation, including extremity trauma, chest trauma, or even postinjury sepsis. The sources of increased intra-abdominal pressure include bowel edema, ascites, bleeding, and packs. A diagnosis of intra-abdominal hypertension cannot reliably be made by physical examination; therefore, it is obtained by measuring the intraperitoneal pressure. The most common technique is to measure the patient’s bladder pressure. Fifty milliliters of saline is instilled into the bladder via the aspiration port of the Foley catheter with the drainage tube clamped, and a three-way stopcock and water manometer is placed at the level of the pubic symphysis. Bladder pressure is then measured on the manometer in centimeters of water (Table 7-12) and

Figure 7-73. A. The anterior and lateral compartments are approached from a lateral incision, with identification of the fascial raphe between the two compartments. Care must be taken to avoid the superficial peroneal nerve running along the raphe. B. To decompress the deep flexor compartment, which contains the tibial nerve and two of the three arteries to the foot, the soleus muscle must be detached from the tibia.
Table 7-12
Abdominal compartment syndrome grading system

<table>
<thead>
<tr>
<th>GRADE</th>
<th>BLADDER PRESSURE</th>
<th>mmHg</th>
<th>cmH₂O</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>10–15</td>
<td>13–20</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>16–25</td>
<td>21–35</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>26–35</td>
<td>36–47</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>&gt;35</td>
<td>&gt;48</td>
<td></td>
</tr>
</tbody>
</table>

Increased abdominal pressure affects multiple organ systems (Fig. 7-74). Abdominal compartment syndrome, as noted earlier, is defined as intra-abdominal hypertension sufficient to produce physiologic deterioration and manifests via such end-organ sequelae as decreased urine output, increased pulmonary inspiratory pressures, decreased cardiac preload, and decreased cardiac output. Because any of these clinical symptoms of abdominal compartment syndrome may be attributed to the primary injury, a heightened awareness of this syndrome must be maintained. Organ failure can occur over a wide range of recorded bladder pressures. Generally, no specific bladder pressure prompts therapeutic intervention, except when the pressure is >35 mmHg. Rather, emergent decompression is carried out when intra-abdominal hypertension reaches a level at which end-organ dysfunction occurs. Mortality is directly affected by the timing of decompression, with 70% mortality in patients with a delay in decompression, and nearly uniform mortality in those not undergoing decompression. Decompression is usually performed operatively, either in the ICU if the patient is hemodynamically unstable or in the OR. ICU bedside laparotomy is easily accomplished, avoids transport of hemodynamically compromised patients, and requires minimal equipment (e.g., scalpel, suction device, cautery, and dressings for temporary abdominal closure). In patients with significant intra-abdominal fluid as the primary component of abdominal compartment syndrome, rather than bowel or retroperitoneal edema, decompression can be accomplished effectively via a percutaneous drain. These patients are identified by bedside ultrasound, and the morbidity of a laparotomy is avoided. When operative decompression is required with egress of the abdominal contents, temporary coverage is obtained using a fenestrated subfascial 45 × 60 cm sterile drape and Ioban application (see Fig. 7-49).

The performance of damage control surgery and recognition of abdominal compartment syndrome have dramatically improved patient survival, but at the cost of an open abdomen. Several management points deserve attention. Despite having a widely open abdomen, patients can develop recurrent abdominal compartment syndrome, which increases their morbidity and mortality; therefore, bladder pressure should be monitored every 4 hours, with significant increases in pressures alerting the clinician to the possible need for repeat operative decompression. Patients with an open abdomen lose between 500 and 2500 mL per day of abdominal effluent. Appropriate volume compensation for this albumin-rich fluid remains controversial, with regard to both the amount administered (replacement based on clinical indices vs. routine 0.5 mL replacement for every milliliter lost) as well as the type of replacement (crystalloid vs. colloid).

Following resuscitation and management of specific injuries, the goal of the operative team is to close the abdomen as quickly as possible. Multiple techniques have been introduced to obtain fascial closure of the open abdomen to minimize morbidity and cost of care. Historically, for patients who could not be closed at repeat operation, approximation of the fascia with mesh (prosthetic or biologic) was used, with planned reoperation. Another option was split-thickness skin grafts applied directly to the exposed bowel for coverage; removal of the skin...
grids was planned 9 to 12 months after the initial surgery, with definitive repair of the hernia by component separation. However, delayed abdominal wall reconstruction was resource intensive, with considerable patient morbidity. The advent of VAC technology has revolutionized fascial closure. The authors currently use a sequential closure technique with the wound VAC device that is based on constant fascial tension and return to the OR every 48 hours until closure is complete (Fig. 7-75). The authors’ success rate with this approach exceeds 95%. This is important because among patients not attaining fascial closure, 20% suffer GI tract complications that prolong their hospital course. These include intra-abdominal abscess, enteric fistula, and bowel perforations (Fig. 7-76). Management requires frequent operative or percutaneous drainage of abscesses, control of fistulas, and prolonged nutritional support.

**SPECIAL POPULATIONS**

**Pregnant Patients**

During pregnancy, 7% of women are injured. Motor vehicle collisions and falls are the leading causes of injury, accounting for 70% of cases. Fetal death after trauma most frequently occurs after motor vehicle collisions, but only 11% of fetal deaths are due to the death of the mother; therefore, early trauma resuscitation and management is directed not only at the mother but also at the fetus. Domestic violence is also common, affecting between 10% and 30% of pregnant women and resulting in fetal mortality of 5%.

Pregnancy results in physiologic changes that may impact postinjury evaluation (Table 7-13). Heart rate increases by 10 to 15 beats per minute during the first trimester and remains elevated until delivery. Blood pressure diminishes during the first two trimesters due to a decrease in systemic vascular resistance and rises again slightly during the third trimester (mean values: first = 105/60, second = 102/55, third = 108/67). Intra-vascular volume is increased by up to 8 L, which results in a relative anemia but also a relative hypervolemia. Consequently, a pregnant woman may lose 35% of her blood volume before exhibiting signs of shock. Pregnant patients have an increase in tidal volume and minute ventilation but a decreased functional residual capacity; this results in a diminished PCO₂ and respiratory alkalosis. Also, pregnant patients may desaturate more rapidly, particularly in the supine position and during intubation. Supplemental oxygen is always warranted in the trauma patient but is particularly critical in the injured pregnant patient because the oxygen dissociation curve is shifted to the left for the fetus compared to the mother (i.e., small changes in maternal oxygenation result in larger changes for the fetus because the fetus is operating in the steep portion of the dissociation curve). Anatomic changes contribute to these pulmonary functional alterations and are relevant in terms of procedures. With the gravid uterus enlarged, DPL should be performed in a supraumbilical site with the catheter directed cephalad. In addition, the upward pressure on the diaphragm calls for caution when placing a thoracostomy tube; standard positioning may result in an infra-abdominal location or perforation of the diaphragm.

Other physiologic changes during pregnancy affect the GI, renal, and hematologic systems. The lower esophageal sphincter has decreased competency, which increases the risk for aspiration. Liver function test values increase, with the alkaline phosphatase level nearly doubling. The high levels of progesterone impair gallbladder contractions, which results in bile stasis and an increased incidence of gallstone formation; this may not affect the trauma bay evaluation but becomes important in a prolonged ICU stay. Plasma albumin level decreases from a normal of around 4.3 g/dL to an average of 3.0 g/dL. Renal blood flow increases by 30% during pregnancy, which causes a decrease in serum level of blood urea nitrogen and creatinine. The uterus may also compress the ureters and bladder, causing hydronephrosis and hydroureter. Finally, as noted earlier there is a relative anemia during pregnancy, but a hemoglobin level of <11 g/dL is considered abnormal. Additional hematologic changes include a moderate leukocytosis (up to 20,000 mm³) and a relative hypercoagulable state due to increased levels of factors VII, VIII, IX, X, and XII and decreased fibrinolytic activity.

During evaluation in the ED, the primary and secondary surveys commence, with mindfulness that the mother always receives priority while conditions are still optimized for the fetus. This management includes provision of supplemental oxygen (to prevent maternal and fetal hypoxia), fluid resuscitation (the hypervolemia of pregnancy may mask signs of shock), and placement of the patient in the left lateral decubitus position (or tilting of the backboard to the left) to avoid caval compression. Assessment of the fetal heart rate is the most valuable information regarding fetal viability. Fetal monitoring should initially be assessed with bedside FAST ultrasound to document the heart rate of the fetus; subsequent monitoring should be performed with a cardiotocographic device that measures both contractions and fetal heart tones (FHTs). Because change in heart rate is the primary response of the fetus to hypoxia or hypotension, anything above an FHT of 160 is a concern, whereas bradycardia (FHT of <120) is considered fetal distress. Indications for emergent cesarean section include: (a) severe maternal shock or impending death (if the fetus is delivered within 5 minutes, survival is estimated at 70%), (b) uterine injury or significant fetal distress (anticipated survival rates of >70% if FHTs are present and fetal gestational age is >28 weeks).

If possible, a member of the obstetrics team should be present during initial evaluation. Vaginal bleeding can signal early cervical dilation and labor, abruptio placentae, or placenta previa. Amniotic sac rupture can result in prolapse of the umbilical cord with fetal compromise. Strong contractions are associated with true labor and should prompt consideration of delivery and resuscitation of the neonate. Focused prenatal history-taking should elicit a history of pregnancy-induced hypertension, gestational diabetes, congenital heart disease, preterm labor, or placental abnormalities. Asking the patient when the baby first moved and if she is currently experiencing movement of the fetus is important. Determining fetal age is key for considerations of viability. Gestational age may be estimated by noting fundal height, with the fundus approximating the umbilicus at 20 weeks and the costal margin at 40 weeks. Discrepancy in dates and size may be due to uterine rupture or hemorrhage.

Initial evaluation for abdominopelvic trauma in pregnant patients should proceed in the standard manner. Ultrasound (FAST) of the abdomen should evaluate the four windows (pericardial, right and left upper quadrant, and bladder) and additionally assess FHTs, fetal movement, and sufficiency of amniotic fluid. DPL can be performed in pregnant women via a supraumbilical, open technique. Trauma radiography of pregnant patients presents a conundrum. Radiation damage has three distinct phases of damage and effect: preimplantation, during the period of organogenesis from 3 to 16 weeks, and after 16 weeks. Generally, it is accepted that “safe” doses of radiation
Figure 7-75. The authors’ sequential closure technique for the open abdomen. A. Multiple white sponges (solid arrow), stapled together, are placed on top of the bowel underneath the fascia. Interrupted No. 1 polydioxanone sutures are placed approximately 5 cm apart (dashed arrow), which puts the fascia under moderate tension over the white sponge. B. After the sticky clear plastic vacuum-assisted closure (VAC) dressing is placed over the white sponges and adjacent 5 cm of skin, the central portion is removed by cutting along the wound edges. C and D. Black VAC sponges are placed on top of the white sponges and plastic-protected skin with standard occlusive dressing and suction. E. On return to the operating room (OR) 48 hours later, fascial sutures are placed from both the superior and inferior directions until tension precludes further closure; skin is closed over the fascial closure with skin staples. F. White sponges (fewer in number) are again applied and fascial retention sutures are placed with planned return to the OR in 48 hours.
Complications after split-thickness skin graft closure of the abdomen include enterocutaneous fistulas (intubated here with a red rubber catheter) (A; arrow), rupture of the graft with exposure of the bowel mucosa (B), enteroatmospheric fistulas with a large ventral hernia and loss of abdominal domain (C).

Figure 7-76. Complications after split-thickness skin graft closure of the abdomen include enterocutaneous fistulas (intubated here with a red rubber catheter) (A; arrow), rupture of the graft with exposure of the bowel mucosa (B), enteroatmospheric fistulas with a large ventral hernia and loss of abdominal domain (C).

loss may be related to both maternal shock and direct injury to the uterus or fetal head. Penetrating injuries in this patient population also carry a high risk. The gravid uterus is a large target, and any penetrating injury to the abdomen may result in fetal injury depending on trajectory and uterine size. Gunshot wounds to the abdomen are associated with a 70% injury rate to the uterus and 35% mortality rate of the fetus. If the bullet traverses the uterus and the fetus is viable, cesarean section should be performed. On the other hand, stab wounds do not often penetrate the thick wall of the uterus.

Any patient with a viable pregnancy should be monitored after trauma, with the length of monitoring determined by the injury mechanism and patient physiology. Patients who are
symptomatic, defined by the presence of uterine irritability or contractions, abdominal tenderness, vaginal bleeding, or blood pressure instability, should be monitored in the hospital for at least 24 hours. In addition, patients at high risk for fetal loss (those experiencing vehicle ejection or involved in motorcycle or pedestrian collisions and those with maternal tachycardia, Injury Severity Score of >9, gestational period of >35 weeks, or history of prior assault) also warrant careful monitoring. Patients without these risk factors who are asymptomatic can be monitored for 6 hours in the ED and sent home if no problems develop. They should be counseled regarding warning signs that mandate prompt return to the ED.

Geriatric Patients
Elderly trauma patients (>65 years of age) are hospitalized twice as often as those in any other age group, and this population accounts for one quarter of all trauma admissions. Although the physiology of aging separates older trauma patients from the younger generation (Table 7-14), treatment must remain individualized (some octogenarians look and physiologically act 50 years old, whereas others appear closer to 100 years). No chronologic age is associated with a higher morbidity or mortality, but a patient’s comorbidities do impact the individual’s postinjury course and outcome. For example, recognition that a patient is taking β-blockers affects the physician’s evaluation of vital signs in the ED and impacts treatment course in the ICU. Early monitoring of arterial blood gas values will identify occult shock. A base deficit of >6 mmol/L is associated with a twofold higher risk of mortality in patients over the age of 55 than in younger patients (67% vs. 30%).

Although the published literature on geriatric traumatic brain injury is relatively sparse and uncontrolled with regard to management, some interesting points are noted. First, as anticipated, outcomes are worse in this age group than in their younger counterparts. Based on data from the Traumatic Coma Databank, mortality in patients with severe head injury more than doubles after the age of 55. Moreover, 25% of patients with a normal GCS score of 15 had intracranial bleeding, with an associated mortality of 50%. Just as there is no absolute age that predicts outcome, admission GCS score is a poor predictor of individual outcome. Therefore, the majority of trauma centers advocate an initial aggressive approach with reevaluation at the 72-hour mark to determine subsequent care.

One of the most common sequelae of blunt thoracic trauma is rib fractures. In the aging population, perhaps due to osteoporosis, less force is required to cause a fracture. In fact, in one study, 50% of patients >65 years old sustained rib fractures from a fall of <6 ft, compared with only 1% of patients <65 years of age. Concurrent pulmonary contusion is noted in up to 35% of patients, and pneumonia complicates the injuries in 10% to 30% of patients with rib fractures, not surprisingly leading to longer ICU stays. Additionally, mortality increases linearly with the number of rib fractures. Patients who sustain more than six rib fractures have pulmonary morbidity rates of >50% and overall mortality rates of >20%.

Chronologic age is not the best predictor of outcome, but the presence of preexisting conditions, which affect a patient’s physiologic age, is associated with increased mortality rates. Injury Severity Score is probably the best overall predictor of patient outcome in the elderly; however, for any given individual its sensitivity may not be precise, and there is a time delay in obtaining sufficient information to calculate the final score. In addition to preexisting conditions and severity of injury, the occurrence of complications compounds the risk for mortality.

### Table 7-13
**Physiologic effects of pregnancy**

**Cardiovascular**
- Increase in heart rate by 10–15 bpm
  - (a) Decreased systemic vascular resistance resulting in:
  - (b) Increased intravascular volume
- Decreased blood pressure during the first two trimesters

**Pulmonary**
- Elevated diaphragm
- Increased tidal volume
- Increased minute ventilation
- Decreased functional residual capacity

**Hematopoietic**
- Relative anemia
- Leukocytosis

**Hypercoagulability**
- (a) Increased levels of factors VII, VIII, IX, X, XII
- (b) Decreased fibrinolytic activity

**Other**
- Decreased competency of lower esophageal sphincter
- Increased enzyme levels on liver function tests
- Impaired gallbladder contractions
- Decreased plasma albumin level
- Decreased blood urea nitrogen and creatinine levels
- Hydronephrosis and hydroureter

### Table 7-14
**Physiologic effects of aging**

**Cardiovascular**
- Atherosclerotic disease that limits cardiac response to stress
- Progressive stiffening and loss of elasticity of the myocardium
- Diminished stroke volume, systolic contraction, and diastolic relaxation
- Decrease in cardiac output of 0.5% per year
- Thickening and calcification of the cardiac valves, which results in valvular incompetence

**Pulmonary**
- Loss of compliance
- Progressive loss of alveolar size and surface area
- Air trapping and atelectasis

**Intracranial**
- Loss of cerebral volume, resulting in:
  - (a) Increased risk of tearing of bridging veins with smaller injuries
  - (b) Accumulation of a significant amount of blood before symptoms occur
  - (c) Muscle loss
- Senescence of the senses

**Other**
- Decline in creatinine clearance by 80%–90%
- Osteoporosis, which causes a greater susceptibility to fractures
Pediatric Patients
Twenty million children, or almost one in four children, are injured each year, with an associated cost of treating the injured child of $16 billion per year. Injury is the leading cause of death among children over the age of 1 year, with 15,000 to 25,000 pediatric deaths per year. Disability after traumatic injury is more devastating, with rates 3 to 10 times that of the death rate. Pediatric trauma involves different mechanisms, different constellations of injury, and the potential for long-term problems related to growth and development. As with adult trauma, over 85% of pediatric trauma has a blunt mechanism, with boys injured twice as often as girls.155 Falls are the most common cause of injury in infants and toddlers. In children, bicycle mishaps are the most common cause of severe injury, whereas motor vehicle-related injury predominates in adolescence. Although unintentional injuries are by far the most common type of injuries in childhood, the number of intentional injuries, such as firearm-related injury and child abuse, is increasing.

ED preparation for the pediatric trauma patient includes assembling age-appropriate equipment (e.g., intubation equipment; IV catheters, including intravenous needles and 4F single-lumen lines), laying out the Broselow Pediatric Emergency Tape (which allows effective approximation of the patient’s weight, medication doses, size of endotracheal tube, and chest tube size), and turning on heat lamps. Upon the pediatric patient’s arrival, the basic tenets of the ABCs apply, with some caveats. In children, the airway is smaller and more cephalad in position compared with that of adults, and in children younger than 10 years, the larynx is funnel shaped rather than cylindrical as in adults. Additionally, the child’s tongue is much larger in relation to the oropharynx. Therefore, a small amount of edema or obstruction can significantly reduce the diameter of the airway (thus increasing the work of breathing), and the tongue may posteriorly obstruct the airway, causing intubation to be difficult. During intubation, a Miller (straight) blade rather than a Macintosh (curved) blade may be more effective due to the acute angle of the cephalad, funnel-shaped larynx. Administration of atropine before rapid-sequence intubation will prevent bradycardia. Adequate ventilation is critical because oxygen consumption in infants and young children is twice that in adults; onset of hypoxemia, followed by cardiac arrest, may be precipitous. Because gastric distension can inhibit adequate ventilation, placement of a nasogastric tube may facilitate effective gas exchange. Approximately one-third of preventable deaths in children are related to airway management; therefore, if airway control cannot be obtained using a standard endotracheal method, surgical establishment of an airway should be considered. In children older than 11 years, standard cricothyroidotomy is performed. Due to the increased incidence of subglottic stenosis in younger patients, needle cricothyroidotomy with either a 14- or 16-gauge catheter is advocated, although it is rarely used. Alternatively, tracheostomy may be performed. In children, the standard physiologic response to hypovolemia is peripheral vasoconstriction and reflex tachycardia; this may mask significant hemorrhagic injury because children can compensate for up to a 25% loss of circulating blood volume with minimal external signs. “Normal” values for vital signs should not necessarily make one feel more secure about the child’s volume status. Volume restoration is based on the child’s weight; two to three boluses of 20 mL/kg of crystalloid is appropriate. Hypotension in children may be due to TBI rather than hypovolemia and should be considered in the appropriate clinic scenario.156

After initial evaluation based on the trauma ABCs, identification and management of specific injuries proceeds. Acute traumatic brain injury is the most common cause of death and disability in any pediatric age group. Although falls are the most common mechanism overall, severe brain injury most often is due to child abuse (in children <2 years) or motor vehicle collisions (in those >2 years). Head CT should be performed to determine intracranial pathology, followed by skull radiography to diagnose skull fractures. As in adults, CPP is monitored, and appropriate resuscitation is critical to prevent the secondary insults of hypoxemia and hypovolemia. Although some data indicate that the pediatric brain recovers from traumatic injury better than the adult brain, this advantage may be eliminated if hypotension is allowed to occur.

As is true in adults, the vast majority of thoracic trauma is also blunt. However, because a child’s skeleton is not completely calcified, it is more pliable. Significant internal organ damage may occur without overlying bony fractures. For example, adult patients with significant chest trauma have a 70% incidence of rib fractures, whereas only 40% of children with significant chest trauma do. Pneumothorax is treated similarly in the pediatric population; patients who are asymptomatic with a pneumothorax of <15% are admitted for observation, whereas those who have a pneumothorax of >15% or who require positive pressure ventilation undergo tube decompression. Presence of a hemothorax in this age group may be particularly problematic because the child’s chest may contain his or her entire blood volume. If the chest tube output is initially 20% of the patient’s blood volume (80 mL/kg) or is persistently >1 to 2 mL/kg per hour, thoracotomy should be considered. Aortic injuries are rare in children, and tracheobronchial injuries are more amenable to nonoperative management. Penetrating thoracic trauma, although uncommon, has 35% operative intervention rate, which is considerably higher than that of the adult population.157 Thoracic injuries are second only to brain injuries as the main cause of death according to the National Pediatric Trauma Registry; however, the overall mortality rate of 15% correlates with the levels in many adult studies.

The evaluation for abdominal trauma in the pediatric patient is similar to that in the adult. FAST is valid in the pediatric age group to detect intra-abdominal fluid. The mechanism of injury often correlates with specific injury patterns. A child sustaining a blow to the epigastrium (e.g., hitting the handlebars during a bike accident) should be evaluated for a duodenal hematoma and/or a pancreatic transection. After a motor vehicle collision in which the patient was wearing a passenger restraint, injuries comprising the “lap belt complex” or “seat belt syndrome” (i.e., abdominal wall contusion, small bowel perforation, flexion-distraction injury of the lumbar spine, diaphragm rupture, and occasionally abdominal aortic dissection) may exist. Nonoperative management of solid organ injuries, first used in children, is the current standard of care in the hemodynamically stable patient. If the patient shows clinical deterioration or hemodynamic lability, has a hollow viscus injury, or requires >40 mL/kg of packed RBCs, continued nonoperative management is not an option. Success rates of nonoperative management approach 95%,158 with an associated 10% to 23% transfusion rate. Findings of a hepatic or splenic blush on CT imaging does not uniformly require intervention; patient physiology should dictate embolization or operative intervention.159
REFERENCES

Entries highlighted in bright blue are key references.


77. Mathew S, Smith BP, Cannon JW, Reilly PM, Schwab CW, Seamon MJ. Temporary arterial shunts in damage


PART I BASIC CONSIDERATIONS


156. Gardner AR, Diz DI, Tooze JA, Miller CD, Petty J. Injury patterns associated with hypotension in pediatric trauma patients:


Surgical care of the burned patient has evolved into a specialized field incorporating the interdisciplinary skills of burn surgeons, nurses, burn therapists, and other healthcare specialists. However, recent mass casualty events have been a reminder that healthcare systems may be rapidly pressed to care for large numbers of burn patients. Naturally, general surgeons may be at the forefront in these events, so it is crucial that they are comfortable with the care of burned patients and well equipped to provide the standard of care.

**BACKGROUND**

Burn injury historically carried a poor prognosis. With advances in fluid resuscitation and the advent of early excision of the burn wound, survival has become an expectation even for patients with severe burns. Continued improvements in critical care and progress in skin bioengineering herald a future in which functional and psychologic outcomes are equally important as survival alone. With this shift in priority, the American Burn Association (ABA) has emphasized referral to specialized burn centers after early stabilization. Specific criteria should guide transfer of patients with more complex injuries or other medical needs to a burn center (Table 8-1). The ABA has published standards of care and created a verification process to ensure that burn centers meet those standards. Because of increased prehospital safety measures, burn patients are transferred longer distances for definitive care at regional burn centers; data from one burn center with a particularly wide catchment area confirmed that even transport times averaging several hours did not affect the long-term outcomes of burn patients.

**INITIAL EVALUATION**

Initial evaluation of the burned patient should follow the same initial priorities of all trauma patients and involves four crucial assessments: airway management, evaluation of other injuries, estimation of burn size, and diagnosis of CO and cyanide poisoning. With direct thermal injury to the upper airway or smoke inhalation, rapid and severe airway edema is potentially lethal. Anticipating the need for intubation and establishing an early airway are critical. Signs of impending respiratory compromise include a hoarse voice, wheezing, or stridor; subjective dyspnea is a particularly concerning symptom and should trigger prompt elective endotracheal intubation. Perioral burns and singed nasal hairs alone do not indicate an upper airway injury, but are signs that the oral cavity and pharynx should be further evaluated for mucosal injury. Orotracheal intubation is the preferred method for securing the airway. Nasotracheal intubation may be useful for patients with associated facial trauma when experienced providers are present, but it should be avoided if oral intubation is safe and easy.

Burned patients are trauma patients and evaluated with a primary survey in accordance with Advanced Trauma Life Support guidelines. Concurrently with the primary survey, large-bore peripheral intravenous (IV) catheters should be placed and fluid resuscitation should be initiated; for a burn larger than 40% total body surface area (TBSA), two large-bore IVs are ideal. IV placement through burned skin is safe and effective but requires attention to securing the catheters. Central venous access and intraosseous (IO) access should be considered when peripheral access cannot be easily obtained. Rarely, IV resuscitation is indicated in patients with burns smaller than 15% who can usually hydrate orally. Pediatric patients with burns larger than 15% may require IO access in emergent situations if venous access cannot be attained. An early and comprehensive secondary survey must be performed on all burn patients, but especially those with a history of associated mechanical trauma such as a motor vehicle collision. Also, patients from structural fires in which the manner of egress is not known should be carefully evaluated for injuries from a possible jump or fall. Urgent radiology studies, such as a chest X-ray, should be performed in the emergency department, but nonurgent skeletal evaluation (i.e., extremity X-rays) can be
done in the intensive care unit (ICU) to avoid hypothermia and delayed resuscitation. Hypothermia is a common prehospital complication that contributes to resuscitation failure. Patients should be wrapped with clean blankets in transport. Cooling should be avoided in patients with moderate or large (>20% TBSA) burns. Patients with acute burn injuries should never receive prophylactic antibiotics. This intervention has been clearly demonstrated to promote development of fungal infections and resistant organisms and was abandoned in the mid-1980s. A tetanus booster should be administered in the emergency department depending on patient immunization status.

The importance of pain management for these patients has been widely recognized over the past 25 years. While pain management is a priority for burn patients, it is important to acknowledge the opioid crisis and the recent push toward decreasing opiate use in general. In order to limit opiate-related morbidity, we recommend responsible opiate use in conjunction with multimodal pain control and development of a weaning plan starting at opioid commencement. Clear expectations around pain medication use should be set with patients. Anxiety is another component of the psychological response of burn patients, seen with both wound care and general postinjury hospital course. Benzodiazepines are a staple in the treatment of acute anxiety; however, they can contribute significantly to hospital delirium. We recommend conservative benzodiazepine use to mitigate the effects of anxiety while minimizing delirogenic effects of benzodiazepines.

**Key Points**

1. Follow American Burn Association criteria for referral of a patient to a regional burn center.
2. Never administer prophylactic antibiotics other than tetanus vaccination.
3. Early excision and grafting of full-thickness and deep partial-thickness burns improve outcomes.
4. Intravenous fluid resuscitation for patients with burns >20% of total body surface area (children with burns >15% of total body surface area) should be titrated to mean arterial pressure (MAP) >60 mmHg and appropriate urine output.

Most burn resuscitation formulas estimate fluid requirements based on burn size measured as a percentage of TBSA (%TBSA). The “rule of nines” is a crude but quick and effective method of estimating burn size (Fig. 8-1). In adults, the anterior and posterior trunk each account for 18%, each lower extremity is 18%, each upper extremity is 9%, and the head is 9%. In children under 3 years old, the head accounts for a larger relative surface area and should be taken into account when estimating burn size. For smaller or odd-shaped burns, the “rule of the palm” where the palmar surface of the hand, including the digits, is 1% TBSA is useful. Diagrams such as the Lund and Browder chart give a more accurate accounting of the true burn size in children and adults. The importance of an accurate burn size assessment cannot be overemphasized. Superficial or first-degree burns should not be included when calculating burn size, and thorough cleaning of soot and debris is mandatory to avoid confusing soiled skin with burns. Examination of referral data suggests that physicians inexperienced with burns tend to overestimate the size of small burns and underestimate the size of large burns, with potentially detrimental effects on pretransfer resuscitation.

**BURN CLASSIFICATION**

Burns are commonly classified as thermal, electrical, or chemical burns, with thermal burns consisting of flame, contact, or scald burns. Flame burns are the most common cause for hospital admission of burns, and have the highest mortality. This is primarily related to their association with structural fires and the accompanying inhalation injury and/or CO poisoning.

Electrical burns make up 3% of U.S. hospital admissions but have special concerns, including cardiac arrhythmia and compartment syndrome with concurrent rhabdomyolysis. A baseline ECG is recommended in all patients with an electrical injury, and a normal ECG in a low-voltage injury (<1000 V) may preclude hospital admission. Because compartment syndrome and rhabdomyolysis are common in high-voltage electrical injuries, vigilance must be maintained for neurologic or vascular compromise, and fasciotomies should be performed even in cases of moderate clinical suspicion. Long-term neurologic symptoms and cataract development are not uncommon with high-voltage electrical injuries, and neurologic and ophthalmologic consultation should be obtained to define baseline patient function.

Chemical burns also comprise 3% of admitted burn patients and result in potentially severe burns. Typically, acid chemical burns result in coagulation necrosis and alkali chemical burns cause liquefactive necrosis (with an exception of hydrofluoric acid, which also causes liquefactive necrosis).

---

**Table 8-1**

<table>
<thead>
<tr>
<th>Guidelines for referral to a burn center</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partial-thickness burns greater than 10% TBSA</td>
</tr>
<tr>
<td>Burns involving the face, hands, feet, genitalia, perineum, or major joints</td>
</tr>
<tr>
<td>Third-degree burns in any age group</td>
</tr>
<tr>
<td>Electrical burns, including lightning injury</td>
</tr>
<tr>
<td>Chemical burns</td>
</tr>
<tr>
<td>Inhalation injury</td>
</tr>
<tr>
<td>Burn injury in patients with complicated preexisting medical disorders</td>
</tr>
<tr>
<td>Patients with burns and concomitant trauma in which the burn is the greatest risk. If the trauma is the greater immediate risk, the patient may be stabilized in a trauma center before transfer to a burn center.</td>
</tr>
<tr>
<td>Burned children in hospitals without qualified personnel for the care of children</td>
</tr>
<tr>
<td>Burn injury in patients who will require special social, emotional, or rehabilitative intervention</td>
</tr>
</tbody>
</table>

TBSA = total body surface area.
The most important components of initial therapy are careful removal of the toxic substance from the patient and irrigation of the affected area with water for a minimum of 30 minutes. In cases of exposure to dry chemicals, such as concrete or powdered forms of lye, the substance should be swept from the patient to avoid a thermal reaction with water. The offending agents in chemical burns can be systemically absorbed and may cause specific metabolic derangements. Formic acid has been known to cause hemolysis and hemoglobinuria, and hydrofluoric acid causes hypocalcemia. Hydrofluoric acid is a particularly common offender due to its widespread industrial uses.

Calcium-based therapies are the mainstay of treating hydrofluoric acid burns, with topical application of calcium gluconate onto wounds and IV administration of calcium gluconate for systemic hypocalcemia symptoms. Intra-arterial calcium gluconate infusion provides effective treatment of progressive tissue injury and intense pain. Patients undergoing intra-arterial therapy need continuous cardiac monitoring. Persistent refractory hypocalcemia with electrocardiographic abnormalities may signal the need for emergent excision of the burned areas.

**BURN DEPTH**

Based on the original burn depth classification by Dupuytren in 1832, burn wounds are commonly classified as superficial (first-degree), partial-thickness (second-degree), full-thickness (third-degree), and fourth-degree burns, which affect underlying soft tissue. Fifth-degree burns (through muscle to bone) and sixth degree burns (charring bone) were also described although are less common. Partial-thickness burns are classified as either superficial or deep partial-thickness burns by depth of involved dermis. Clinically, first-degree burns are painful but do not blister, second-degree burns have dermal involvement and are extremely painful with weeping and blisters, and third-degree burns are leathery, painless, and nonblanching. Jackson described three zones of tissue injury following burn injury. The zone of coagulation is the most severely burned portion and is typically in the center of the wound. As the name implies, the affected tissue is coagulated and sometimes frankly necrotic, much like a full thickness burn, and will need excision and grafting. Peripheral to that is a zone of stasis, with variable degrees of vasoconstriction and resultant ischemia, much like a second-degree burn. Appropriate resuscitation and wound care may prevent conversion to a deeper wound, but infection or suboptimal perfusion may result in an increase in burn depth. This is clinically relevant because many superficial partial-thickness burns will heal with nonoperative management, and the majority of deep partial-thickness burns benefit from excision and skin grafting. The outermost area of a burn is called the zone of hyperemia, which will heal with minimal or no scarring and is most like a superficial partial thickness burn or first-degree burn.

Unfortunately, even experienced burn surgeons have limited ability to accurately predict the healing potential of partial-thickness burns soon after injury; one reason is that burn wounds evolve over the 48 to 72 hours after injury. Numerous burn depth assessment tools have been developed with the idea that...
earlier burn depth definition will expedite appropriate surgical decision-making. One of the most effective ways to determine burn depth is full-thickness biopsy, but this has several limitations; not only is the procedure painful and potentially scarring, but accurate interpretation of the histopathology requires a specialized pathologist and may have slow turnaround times. Laser Doppler can measure skin perfusion to predict burn depth with sensitivities and specificity of up to 83% and 97%, respectively. Noncontact ultrasound has been postulated as a painless modality to predict nonhealing wounds and has the advantage of easily performed serial measurements. Unfortunately, none of these newer therapies have proven adequately superior to justify their cost and as yet have not substituted serial examination by experienced burn surgeons.

**PROGNOSIS**

The Baux score (mortality risk equals age plus %TBSA) was used for many years to predict mortality in burns. Analysis of multiple risk factors for burn mortality has validated age and burn size as the strongest predictors of mortality. Advance-ments in burn care have lowered overall mortality to the point that the original Baux score may no longer be accurate. However, the Revised Baux Score, which accounts for age, burn size, and inhalation injury, has been found to be independently associated with mortality. As such, age, burn size, and inhalation injury continue to be the most robust indicators for burn mortality. Age even as a single variable strongly predicts mortality in burns, and in-hospital mortality in elderly burn patients is a function of age regardless of other comorbidities. In nonelderly patients, comorbidities such as preinjury human immunodeficiency virus (HIV), metastatic cancer, and kidney or liver disease may influence mortality and length of stay. A large database study of 68,661 burn patients found that the variables with the highest predictive value for mortality were age, %TBSA, inhalation injury, coexistent trauma, and pneumonia. A more recent study analyzing 506,628 burn inpatients between 1998 and 2008 demonstrated an association between burn size, age, inhalation injury, and mortality. Other factors associated with mortality included African American race, female gender, and treatment in urban private hospitals (as opposed to urban academic hospitals).

Mortality is not the only outcome of interest in the burn population. Burn injury can significantly impact the subsequent quality of life for survivors, including but not limited to appearance, mobility, functional status, and ability to work. One study found that burn injury reduces short term quality of life by 30% and long-term quality of life by approximately 11%. Predictors of poorer physical and mental health 12 months removed from burn injury include older age, female gender, and greater %TBSA burn size. One factor impacting quality of life is itching—a late and bothersome consequence of burn injury that affects both adult and pediatric population. Other factors discussed later in this chapter include hypertrophic scarring, contracture, and heterotopic ossification. Finally, return to work or school has been a useful tool to evaluate recovery and prognosis. A recent meta-analysis found that approximately 28% of burn survivors never return to work. A recent study using an interventional bundle involving the patient, the employee, worker’s compensation, and burn clinic staff demonstrated a return to work rate of 93%. The return to school for pediatric patients is actually very prompt, averaging about 10 days after discharge. However, further study is needed to determine whether attendance and performance suffer despite early reentry to school.

It is important to recognize these potential quality-of-life issues in burn patients and take necessary steps to diminish the impact that burn injury has on quality of life both in the hospital and following discharge.

**RESUSCITATION**

A myriad of formulas exist for calculating fluid needs during burn resuscitation, suggesting that no one formula benefits all patients. The most commonly used formula, the Parkland or Baxter formula, consists of 3 to 4 mL/kg per % burn of Lactated Ringer’s, of which half is given during the first 8 hours after burn and the remaining half is given over the subsequent 16 hours. The most recent American Burn Association consensus formula recommends 2 mL/kg per % burn of Lactated Ringers given the tendency toward excessive fluid administration with the traditional formulas. The concept behind continuous fluid requirements is simple. The burn (and/or inhalation injury) drives an inflammatory response that leads to capillary leak; as plasma leaks into the extravascular space, crystalloid administration maintains the intravascular volume. Therefore, if a patient receives a large fluid bolus in a prehospital setting or emergency department, the fluid has likely leaked into the interstitium, and the patient still requires ongoing burn resuscitation according to the estimates. Continuation of fluid volumes should depend on the time since injury, urine output, and mean arterial pressure (MAP). As the capillary leak closes, the patient will require less volume to maintain these two resuscitation endpoints. Children under 20 kg have the additional requirement that they do not have sufficient glycogen stores to maintain an adequate glucose level in response to the inflammatory response. Specific pediatric formulas have been described, but the simplest approach is to deliver a weight-based maintenance IV fluid with glucose supplementation in addition to the calculated resuscitation with lactated Ringer’s.

It is important to remember that any formula for burn resuscitation is merely a guideline, and fluid must be titrated based on appropriate response to therapy. A number of parameters are widely used to gauge burn resuscitation, but the most common remain the simple outcomes of blood pressure and urine output. As in any critically ill patient, a target MAP of 60 mmHg ensures optimal end-organ perfusion. Goals for urine output should be 30 mL/h in adults and 1 to 1.5 mL/kg per h in pediatric patients. Because blood pressure and urine output may not correlate perfectly with true tissue perfusion, the search continues for other adjunctive parameters that may more accurately reflect adequate resuscitation. Some centers have found serum lactate to be a better predictor of mortality in severe burns, and others have found that base deficit predicts eventual organ dysfunction and mortality. Because burned patients with normal blood pressure and serum lactate levels may have compromised gastric mucosal perfusion, continuous measurement of mucosal pH with its logistical difficulties has garnered limited popularity. Invasive monitoring with pulmonary artery catheters typically results in significant excessive fluid administration without improved cardiac output or preload measurements; use of invasive monitoring seems to have variable effects on long-term outcomes.

Actual administrated fluid volumes typically exceed volumes predicted by standard formulas. One survey of burn
centers showed that 58% of patients end up getting more fluids than would be predicted by Baxter’s formula. Comparison of modern-day patients with historical controls shows that over-resuscitation may be a relatively recent trend. One theory is that increased opioid analgesic use results in peripheral vasodilation and hypotension and the need for greater volumes of bolused resuscitative fluids. A classic study by Navar et al showed that burned patients with inhalation injury required an average of 5.76 mL/kg per % burn, vs. 3.98 mL/kg per % burn for patients without inhalation injury, and this has been corroborated by subsequent studies. Prolonged mechanical ventilation may also play a role in increased fluid needs. A multicenter study found that age, weight, %TBSA, and intubation on admission were significant predictors of more fluid delivery during the resuscitation period. Those patients receiving higher fluid volumes were at increased risk of complications and death. Common complications include abdominal compartment syndrome, extremity compartment syndrome, intraocular compartment syndrome, and pleural effusions. Monitoring bladder pressures can provide valuable information about development of intra-abdominal hypertension.

The use of colloid as part of the burn resuscitation has generated much interest over the years. In late resuscitation when the capillary leak has closed, colloid administration may decrease overall fluid volumes and potentially may decrease associated complications such as intra-abdominal hypertension. A recent meta-analysis accounting for statistical heterogeneity among studies included demonstrated a trend toward mortality benefit for patients receiving albumin. However, albumin use has never been shown to definitively improve mortality in burn patients and has controversial effects on mortality in critically ill patients. Still, many burn centers including ours continue to use albumin as an adjunct during burn resuscitation. Attempts to minimize fluid volumes in burn resuscitation have included study of hypertonic solutions. A recent meta-analysis evaluating hyperosmotic vs. isosmotic fluid resuscitation demonstrates decreased total fluid load (vol/%TBSA per weight) over the first 24 hours with use of hyperosmotic fluid with no difference in total fluid, urine output, creatinine, or mortality. A described downside of hypertonic fluid administration is hyperchloremic acidosis.

Other adjuncts are being increasingly used during initial burn resuscitation. High-dose ascorbic acid (vitamin C) may decrease fluid volume requirements and ameliorate respiratory embarrassment during resuscitation, although no mortality benefit has been noted thus far in two trials. Plasmapheresis has also been associated with decreased fluid requirements and increased urine output in patients who require higher resuscitative volumes than predicted to maintain adequate urine output and MAP. It is postulated that plasmapheresis may filter out inflammatory mediators, thus decreasing ongoing vasodilation and capillary leak.

One adjunct that has found increasing utility in surgical ICUs has been the application of bedside ultrasound. Ultrasound offers the potential to make rapid, noninvasive assessments during acute changes in clinical condition. For burn patients, bedside ultrasonography may be indicated for evaluation of volume status, gross assessment of cardiac function, and diagnosis of pneumothorax. Determining patient cardiac function and volume status may guide fluid resuscitation. Cardiac function can be evaluated with three common heart views: the parasternal long axis, parasternal short axis, and apical four-chamber views. Whereas no study has used ultrasound to guide fluid resuscitation in burn patients, volume status can be estimated by examination of cardiac function and evaluation of the inferior vena cava (IVC) diameter with changes in respiration, as has been done in patients with hemorrhage and shock. Ultrasound also allows timely diagnosis of pneumothorax. A high-frequency probe with an adequate window between ribs permits identification of lung parenchyma against the chest well. A pneumothorax appears as a transition on ultrasound between lung parenchyma, which has a heterogeneous appearance, and air, which has a hypoechoic appearance. Further studies are warranted to identify indications for the use of ultrasound in burned patients.

Machine learning and bedside computer decision support are other adjuncts gaining traction in caring for burn patients. These modalities can enhance patient care and aid in diagnosis, treatment, and research. The use of bedside computer decision support has been particularly appealing for resuscitation of burn patients in the first 48 hours and has been shown to improve fluid management during initial resuscitation.

The role of blood transfusion in critically injured patients has undergone a reevaluation in recent years. Blood transfusions are considered to be immunomodulatory and potentially immunosuppressive, which is one explanation to the links between blood transfusions and increased infection and shorter time to recurrence after oncologic surgery. A large multicenter study of blood transfusions in burn patients found that increased numbers of transfusions were associated with increased infections and higher mortality in burn patients, even when correcting for burn severity. A follow-up study implementing a restrictive transfusion policy in burned children showed that a hemoglobin threshold of 7 g/dL had no more adverse outcomes vs. a traditional transfusion trigger of 10 g/dL. In addition, costs incurred to the institution were significantly less. A recent randomized control trial in patients with >20% TBSA compared outcomes of a restrictive to a liberal red blood cell transfusion strategy (hemoglobin 7–8 vs. 10–11, respectively). There were no differences in blood stream infection, organ dysfunction, ventilator days, time to wound healing, or 30-day mortality between both groups. These data, in concert with other reported complications such as transfusion-related lung injury, have led to recommendations that blood transfusions be used only when there is an apparent physiologic need. Attempts to minimize blood transfusion in nonburned critically ill patients have led to use of erythropoietin by some centers. However, burn patients often have elevated erythropoietin levels, and a randomized study in burn patients showed that recombinant human erythropoietin did not effectively prevent anemia or decrease the number of transfusions given. Promising animal studies demonstrating erythropoietin-mediated prevention of secondary burn progression have yet to be validated in humans.

INHALATION INJURY AND VENTILATOR MANAGEMENT

Inhalation injuries are commonly seen in tandem with burn injuries and are known to increase mortality in burned patients. Smoke inhalation is present in as many as 35% of hospitalized burn patients and may triple the hospital stay compared to isolated burn injuries. Mortality for inhalation injury has been reported to be as high as 25%, with this increasing to 50% in patients with ≥20% TBSA burns. The pneumonia rate in
patients with inhalation injury has been reported to be three times higher than those without inhalation injury, and it has been associated with increased length of stay, increased ventilator days, and need for tracheostomy.91,92 The combination of burns, inhalation injury, and pneumonia increases mortality by up to 60% over burns alone.93 Subsequent development of the adult respiratory distress syndrome (ARDS) is common in these patients and may be caused in part by recruitment of alveolar leukocytes with an enhanced endotoxin-activated cytokine response.94 When ARDS complicates burns and inhalation injury, mortality approaches 66%; in one study, patients with burns ≥60% TBSA in combination with inhalation injury and ARDS had 100% mortality.95

Smoke inhalation causes injury in two ways: by direct heat injury to the upper airways and inhalation of combustion products into the lower airways. Direct injury to the upper airway causes airway swelling that typically leads to maximal edema in the first 24 to 48 hours after injury and often requires a short course of endotracheal intubation for airway protection. Combustion products found in smoke, most commonly from synthetics in structural fires, cause lower airway injury. These irritants cause direct mucosal injury, which in turn leads to mucosal sloughing, edema, reactive bronchoconstriction, and finally obstruction of the lower airways. Injury to both the epithelium and pulmonary alveolar macrophages causes release of prostanoids, chemokines, and other inflammatory mediators; neutrophil migration; increased tracheobronchial blood flow; and, finally, increased capillary permeability. All of these components of acute lung injury increase the risk of pneumonia and ARDS following an inhalation injury.

The physiologic effects of smoke inhalation are numerous. Inhalation injury decreases lung compliance⁹⁸ and increases airway resistance work of breathing.⁹⁷ Inhalation injury in the presence of burns also increases overall metabolic demands.⁹⁹ The most common physiologic derangement seen with inhalation injury is increased fluid requirement during resuscitation. Since severe inhalation injury may result in mucosal sloughing with obstruction of smaller airways, bronchoscopy findings including carbon deposits, erythema, edema, bronchorrhea, and a hemorrhagic appearance may be useful for staging inhalation injury. The Abbreviated Injury Score—a scale from 0 to 4, with 0 representing no injury and 4 representing massive injury—is commonly used for grading inhalation injury. Higher grades of bronchoscopic inhalation injury have been associated with increased incidence of ARDS, increased ventilator days, higher rate of multiple organ dysfunction syndrome, and higher mortality.⁹⁸ Bronchoscopic evaluation can also help isolate organisms early in the course of a potential pneumonia. Bronchoalveolar lavage (BAL) within 24 hours after an inhalation injury demonstrates a high rate of positive quantitative cultures,¹⁰⁰ suggesting that pneumonia develops soon after the acute lung injury. Bacterial contamination from urgent intubation may contribute to early development of pneumonia in patients with inhalation injury.¹⁰⁶ Early evaluation with bronchoscopy can identify causative organisms and guide appropriate antibiotic therapy.

Because bronchoscopy is an invasive test, attempts have been made to utilize other diagnostic modalities, such as thoracic computed tomography (CT) scans¹⁰¹ and xenon ventilation-perfusion scanning.¹⁰²,¹⁰³ However, these are generally not utilized unless otherwise indicated, and the best tools available for diagnosing inhalation injury remain clinical presentation and bronchoscopic evaluation. Decreased PaO₂/FiO₂ ratio (<350) on admission may not only predict inhalation injury but also indicate increased fluid needs more accurately than bronchoscopic grading of the severity of inhalation.¹⁰⁴

Treatment of inhalation injury consists primarily of supportive care. Aggressive pulmonary toilet and routine use of nebulized bronchodilators such as albuterol are recommended. Nebulized N-acetylcysteine is an antioxidant free radical scavenger designed to decrease the toxicity of high oxygen concentrations. Aerosolized heparin aims to prevent formation of fibrin plugs and decrease the formation of airway casts and has been associated with increased number of ventilator-free days.¹⁰⁵ A recent meta-analysis demonstrated improved mortality with the use of inhaled anticoagulation regimens.¹⁰⁶ Aerosolized tissue plasminogen activator¹⁰⁷ and recombinant human antithrombin¹⁰⁸ have shown promise in sheep models but have not yet seen widespread clinical use. Administration of intrabronchial surfactant has been used as a salvage therapy in patients with severe burns and inhalation injury.¹⁰⁹ Inhaled nitric oxide may also be useful as a last effort in burn patients with severe lung injury who are failing other means of ventilator support.¹¹⁰ The use of steroids has traditionally been avoided due to the worse outcomes in burn patients¹¹¹, however, some data demonstrate selectively improved outcomes with septic shock requiring vasopressor circulatory.¹¹²

An important contributor to early mortality in burn patients and often seen in patients with inhalation injury is carbon monoxide (CO) poisoning. This clear, odorless gas has an affinity for hemoglobin is approximately 200 to 250 times more than that of oxygen. Carboxyhemoglobin decreases the levels of normal oxygenated hemoglobin and can quickly lead to anoxia and death.¹¹³ CO also causes uncoupling of oxidative phosphorylation in mitochondria, free radical generation, and increased systemic inflammatory response via platelet activation—all of which may increase cardiac and neurologic morbidity and mortality in CO toxicity.¹¹⁴ Unexpected neurologic or cardiac symptoms should raise the level of suspicion, and an arterial carboxyhemoglobin level must be obtained because pulse oximetry can be falsely elevated. Administration of 100% normobaric oxygen is the gold standard for treating CO poisoning and reduces the half-life of CO from 250 minutes in room air to 40 to 60 minutes.¹¹⁵ Some authors have proposed hyperbaric oxygen as an adjunctive therapy for CO poisoning.¹¹⁶ However, a recent meta-analysis offers mixed results regarding the success and long-term outcomes of hyperbaric oxygen, and its associated logistical difficulties and complications have limited its usefulness for patients with moderate or large burns.¹¹⁷ Patients who sustain a cardiac arrest as a result of their CO poisoning have an extremely poor prognosis regardless of the success of initial resuscitation attempts.¹¹⁸

Hydrogen cyanide toxicity may also be a component of an overwhelming smoke inhalation injury. Cyanide inhibits cytochrome oxidase, which is required for oxidative phosphorylation.¹¹⁹ Afflicted patients may have a persistent, severe lactic acidosis, neurologic symptoms, pulmonary edema, or cardiac sequelae (ST elevation on electrocardiogram).¹²⁰,¹²¹ Classic signs of cyanide poisoning—including bitter almond breath and cherry-red skin changes—are rare and should not be used as the sole diagnostic criteria. Treatment consists of sodium thiosulfate, hydroxocobalamin, and 100% oxygen. Sodium thiosulfate works as a substrate for the metabolism cyanide into a nontoxic derivative, but it works slowly and is not effective for acute therapy.¹²¹ Hydroxocobalamina vitamin B₁₂ precursor—quickly
complexes with cyanide, is excreted by the kidney, and is recommended for immediate therapy.\(^{122}\) In the majority of patients, lactic acidosis will resolve with ventilation, and sodium thiosulfate treatment becomes unnecessary.\(^{123}\) Given the unknown side-effects of hydroxocobalamin administration, it should be reserved only for patients with a strong suspicion of cyanide poisoning.

New ventilator strategies have contributed to the improved mortality with ARDS. Although ARDS still contributes to mortality in burn patients, treatments have improved so that mortality is primarily from multisystem organ failure rather than isolated respiratory causes.\(^{124}\) The ARDS Network Study finding that low tidal volume (6 cc/kg) or “lung-protective ventilation” had a 22% lower mortality than patients with traditional tidal volumes (12 cc/kg)\(^{124}\) has dramatically changed the management of patients with acute lung injury. A similar approach had previously been shown to improve outcomes in pediatric burn patients.\(^{125}\) In patients with refractory hypoxemia despite lung-protective ventilation, prone positioning may improve oxygenation and mortality.\(^{126,127}\) No specific studies have examined prone positioning in burned patients, and in fact exclusion criteria from a large prone positioning trial included patients with ≥20% TBSA.\(^{127}\) Select reports demonstrate the feasibility of prone positioning in burn patients,\(^{128}\) although they present logistical challenges and caution must be used in patients with frontal and facial burns who are already at risk for loss of the grafts, invasive catheters, and the endotracheal tube. High-frequency percussive ventilation (HFPV) has shown early promise in patients with inhalation injury.\(^{129}\) One study showed notable decreases in both morbidity and mortality with HFPV, especially in patients with burns <40% TBSA and inhalation injury.\(^{130}\) A randomized controlled trial between low-tidal volume ventilation and HFPV in burn patients requiring mechanical ventilation demonstrated no significant difference in primary clinical outcomes.\(^{131}\) A related technique is high-frequency oscillatory ventilation (HFOV), which has been used primarily as a salvage modality in patients refractory to more conventional measures.\(^{132}\) However, two recent studies and a recent meta-analysis have concluded that HFOV yields no mortality benefit and in fact may actually increase patient mortality in patients with ARDS.\(^{133,134}\) Extracorporeal membrane oxygenation (ECMO) is typically reserved for salvage situations, although utilization of ECMO for burn patients is increasing and outcomes have been shown to be similar to other ECMO patients.\(^{135}\)

TREATMENT OF THE BURN WOUND

Multitudes of topical therapies exist for the treatment of burn wounds, many of which contain antimicrobial properties. A recent Cochrane Database Review nicely summarizes the data surrounding antisepsis for burns; however, much of the data is inconclusive.\(^{137}\) Silver sulfadiazine is one of the most widely used in clinical practice. Silver sulfadiazine has a wide range of antimicrobial activity, primarily as prophylaxis against burn wound infections rather than treatment of existing infections. It has the added benefits of being inexpensive, being easily applied, and having soothing qualities. It is not significantly absorbed systemically and thus has minimal metabolic derangements. Silver sulfadiazine has a reputation for causing neutropenia, but this association is more likely due to neutrophil margination from the inflammatory response following burn injury. True allergic reactions to the sulfa component of silver sulfadiazine are rare, and at-risk patients can have a small test patch applied to identify a burning sensation or rash. Silver sulfadiazine destroys skin grafts and is contraindicated on burns or donor sites in proximity to newly grafted areas. Also, silver sulfadiazine may retard epithelial migration in healing partial-thickness wounds.

Mafenide acetate, either in cream or solution form, is an effective topical antimicrobial. It is effective even in the presence of eschar and can be used in both treating and preventing wound infections; the solution formulation is an excellent antimicrobial for fresh skin grafts. Use of mafenide acetate may be limited by pain with application to partial-thickness burns. As mafenide is a carbonic anhydrase inhibitor, a historically described side effect is metabolic acidosis. However, multiple studies have been performed using mafenide to treat burn wounds without any significant incidence of metabolic acidosis.\(^{138,139}\)

Silver nitrate has broad-spectrum antimicrobial activity as a topical solution. The solution used must be dilute (0.5%), and prolonged topical application leads to electrolyte extravasation with resulting hyponatremia. A rare complication is methemoglobinemia.\(^{140}\) Although inexpensive, silver nitrate solution causes black stains, and laundry costs may offset any fiscal benefit to the hospital. Although there is no definitive evidence regarding use in the burn population, Dakin’s solution (0.5% sodium hypochlorite solution) is an acceptable alternative as an inexpensive topical antimicrobial.

For smaller burns or larger burns that are nearly healed, topical ointments such as bacitracin, neomycin, and polymyxin B can be used. These are also useful for superficial partial-thickness facial burns as they can be applied and left open to air without dressing coverage. Meshed skin grafts in which the interstices are nearly closed are another indication for use of these agents, preferably with greasy gauze to help retain the ointment in the affected area. All three have been reported to cause nephrotoxicity and should be used sparingly in large burns. Recent media coverage of mexitilin-resistant Staphylococcus aureus (MRSA) has led to widespread use by community practitioners of mupirocin for new burns. Unless the patient has known risk factors for MRSA, mupirocin should only be used in culture-positive burn wound infections to prevent emergence of further resistance.

Silver-impregnated dressings are increasingly being used for donor sites, skin grafts, and partial-thickness burns because of their potential to avoid daily dressing changes. These may be more comfortable for the patient, reduce the number of dressing changes, and shorten hospital length of stay, but they limit serial wound examinations. Biologic membranes such as Biobrane (Smith & Nephew Global Products) provide a prolonged barrier under which wounds may heal. Because of the occlusive nature of these dressings, these are typically used only on fresh, superficial, partial-thickness burns that are clearly not contaminated.

NUTRITION

Nutritional support may be more important in patients with large burns than in any other patient population. Not only does adequate nutrition play a role in acute issues such as immune responsiveness, but the hypermetabolic response in burn injury may raise baseline metabolic rates by as much as 200%.\(^{141}\) This can lead to catabolism of muscle proteins and decreased lean body mass that may delay functional recovery.\(^{142}\) Early enteral
feeding for patients with burns >20% TBSA is safe and may reduce loss of lean body mass. Early enteral feeds have also been associated with shorter duration of ICU stay and decreased rates of wound infection. If the enteral feeds are started within the first few hours after admission, gastric ileus may be avoided. Adjuncts such as metoclopramide promote gastrointestinal motility; if other measures for gastric feeding are unsuccessful, advancing the tube into the small bowel with nasojejunal feeding can be attempted. In endotracheally intubated patients, trips to the operating room do not necessitate holding enteral feedings. Immune-modulating supplements such as glutamine may decrease infectious complications in burn patients, although the effect on mortality and wound closure remains unknown. One proposed mechanism for glutamine’s immune modulating properties is via prevention of T-cell suppression in mesenteric lymph nodes. There is currently a multicenter randomized control trial recruiting to determine the effect of glutamine on mortality, blood stream infections, and health-related quality of life (https://clinicaltrials.gov/ct2/show/NCT00985205). Micronutrient supplementation with antioxidant vitamins (vitamin E and ascorbic acid) and trace minerals (selenium, zinc, and copper) optimizes wound healing, enhances immune function, and fights oxidative stress.

Calculating the appropriate caloric needs of the burn patient can be challenging. A commonly used formula in non-burned patients is the Harris-Benedict equation, which calculates caloric needs using factors such as gender, age, height, and weight. This formula uses an activity factor for specific injuries, and for burns, the basal energy expenditure is multiplied by two. The Harris-Benedict equation may be inaccurate in burns of <40% TBSA, and in these patients, the Curreri formula may be more appropriate. This formula estimates caloric needs to be 25 kcal/kg per d plus 40 kcal/%TBSA per d. Indirect calorimetry can also be used to calculate resting energy expenditure, but in burn patients, a “metabolic cart” has not been documented to be more beneficial than the predictive equations. Titrating caloric needs closely is important because overfeeding patients will lead to storage of fat instead of muscle anabolism.

Modifying the hypermetabolic response is an area of intense study. ß-Blocker use in pediatric patients decreases heart rate and resting energy expenditure and abrogates protein catabolism, even in long-term use. There may be benefits to ß-blockade in adult patients, and many centers use ß-blockers routinely in the adult population with limited safety and efficacy data. Some data suggests that ß-blocker use in the adult burn population has a greater incidence of iatrogenic hypotension and bradycardia. As such, it is important to monitor hemodynamic status when starting ß-blockers in these populations.

The anabolic steroid oxandrolone has been extensively studied in burn patients as well and has demonstrated improvements in lean body mass and bone density in severely burned children. The weight gain and functional improvements seen with oxandrolone may persist even after stopping administration of the drug. A double-blind, randomized study of oxandrolone showed decreased length of stay, improved hepatic protein synthesis, and no adverse effects on endocrine function, although the authors noted a rise in transaminases with unclear clinical significance. Oxandrolone therapy has also been associated with overall decreased mortality in patients with large burns. 

Hyperglycemia has been associated with increased mortality after burn injury, and intensive insulin therapy in critically ill patients has shown benefit, presumably from avoidance of hyperglycemia. However, in burn patients, the insulin itself may have a metabolic benefit, with improvements in lean body mass and amelioration of the inflammatory response to burn injury. Oral hypoglycemic agents such as metformin also help to avoid hyperglycemia and may contribute to prevention of muscle catabolism.

### Complications in Burn Care

There are several complications commonly associated with treatment of burn patients. Though not always avoidable, maintaining vigilance for typical complications and using appropriate techniques for prevention may limit the frequency and severity of complications. Ventilator-associated pneumonia, as in critically ill patients, is a significant problem in burned patients. However, it is so common in patients with inhalation injury that a better nomenclature may be postinjury pneumonia. Unfortunately, commonly used scores in critical illness such as the Clinical Pulmonary Infection Score (CPIS) have not been shown to be reliable in burn patients. Quantitative bronchoscopic cultures in the setting of clinical suspicion of pneumonia should guide treatment of pneumonia. Simple measures such as elevating the head of the bed and maintaining excellent oral hygiene and pulmonary toilet are recommended to help decrease the risk of postinjury pneumonia. There is some question as to whether early tracheostomy decreases infectious morbidity in burn patients and whether it improves long-term outcomes. There do not seem to be any major differences in the rates of pneumonia with early tracheostomy, though there may be reduced development of subglottic stenosis compared with prolonged endotracheal intubation. Practical considerations such as protection of facial skin grafts may influence the decision for tracheostomy placement. One major consideration in deciding whether to perform a tracheostomy has been the presence of eschar at the insertion site, which complicates tracheostomy site care and increases the risk of airway infection. Bedside percutaneous dilatational tracheostomy is a facile method for performing tracheostomy and is reported to be as safe as open tracheostomy in the burn population.

Massive resuscitation of burned patients may lead to an abdominal compartment syndrome characterized by increased airway pressures with hypoventilation and decreased urine output and hemodynamic compromise. Decompressive laparotomy is the standard of care for refractory abdominal compartment syndrome but carries an especially poor prognosis in burn patients. Adjunctive measures such as minimizing fluid, performing torso escharotomies, decreasing tidal volumes, and chemical paralysis should be initiated before resorting to decompressive laparotomy. Patients undergoing massive resuscitation also develop elevated intraocular pressures and may require lateral canthotomy.

Deep vein thrombosis (DVT) and prophylaxis in the burn population has received increasing attention in the literature recently. Up to 25% of burn patients develop DVT, and fatal pulmonary emboli have been reported in burn patients. A recent prospective trial demonstrated an 8% incidence of DVT in patients with 30% to 60% TBSA burns not receiving low molecular weight heparin prophylaxis with no evidence of DVT in patients receiving prophylaxis. There were no complications.
from low molecular weight heparin prophylaxis. Thus, it appears that heparin prophylaxis is safe in burn patients and may help prevent thrombotic complications.

Unfortunately, the use of both prophylactic and therapeutic heparin may be associated with heparin-associated thrombocytopenia (HIT). One study of HIT in burn patients showed an incidence of 1.6% in heparinized burn patients. Thrombotic complications included DVT, pulmonary embolus, and even arterial thrombosis requiring limb amputation. Nonheparin anticoagulation for HIT commonly caused bleeding complications requiring transfusion. Although rare, a high index of suspicion for HIT should be maintained in thrombocytopenic burn patients, particularly if the platelet counts drop at hospital days 7 to 10.

Burn patients often require central venous access for fluid resuscitation and hemodynamic monitoring. Because of the anatomic relation of their burns to commonly used access sites, burn patients may be at higher risk for catheter-related bloodstream infections. The 2012 Centers for Disease Control and Prevention National Healthcare Safety Network report indicates that American burn centers have higher infectious complication rates than any other ICUs. Because burn patients may commonly exhibit leukocytosis with a documented bloodstream infection, practice has been to rewire lines over a guide wire and to culture the catheter tip. However, this may increase the risk of catheter-related infections in burned patients, and a new site should be used if at all possible.

**SURGERY**

Full-thickness burns with a rigid eschar can form a tourniquet effect as the edema progresses, leading to compromised venous outflow and eventually arterial inflow. The resulting compartment syndrome is most common in circumferential extremity burns, but abdominal and thoracic compartment syndromes also occur. Warning signs of impending compartment syndrome may include paresthesias, pain, decreased capillary refill, and progression to loss of distal pulses; in an intubated patient, the surgeon should anticipate the compartment syndrome and perform frequent neurovascular evaluations. Abdominal compartment syndrome should be suspected with decreased urine output, increased ventilator airway pressures, and hypotension. Hypoventilation, increased airway pressures, and hypotension may also characterize thoracic compartment syndrome. Escharotomies are rarely needed within the first 8 hours following injury and should not be performed unless indicated because of the terrible aesthetic sequelae. When indicated, they are usually performed at the bedside, preferably with electrocautery to minimize blood loss. Extremity incisions are made on the lateral and medial aspects of the limbs in an anatomic position and may extend onto the hand and hypothenar eminences of the hand. Digital escharotomies do not usually result in any meaningful salvage of functional tissue and are not recommended. Inadequate perfusion despite proper escharotomies may indicate the need for fasciotomy, but this procedure should not be routinely performed as part of the eschar release. Thoracic escharotomies should be placed along the anterior axillary lines with bilateral subcostal and subclavicular extensions. Extension of the anterior axillary incisions down the lateral abdomen typically will allow adequate release of abdominal eschar.

The strategy of early excision and grafting in burned patients revolutionized survival outcomes in burn care. Not only did it improve mortality, but early excision also decreased reconstruction surgery, hospital length of stay, and costs of care. Once the initial resuscitation is complete and the patient is hemodynamically stable, attention should be turned to excising the burn wound. Burn excision and wound coverage should ideally start within the first several days, and in larger burns, serial excisions can be performed as patient condition allows. Excision is performed with repeated tangential slices using a Watson or Goulian blade until viable, diffusely bleeding tissue remains. It is appropriate to leave healthy dermis, which will appear white with punctate areas of bleeding. Excision to fat or fascia may be necessary in deeper burns. The downside of tangential excision is a high blood loss, though this may be ameliorated using techniques such as instillation of an epinephrine tumescence solution underneath the burn. Pneumatic tourniquets are helpful in extremity burns, and compresses soaked in a dilute epinephrine solution are necessary adjuncts after excision. A fibrinogen and thrombin spray sealant (Tisseel Fibrin Sealant; Baxter, Deerfield, IL) also has beneficial effects on both hemostasis and graft adherence to the wound bed. The use of these techniques has markedly decreased the number of blood transfusions given during burn surgery. For patients with clearly deep burns and concern for excessive blood loss, fascial excision may be employed. In this technique, electrocautery is used to excise the burned tissue and the underlying subcutaneous tissue down to muscle fascia. This technique markedly decreases blood loss but results in a cosmetically inferior appearance due to the loss of subcutaneous tissue. For excision of burns in difficult anatomic areas, such as the face, eyelids, or hands, a pressurized water dissector may offer more precision but is time consuming, has a steep learning curve, and is expensive.

**WOUND COVERAGE**

Since full-thickness burns are impractical for most burn wounds, split-thickness sheet autografts harvested with a power dermatome make the most durable wound coverings and have a decent cosmetic appearance. In larger burns, meshed autografted skin provides a larger area of wound coverage. This also allows drainage of blood and serous fluid to prevent accumulation under the skin graft with subsequent graft loss. Areas of cosmetic importance such as the face, neck, and hands should be grafted with nonmeshed sheet grafts to ensure optimal appearance and function. Unfortunately, even extensive meshing of skin grafts in patients with limited donor sites may not provide adequate amounts of skin. One emerging technique for large burns with limited donor sites is the Meek micrografting technique, or “postage-stamp” technique, where expansion ratios of up to 9:1 are able to be achieved. This technique has a considerable learning curve and requires further investigation to determine whether it is the optimal technique for large surface area burns with limited donor sites. Options for temporary wound coverage include human cadaveric allograft, which is incorporated into the wound but is rejected by the immune system and must be eventually replaced. This allows temporary biologic wound coverage until donor sites heal enough so that they may be reharvested. Xenograft appears to function as well as allograft for temporary wound coverage and is considerably less expensive.

The search for a perfect permanent synthetic skin substitute remains elusive. Integra (Integra LifeSciences Corporation, Plainsboro, NJ) is a bilayer product with a porous
collagen-chondroitin 6-sulphate inner layer that is attached to an outer silastic sheet, which helps prevent fluid loss and infection as the inner layer becomes vascularized, creating an artificial neodermis. At approximately 2 weeks after placement, the silastic layer can be removed and a thin autograft can be placed over the neodermis. This results in faster healing of the more superficial donor sites and has been increasingly utilized for treatment of complex wounds and injuries. Alloderm (LifeCell Corporation, The Woodlands, TX) is a dermal substitute consisting of cryopreserved acellular human dermis. NovoSorb™ Biodegradable Temporizing Matrix (PolyNovo Limited, Melbourne, Australia) is a biodegradable polyurethane dermal substitute newly available and recently approved by the United States Food and Drug Administration (FDA). These dermal substitutes should also be used in combination with thin split-thickness skin grafts for final wound coverage.

Epidermal skin substitutes such as cultured epithelial autografts are an option in patients with massive burns and very limited donor sites. Their clinical use has been limited by a long turnaround time for culturing, as well as the fragility of the cultured skin, which creates great difficulty with intraoperative handling and graft take. There are promising developments in skin culturing techniques and engineered skin development, but no other products are FDA approved and commercially available.

Thighs make convenient anatomic donor sites; they are easily harvested and relatively hidden from an aesthetic standpoint. The thicker skin of the back is useful in older patients, who have thinner skin elsewhere and may have difficulty with healing of donor sites. The buttocks are an excellent donor site in infants and toddlers; silver sulfadiazine can be applied to the donor site with a diaper as coverage. The scalp is also an excellent donor site; the skin is thick and the many hair follicles allow rapid healing, with the added advantage of being completely hidden once hair regrows. Epinephrine tumesence is necessary for harvesting the scalp, for both hemostasis of this hypervascular area and also to create a smooth contoured surface for harvesting.

The list of commonly used donor site dressings is long and includes simple transparent films to hydrocolloids, petrolatum gauzes, and silver-impregnated dressings. Donor sites close to fresh grafts may be dressed with a porous nonadherent gauze, and both the donors and grafts are soaked with an antimicrobial solution. Principles behind choosing a dressing should balance ease of care, comfort, infection control, and cost. The choice of donor site dressing is largely institution dependent, and few data support the clear superiority of any single treatment plan.

**REHABILITATION**

Rehabilitation is an integral part of the clinical care plan for the burn patient and should be initiated on admission. Immediate and ongoing physical and occupational therapy is mandatory to prevent functional loss. Patients who are unable to actively participate should have passive range-of-motion exercises done at least twice a day. This includes patients with burns over joints, such as with hand burns. Patients should be taught exercises they can do themselves to maintain full range of motion. Patients with foot and extremity burns should be instructed to walk independently without crutches or other assistive devices to prevent extremity swelling, desensitize the burned areas, and prevent disuse atrophy; when patients are not ambulating, they must elevate the affected extremity to minimize swelling. If postoperative immobilization is used for graft protection, the graft should be evaluated early and at frequent intervals so that active exercise can be resumed at the earliest possible occasion. The transition to outpatient care should also include physical and occupational therapy, with introduction of exercises designed to accelerate return to activities of daily living as well as specific job-related tasks. Tight-fitting pressure garments provide vascular support in burns that are further along in the healing process. Whether they prevent hypertrophic scar formation has been long debated. However, they do provide vascular support that many patients find more comfortable.

**LATE COMPLICATIONS: HYPERTROPHIC SCAR, CONTRACTURES, AND HETEROOTOPIC OSSIFICATION**

Once patients have recovered from their acute burns, many face management of the hypertrophic burn scars. In patients with healed burns or donor sites, hypertrophic scar-related morbidity includes pruritus, erythema, pain, thickened tight skin, and even contractures. Within these scars, there is believed to be an increased inflammatory response, irregular neovascularization, aberrant cytokine and Toll-like receptor expression, abundant collagen production, and abnormal extracellular matrix structure. Treatment for these scars has included nonsurgical therapies such as compression garments, silicone gel sheeting, massage, physical therapy, and corticosteroid. Surgical excision and scar revision represent more invasive scar management approaches that are often necessary for functional and aesthetic recovery.

Laser-based therapies provide additional treatment options for symptomatic hypertrophic scars. Two of the most common ones are the pulsed dye laser (PDL) and the ablative carbon dioxide (CO2) laser. The PDL causes photothermolysis of hemoglobin, resulting in coagulative necrosis. It obliterates small capillaries close to the skin and has had success treating congenital, cutaneous vascular malformations. The CO2 laser has been used for treatment of acne and recently has been gaining acceptance for its use to treat hypertrophic burn scars. It works by ablating microscopic columns of tissue to flatten scars and is also believed to stimulate matrix metalloproteinases and other signaling pathways to induce collagen reorganization. Lasers are ultimately believed to help with scar remodeling and collagen reorganization. CO2 laser therapy has been shown to decrease symptoms associated with hypertrophic scarring, including scar appearance, pliability, contracture, neuropathic pain, and pruritus. A recent prospective study utilizing PDL and CO2 laser therapy demonstrated improved signs and symptoms of hypertrophic scars based on the Vancouver Scar Scale and the University of North Carolina 4P Scar Scale. Outpatient and office-based treatment sessions are tolerated well by most patients. There is wide practice variation on when to start therapy and the number of treatments, but the literature has general support for starting treatment at 6 to 12 months and offering three treatments. More research is needed to determine the full potential of laser therapy to provide burn survivors a less invasive treatment of hypertrophic scars with improved symptoms and quality of life.

Contractures are another long-term complication of burn injury that can result in significant morbidity. Contractures result from both wound contracture and scar contracture and prevents range of motion of a particular joint. Factors influencing contracture development include burn depth and activation...
of dermal fibroblasts, myofibroblasts, fibrocytes, and helper T cells.\textsuperscript{196} Despite aggressive physiotherapy, contractures have been reported to develop in as many as one-third of burn patients. A recent study of 1,865 patients demonstrated that the shoulder is the most affected joint, followed by the elbow, wrist, ankle, and knee.\textsuperscript{197} A similar study in the pediatric population yielded similar results. Gender, race, and %TBSA were associated with contracture development in the adult population. Age and length of stay in the ICU were associated with contracture development, severity of contracture, and total number of contractures in the pediatric population. Treatment of contractures includes both nonsurgical and surgical options, ranging from pressure garments and splints to laser therapy and contracture excision.

Heterotopic ossification (HO) is another long-term morbidity associated with burn injury. HO is the pathologic development of lamellar bone in peripheral tissue. Its incidence has been reported to be between 1% and 3% of burn patients.\textsuperscript{198} Symptoms include decreased range of motion, pain, and swelling overlying the affected joints. Often times, the pathologic bone formation can be visualized radiographically with plain X-rays. Risk factors include >30% TBSA, arm burns, arm grafts, ventilator days, and number of trips to the operating room.\textsuperscript{199} Treatment includes aggressive physiotherapy, NSAIDs, bisphosphonates, radiation therapy, and rarely surgical excision. A risk scoring system has been developed to predict which burn patients are at risk of developing HO based on admission criteria; however, further validation is warranted.\textsuperscript{200}

**PSYCHOLOGICAL RECOVERY**

Psychological rehabilitation is equally important in the burn patient. Depression, posttraumatic stress disorder (PTSD), concerns about image, and anxiety about returning to society constitute predictable barriers to progress in both the inpatient and outpatient setting. Psychological distress occurs in as many as 38% of burn patients and persists in severity long after discharge.\textsuperscript{201} Rates of depression vary between 4% and 54% following injury, although these numbers vary dramatically based on the methodology used to diagnose depression.\textsuperscript{202} Still, depressive symptoms have been documented in up to 43% of patients 2 years following injury and have been associated with the female gender. Factors such as gender, extraversion, capacity for forgiveness, the event as a disaster or nondisaster, alcohol use, and peritraumatic emotional response have been identified as contributing factors to PTSD.\textsuperscript{203} Despite the psychological impact of burn injury, many patients will be able to quickly return to work or school, and goals should be set accordingly. The involvement of clinical psychologists and psychiatrists is invaluable in providing guidance and coping techniques to lessen the significant psychological burden of burn injury.

**PREVENTION**

Despite many areas of progress in prevention over the past century, burns continue to be a common source of injury. The cornerstone for burn prevention programs has been “The Five Step Process,” a systematic method of assessing, implementing, and evaluating burn hazards and subsequent intervention impact, and The Five E’s—engineering/environment, enforcement, education, emergency response, and economic initiative.\textsuperscript{204} It has been shown that patients who live in environments optimal for sustaining burn injury have decreased knowledge of burn prevention strategies.\textsuperscript{205} Some successful initiatives have included school-based education and community-based interventions targeting simple home safety measures. A 6-year study of second-graders demonstrated both short- and long-term retention of information related to burn, fire, and life safety following multiple educational sessions. Smoke alarms are known to decrease mortality from structural fires, but not all homes are equipped with proper smoke alarms, particularly in low-income households. Mandatory smoke alarm installation via community initiatives can be successful but seems to be contingent on close, long-term follow-up to ensure proper maintenance and function.\textsuperscript{206,207} Regulation of hot water heater temperatures has had some success and may be even more effective in conjunction with community-based programs emphasizing education and in-home inspections.\textsuperscript{208,209} A recent systematic review of prevention in low- and middle-income countries identified multiple successful prevention programs.\textsuperscript{210} Burn professionals have also demonstrated incomplete knowledge on best practices for fire safety and burn prevention.\textsuperscript{211} As such, appropriate education of burn professionals participating in prevention programs is necessary.

**BURN DISASTERS**

Although rare, burn disasters can be devastating to those involved due to the sudden nature of the event, the difficulty of managing personnel and resources,\textsuperscript{212} a deficit of staff experience in burn management, and relatively small resource availability for a potentially large number of patients.\textsuperscript{213} The American Burn Association has estimated that up to 30% of patients in mass casualty incidents suffer from burn injury.\textsuperscript{214} A recent review of the literature between 1990 and 2016 identified 752 burn disasters world-wide, defined as an incident with ≥50 burn injuries and/or ≥30 burn-related deaths. The majority occurred in Asia and the Middle East and are thought to be secondary to rapid industrialization, inadequate fire-prevention strategies, and poor building codes. There was a significant increase in terrorist-related incidences from 2000 to 2015. Finally, the authors demonstrated that international adoption of the U.S. Health and Human Services guidelines on bed availability for burns and trauma dramatically underestimated the number of beds needed for burn disasters.\textsuperscript{215}

Preparedness is paramount for reacting quickly, efficiently, and effectively to a burn disaster. General surgeons not trained in burn care may feel uncomfortable longitudinally caring for severely burned patients following a burn disaster. However, due to resource limitations, they should be prepared to care for burn patients for the first 72 hours of resuscitation or until the patients can be transferred to a center that specializes in burn care. This will involve initial evaluation, resuscitation, and potential interventions including central line placement, intubation, and escharotomies. Coordination for burn disasters should take place at three levels: institutional/intrafacility, interfacility/intrastate, and interstate/regional. It is important to have multiple stakeholders involved in the development of a disaster plan, from the burn surgeons to the emergency department personnel to the emergency medical personnel who are first responders.\textsuperscript{216} Resource utilization can be guided by The American Burn Association Age/TBSA survival grid, which stratifies patients into benefit-to-resource categories (outpatient, high, low, and expectant) based on age and %TBSA.\textsuperscript{3} This allows providers to allocate resources during burn disasters based on the severity of injury.
and expected survival. Another important consideration is the involvement of nonburn hospitals in the planning for burn disasters as burn centers do not possess enough resources to be sole providers in these events. Multiple strategies have been adopted by local burn centers, including the development of a consortium of hospitals surrounding one burn center in New Jersey to allow transfer of patients when resource capacity is in jeopardy.

Interest in mass burn casualty disaster planning invariably includes a discussion of radiation burns. Radioactive material results in both acute injury from immediate exposure and more prolonged injury from delayed exposure to radioactive fallout or contamination. When a 10-kiloton nuclear bomb is detonated, people at a distance 0.7 miles from ground zero absorb 4.5 Gy. At 60 days, the median lethal radiation dose (LD50) is 3.5 Sv; with aggressive medical care, this dose might be doubled to nearly 7 Sv. To put this in context, radiation exposure from a diagnostic CT of the chest or abdomen is 5 mSv, and the average annual background absorbed radiation dose is 3.6 mSv. Radiation is known to impact several organ systems and result in several syndromes based on increasing exposure doses. These syndromes include hematologic (1–8 Sv exposure), gastrointestinal (8–30 Sv exposure), and cardiovascular/neurologic syndromes (>30 Sv exposure), with the latter two being nonsurvivable.

After initial evaluation and decontamination by removing clothing, a useful way to estimate exposure is by determining the time to emesis. Patients who do not experience emesis within 4 hours of exposure are unlikely to have severe clinical effects. Emesis within 2 hours suggests a dose of at least 3 Sv, and emesis within 1 hour suggests at least 4 Sv. The hematologic system follows a similar dose-dependent temporal pattern for predicting radiation exposure, mortality, and treatment. These have been determined based on the Armed Forces Radiobiology Research Institute’s Biodosimetry Assessment Tool, which can be downloaded from www.afri.usuhs.mil.

The combination of radiation exposure and burn wounds has the potential to increase mortality compared with traditional burns. Early closure of wounds before radiation depletes circulating lymphocytes may be needed for wound healing (which occurs within 48 hours). Also, in radiation injuries combined with burn or trauma, laboratory lymphocyte counts may be unreliable. A significant difference between burn/traumatic injuries and radiation injuries is that burn/traumatic injuries can result in higher mortality when not treated within hours.

Decontamination and triage are vital to maximize the number of survivors. Initial decontamination requires removal of clothing and washing wounds with water. Irrigation fluid should be collected to prevent radiation spread into the water supply. Work by many professional organizations, including the ABA, has focused on nationwide triage for disasters and will be vital to save as many lives as possible. Yet, it is likely that expectant or comfort care could be offered to more patients than typically seen in civilian hospitals, due to resource availability after the disaster.

Finally, agents used in warfare—including white phosphorus and sulfur mustard—can cause significant morbidity and mortality. White phosphorus oxidizes when exposed to the atmosphere, creating the highly corrosive phosphorus pentoxide. Absorption of even small amounts of white phosphorus can result in hypocalcemia and hyperphosphatemia and their subsequent cardiac side effects. Treatment consists of removal of all clothing, irrigation with cool liquid (as phosphorus pentoxide liquidizes above 44°C), application of saline soaked gauze to prevent drying out and reignition, and potential surgical excision. Sulfur mustard, more commonly known as mustard gas, is another chemical warfare agent that can cause lesions similar to burn lesions. The gas infiltrates the skin surface, causing degranulation of mast cells, leukocyte invasion, and subsequent blistering of the skin. Treatment includes scrubbing to relieve the remaining skin of sulfur mustard, irrigation, and traditional burn therapy depending on the depth of the lesion.

**FUTURE AREAS OF STUDY**

It has long been anecdotally noted that two patients of similar ages and burn size may have very divergent responses to their burn injuries. Attention is being increasingly turned to identifying genetic differences among burn patients and how they affect response to injury. Specific allele variants have been linked with increased mortality in burned patients. It may be that genetic differences may predispose burn patients to severe sepsis, perhaps by downregulating the immune response. The Inflammation and the Host Response to Injury trial was a prospective, multicenter, federally funded study that aimed to define specific genetic pathways that differ in the response to both burns and traumatic injury. Blood and tissue samples from a strictly defined patient population were analyzed using gene arrays to determine whether differential expression in certain genetic pathways affects clinical outcomes. Although data from this study are still being analyzed, some interesting findings suggest that sepsis, trauma, and burn patients share common gene expression patterns, starting early after injury. These genes can upregulate proinflammatory pathways as well as disrupt antigen presentation pathways. A better understanding of these common genomic responses may allow for the targeted treatment of immunologic and signal pathways to help improve patient survival from burn injuries.

Another area of increasing interest includes integration of technology to burn size estimation and resuscitation. These include the use of smart device applications to assist with estimation of burn size and resuscitation recommendations. Further investigation is needed to determine the applicability of these models to burn estimation and resuscitation. However, as these models can include hourly updates and recommendations, they nudge clinicians to frequently reconsider fluid parameters during the critical stages of resuscitation.

With the dramatic progress in improving survival following a major burn injury during the twentieth century, understanding and addressing functional and psychological outcomes is critical to the well-being of burn survivors. Since 1993, the National Institute of Disability and Rehabilitation Research has funded four burn model systems to identify long-term sequelae of burn injuries and to develop ways to improve outcomes for survivors. Ongoing outcome studies are crucial for dismantling barriers that our patients face in returning to their communities and to the workplace or to school.

**REFERENCES**

Entries highlighted in bright blue are key references.


CHAPTER 8


169. Gore DC, Wolf SE, Sanford A, Herndon DN, Wolfe RR. Influence of metformin on glucose intolerance and...


This page intentionally left blank
HISTORY OF WOUND HEALING

The earliest accounts of wound healing date back to about 2000 B.C., when the Sumerians employed two modes of treatment: a spiritual method consisting of incantations, and a physical method of applying poultice-like materials to the wound. The Egyptians were the first to differentiate between infected and diseased wounds compared to noninfected wounds. The Edwin Smith Surgical Papyrus, a copy of a much older document, describes at least 48 different types of wounds. A later document (Ebers Papyrus, 1550 B.C.) relates the use of concoctions containing honey (antibacterial properties), lint (absorbent properties), and grease (barrier) for treating wounds. These same properties are still considered essential in contemporary daily wound management.

The Greeks, equipped with the knowledge bequeathed by the Egyptians, went even further and classified wounds as acute or chronic in nature. Galen of Pergamum (120–201 A.D.), appointed as the doctor to the Roman gladiators, had an enormous number of wounds to deal with following gladiatorial combats. He emphasized the importance of maintaining a moist environment to ensure adequate healing. It took almost 19 centuries for this important concept to be proven scientifically, when it was shown that the epithelialization rate increases by 50% in a moist wound environment when compared to a dry wound environment.1

The next major stride in the history of wound healing was the discovery of antiseptics and their importance in reducing wound infections. Ignaz Philipp Semmelweis, a Hungarian obstetrician (1818–1865), noted that the incidence of puerperal fever was much lower if medical students, following cadaver-dissection class and prior to attending childbirth, washed their hands with soap and hypochlorite. Louis Pasteur (1822–1895) was instrumental in dispelling the theory of spontaneous generation of germs and proving that germs existed in and were always introduced from the environment. Joseph Lister probably made one of the most significant contributions to wound healing. On a visit to Glasgow, Scotland, Lister noted that some areas of the city’s sewer system were less murky than the rest. He discovered that the water from pipes that were dumping waste containing carbolic acid (phenol) was clear. In 1865, Lister began soaking his surgical instruments in phenol and spraying the operating rooms, reducing the postoperative mortality rates from 50% to 15%. After attending an impressive lecture by Lister in 1876, Robert Wood Johnson left the meeting and began 10 years of research that would ultimately result in the production of an antiseptic dressing in the form of cotton gauze impregnated with iodoform. Since then, several other materials have been used to impregnate cotton gauze to achieve antisepsis.

The 1960s and 1970s led to the development of polymeric dressings. These polymeric dressings can be custom made to specific parameters, such as permeability to gases (occlusive vs. semi-occlusive), varying degrees of absorbency, and different physical forms. Due to the ability to customize, the available range of materials that aid in wound care has grown exponentially to include an ever-expanding variety. Currently, the practice of wound healing encompasses manipulation and/or use of, among others, inflammatory cytokines, growth factors, and bioengineered tissue. It is the combination of all these modalities that enables optimal wound healing. The role of organism in the perpetuation of nonhealing of chronic wounds
Wound healing is a complex cellular and biochemical cascade that leads to restitution of integrity and function. Although wounds are classified under one entity, it is believed that they behave differently based on the host and organism involved. The future of wound healing is in “precision medicine” in which treatment strategies will be based on the host, the underlying mechanism, and the organisms in the wound bed and tissue.

PHASES OF WOUND HEALING
Wound healing is a complex process of overlapping phases that is initiated by an injury or wound. Normal wound healing is divided into phases defined by characteristic cellular populations and biochemical activities: (a) hemostasis and inflammation, (b) proliferation, and (c) maturation and remodeling. An approximate timeline of these events is depicted in Fig. 9-1. This sequence of events in most circumstances spans the time from injury to resolution of acute wounds. All wounds need to progress through this series of cellular and biochemical events that characterize the phases of healing in order to successfully reestablish tissue integrity. However, multiple factors can interfere with this sequence and can lead to lengthy healing (chronic wounds) or nonhealing.

Hemostasis and Inflammation
Hemostasis precedes and initiates inflammation with the ensuing release of chemotactic factors from the wound site (Fig. 9-2A). Wounding by definition disrupts tissue integrity, leading to division of blood vessels and direct exposure of extracellular matrix to platelets. Exposure of subendothelial collagen to platelets results in platelet aggregation, degranulation, and activation of the coagulation cascade. Platelet α granules release a number of wound-active substances, such as platelet-derived growth factor (PDGF), transforming growth factor-β (TGF-β), platelet-activating factor (PAF), fibronectin, and serotonin. In addition to achieving hemostasis, the fibrin clot serves as scaffolding for the migration into the wound of inflammatory cells such as polymorphonuclear leukocytes (PMNs, neutrophils) and monocytes.

Cellular infiltration after injury follows a characteristic, predetermined sequence (see Fig. 9-1). PMNs are the first infiltrating cells to enter the wound site, peaking at 24 to 48 hours. Increased vascular permeability, local prostaglandin release, and the presence of chemotactic substances such as complement factors, interleukin-1 (IL-1), tumor necrosis factor-α (TNF-α), TGF-β, platelet factor 4, or bacterial products all stimulate neutrophil migration.

The postulated primary role of neutrophils is phagocytosis of bacteria and tissue debris. PMNs are also a major source of cytokines early during inflammation, especially TNF-α which may have a significant influence on subsequent angiogenesis and collagen synthesis (see Fig. 9-2B). PMNs also release proteases such as collagenases, which participate in matrix and ground substance degradation in the early phase of wound healing. Other than their role in limiting infections, these cells do not appear to play a role in collagen deposition or acquisition of mechanical wound strength. On the contrary, neutrophil factors have been implicated in delaying the epithelial closure of wounds.

The second population of inflammatory cells that invades the wound consists of macrophages, which are recognized as being essential to successful healing. Derived from circulating monocytes, macrophages achieve significant numbers in the wound by 48 to 96 hours post injury and remain present until wound healing is complete.

Macrophages, like neutrophils, participate in wound debridement via phagocytosis and contribute to microbial stasis via oxygen radical and nitric oxide synthesis (see Fig. 9-2B,C). The macrophage’s central function is activation and recruitment of other cells via mediators such as cytokines and growth factors, as well as directly by cell-cell interaction and intercellular adhesion molecules (ICAM). By releasing such mediators as TGF-β, vascular endothelial growth factor (VEGF), insulin-like growth factor (IGF), epithelial growth factor (EGF), and lactate, macrophages regulate cell proliferation, matrix synthesis, and angiogenesis. Macrophages also play a significant role in regulating angiogenesis and matrix deposition and remodeling (Table 9-1).

T lymphocytes comprise another population of inflammatory/immune cells that routinely invades the wound. Less
numerous than macrophages, T-lymphocyte numbers peak at about 1 week post injury and truly bridge the transition from the inflammatory to the proliferative phase of healing. Though known to be essential to wound healing, the role of lymphocytes in wound healing is not fully defined. A significant body of data supports the hypothesis that T lymphocytes play an active role in the modulation of the wound environment. Depletion of most wound T lymphocytes decreases wound strength and collagen content, while selective depletion of the CD8+ suppressor subset of T lymphocytes enhances wound healing. However, depletion of the CD4+ helper subset has no effect. Lymphocytes also exert a downregulating effect on fibroblast collagen synthesis by cell-associated interferon IFN-γ, TNF-α, and IL-1. This effect is lost if the cells are physically separated, suggesting that extracellular matrix synthesis is regulated not only via soluble factors but also by direct cell-cell contact between lymphocytes and fibroblasts.

**Proliferation**

The proliferative phase is the second phase of wound healing and roughly spans days 4 through 12 (see Fig. 9-2C). It is during this phase that tissue continuity is reestablished. Fibroblasts and endothelial cells are the last cell populations to infiltrate the healing wound, and the strongest chemotactic factor for fibroblasts is PDGF. Upon entering the wound environment, recruited fibroblasts first need to proliferate, and then become activated, to carry out their primary function of matrix synthesis remodeling. This activation is mediated mainly by the cytokines and growth factors released from wound macrophages.

Fibroblasts isolated from wounds synthesize more collagen than nonwound fibroblasts, they proliferate less, and they actively carry out matrix contraction. Although it is clear that the cytokine-rich wound environment plays a significant role in this phenotypic alteration and activation, the exact mediators are only partially characterized. Additionally, lactate, which accumulates in significant amounts in the wound environment over time (~10 mmol), is a potent regulator of collagen synthesis through a mechanism involving adenosine diphosphate (ADP)-ribosylation.

Endothelial cells also proliferate extensively during this phase of healing. These cells participate in the formation of new capillaries (angiogenesis), a process essential to successful wound healing. Endothelial cells migrate from intact venules close to the wound. Their migration, replication, and new capillary tubule formation is under the influence of such cytokines and growth factors as TNF-α, TGF-β, and VEGF. Although
many cells produce VEGF, macrophages represent a major source in the healing wound, and VEGF receptors are located specifically on endothelial cells. 18,19

**Matrix Synthesis**

**Biochemistry of Collagen.** Collagen, the most abundant protein in the body, plays a critical role in the successful completion of adult wound healing. Its deposition, maturation, and subsequent remodeling are essential to the functional integrity of the wound.

Although there are at least 18 types of collagen described, the main ones of interest to wound repair are types I and III. Type I collagen is the major component of extracellular matrix in skin. Type III, which is also normally present in skin, becomes more prominent and important during the repair process.

Biochemically, each chain of collagen is composed of a glycine residue in every third position. The second position in the triplet is made up of proline or lysine during the translation process. The polypeptide chain that is translated from mRNA contains approximately 1000 amino acid residues and is called protocollagen. Release of protocollagen into the endoplasmic reticulum results in the hydroxylation of proline and lysine by specific hydroxylases (Fig. 9-3). Prolyl hydroxylase requires oxygen and iron as cofactors, α-ketoglutarate as co-substrate, and ascorbic acid (vitamin C) as an electron donor. In the endoplasmic reticulum, the protocollagen chain is also glycosylated by the linking of galactose and glucose at specific hydroxyl-sine residues. These steps of hydroxylation and glycosylation alter the hydrogen bonding forces within the chain, imposing steric changes that force the protocollagen chain to assume an α-helical configuration. Three α-helical chains entwine to form a right-handed superhelical structure called procollagen. At both ends, this structure contains nonhelical peptide domains called registration peptides. Although initially joined by weak, ionic bonds, the procollagen molecule becomes much stronger by the covalent cross-linking of lysine residues.

Extracellularly, the nonhelical registration peptides are cleaved by a procollagen peptidase, and the procollagen strands undergo further polymerization and cross-linking. The resulting collagen monomer is further polymerized and cross-linked by the formation of intra- and intermolecular covalent bonds.

Collagen synthesis, as well as posttranslational modifications, are highly dependent on systemic factors such as an adequate oxygen supply; the presence of sufficient nutrients (amino acids and carbohydrates) and cofactors (vitamins and trace metals); and the local wound environment (vascular supply and lack of infection). Addressing these factors and reversing nutritional deficiencies can optimize collagen synthesis and deposition.

**Proteoglycan Synthesis.** Glycosaminoglycans comprise a large portion of the “ground substance” that makes up granulation tissue. Rarely found free, they couple with proteins to form proteoglycans. The polysaccharide chain is made up of repeating disaccharide units composed of glucuronic or iduronic acid and a hexosamine, which is usually sulfated. The disaccharide
The composition of proteoglycans varies from about 10 units in the case of heparin sulfate to as much as 2000 units in the case of hyaluronic acid.

The major glycosaminoglycans present in wounds are dermatan and chondroitin sulfate. Fibroblasts synthesize these compounds, increasing their concentration greatly during the first 3 weeks of healing. The interaction between collagen and proteoglycans is being actively studied. It is thought that the assembly of collagen subunits into fibrils and fibers is dependent upon the lattice provided by the sulfated proteoglycans.

Furthermore, it appears that the extent of sulfation is critical in determining the configuration of the collagen fibrils. As scar collagen is deposited, the proteoglycans are incorporated into the collagen scaffolding. However, with scar maturation and collagen remodeling, the content of proteoglycans gradually diminishes.

**Maturation and Remodeling**

The maturation and remodeling of the scar begins during the fibroplastic phase and is characterized by a reorganization of previously synthesized collagen. Collagen is broken down by matrix metalloproteinases (MMPs), and the net wound collagen content is the result of a balance between collagenolysis and collagen synthesis. There is a net shift toward collagen synthesis and eventually the reestablishment of extracellular matrix composed of a relatively acellular collagen-rich scar.

Wound strength and mechanical integrity in the fresh wound are determined by both the quantity and quality of the newly deposited collagen. The deposition of matrix at the wound site follows a characteristic pattern: fibronectin and collagen type III constitute the early matrix scaffolding; glycosaminoglycans and proteoglycans represent the next significant matrix components; and collagen type I is the final matrix. By several weeks post injury, the amount of collagen in the wound reaches a plateau, but the tensile strength continues to increase for several more months. Fibril formation and fibril cross-linking result in decreased collagen solubility, increased strength, and increased resistance to enzymatic degradation of the collagen matrix. Fibrillin, a glycoprotein secreted by fibroblasts, is essential for the formation of elastic fibers found in connective tissue. Scar remodeling continues for many (6 to 12) months post injury, gradually resulting in a mature, avascular, and acellular scar. The mechanical strength of the scar never achieves that of the uninjured tissue.

There is a constant turnover of collagen in the extracellular matrix, both in the healing wound as well as during normal tissue homeostasis. Collagenolysis is the result of collagenase activity, a class of MMPs that require activation. Both collagen synthesis and lysis are strictly controlled by cytokines and growth factors. Some factors affect both aspects of collagen remodeling. For example, TGF-β increases new collagen transcription and also decreases collagen breakdown by stimulating synthesis of tissue inhibitors of metalloproteinase. This balance of collagen deposition and degradation is the ultimate determinant of wound strength and integrity.

**Epithelialization**

Epithelialization is the final step in establishing tissue integrity. This process is characterized primarily by proliferation and migration of epithelial cells adjacent to the wound (Fig. 9-4). The process begins within 1 day of injury and is seen as thickening of the epidermis at the wound edge. Marginal basal cells at the edge of the wound lose their firm attachment to the underlying dermis, enlarge, and begin to migrate across the surface of the provisional matrix. Fixed basal cells in a zone near the cut edge undergo a series of rapid mitotic divisions, and these cells appear to migrate by moving over one another in a leapfrog fashion until the defect is covered. Once the defect is bridged, the migrating epithelial cells lose their flattened appearance, become more columnar in shape, and increase their mitotic activity. Layering of the epithelium is reestablished, and the surface layer eventually keratinizes.

Figure 9-3. The steps of collagen synthesis. mRNA = messenger RNA.
Reepithelialization is complete in less than 48 hours in the case of approximated incised wounds but may take substantially longer in the case of larger wounds, where there is a significant epidermal/dermal defect. If only the epithelium and superficial dermis are damaged, such as occurs in split-thickness skin graft donor sites or in superficial second-degree burns, then repair consists primarily of reepithelialization with minimal or no fibroplasia and granulation tissue formation. The stimuli for reepithelialization remain incompletely defined; however, it appears that the process is mediated by a combination of a loss of contact inhibition; exposure to constituents of the extracellular matrix, particularly fibronectin; and cytokines produced by immune mononuclear cells. In particular EGF, TGF-β, basic fibroblast growth factor (bFGF), PDGF, and IGF-1 have been shown to promote epithelialization.

**Role of Growth Factors in Normal Healing**

Growth factors and cytokines are polypeptides produced in normal and wounded tissue that stimulate cellular migration, proliferation, and function. They often are named for the cells from which they were first derived (e.g., platelet-derived growth factor, PDGF) or for their initially identified function (e.g., fibroblast growth factor, FGF). These names are often misleading because growth factors have been demonstrated to have multiple functions. Most growth factors are extremely potent and produce significant effects in nanomolar concentrations.

They may act in an autocrine manner (where the growth factor acts on the cell producing it), a paracrine manner (by release into the extracellular environment, where it acts on the immediately neighboring cells), or in an endocrine manner (where the effect of the substance is distant to the site of release, and the substance is carried to the effector site through the blood stream). In addition to the concentration of the growth factor, the timing of release is as important to determine their effectiveness. As these growth factors exert their effects by cell-surface receptor binding, the appropriate receptor on the responding cells must be present at the time of release in order for the biologic effect to occur. Table 9-2 summarizes the principal growth factors found in healing wounds and their known effects on cells participating in the healing process. Growth factors have divergent actions on different cells; they can be chemotaxative to one cell type while stimulating replication of a different cell type. Little is known about the ratio of growth factor concentrations, which may be as important as the absolute concentration of individual growth factors.

Growth factors act on cells via surface receptor binding. Various receptor types have been described, such as ion channels, G-protein linked, or enzyme linked. The response elicited in the cell is usually one of phosphorylation or dephosphorylation of second-messenger molecules through the action of phosphatases or kinases, resulting in activation or deactivation of proteins in the cytosol or nucleus of the target cell. Phosphorylation of nuclear proteins is followed by the initiation of transcription of target genes. The signal is stopped by internalization of the receptor-ligand complex.

**Wound Contraction**

All wounds undergo some degree of contraction. For wounds that do not have surgically approximated edges, the area of the wound will be decreased by this action (healing by secondary intention). The myofibroblast has been postulated as the major cell responsible for contraction, and it differs from the normal fibroblast in that it possesses a cytoskeletal structure. Typically this cell contains α-smooth muscle actin in thick bundles called stress fibers, giving myofibroblasts contractile capability. The α-smooth muscle actin is undetectable until day 6, and then it is increasingly expressed for the next 15 days of wound healing. After 4 weeks, this expression fades, and the cells are believed to undergo apoptosis. A puzzling point is that the identification of myofibroblasts in the wound does not correspond directly to the initiation of wound contraction, which starts almost immediately after injury. It is believed that fibroblasts might play a role in contraction. In vitro, fibroblasts placed in a collagen lattice actively move in the lattice and contract it without expressing stress fibers. It is postulated that the movement of cells with concomitant reorganization of the cytoskeleton is responsible for contraction.

**HERITABLE DISEASES OF CONNECTIVE TISSUE**

Heritable diseases of connective tissue consist of a group of generalized, genetically determined, primary disorders of one of the elements of connective tissue: collagen, elastin, or mucopolysaccharide. Five major types, Ehlers-Danlos syndrome, Marfan’s syndrome, osteogenesis imperfecta, epidermolysis bullosa, and acrodermatitis enteropathica, will be discussed, as each provides unique challenges to the surgeon.

**Ehlers-Danlos Syndrome**

Ehlers-Danlos syndrome (EDS) is a group of 10 disorders that present as a defect in collagen formation. Over half of the
<table>
<thead>
<tr>
<th>GROWTH FACTOR</th>
<th>WOUND CELL ORIGIN</th>
<th>CELLULAR AND BIOLOGIC EFFECTS</th>
</tr>
</thead>
</table>
| PDGF             | Platelets, macrophages, monocytes, smooth muscle cells, endothelial cells          | Chemotaxis: fibroblasts, smooth muscle, monocytes, neutrophils  
Mitogenesis: fibroblasts, smooth muscle cells  
Stimulation of angiogenesis  
Stimulation of collagen synthesis  
Enhance reepithelialization  
Modulate tissue remodeling |
| FGF              | Fibroblasts, endothelial cells, keratinocytes, smooth muscle cells, chondrocytes   | Stimulation of angiogenesis (by stimulation of endothelial cell proliferation and migration)  
Mitogenesis: mesoderm and neuroectoderm |
| HGF              | Fibroblasts                                                                        | Stimulates fibroblasts, keratinocytes, chondrocytes, myoblasts  
Suppresses inflammation, granulation tissue formation, angiogenesis, reepithelialization |
| Keratinocyte growth factor | Keratinocytes, fibroblasts                                                      | Significant homology with FGF; stimulates keratinocytes |
| EGF              | Platelets, macrophages, monocytes (also identified in salivary glands, duodenal glands, kidney, and lacrimal glands) | Stimulates proliferation and migration of all epithelial cell types |
| TGF-α            | Keratinocytes, platelets, macrophages                                             | Homology with EGF; binds to EGF receptor  
Mitogenic and chemotactic for epidermal and endothelial cells |
| TGF-β (three isoforms: β1, β2, β3) | Platelets, T lymphocytes, macrophages, monocytes, neutrophils, fibroblasts, keratinocytes | Stimulates angiogenesis  
Stimulates leukocyte chemotaxis  
TGF-β1 stimulates wound matrix production (fibronectin, collagen glycosaminoglycans); regulation of inflammation  
TGF-β1 inhibits scar formation |
| Insulin-like growth factors (IGF-1, IGF-2) | Platelets (IGF-1 in high concentrations in liver; IGF-2 in high concentrations in fetal growth); likely the effector of growth hormone action | Promote protein/extracellular matrix synthesis  
Increase membrane glucose transport |
| Vascular endothelial growth factor | Macrophages, fibroblasts, endothelial cells, keratinocytes                  | Mitogen for endothelial cells (not fibroblasts)  
Stimulates angiogenesis  
Proinflammatory |
| IL-1             | Macrophages, leukocytes, keratinocytes, fibroblasts                             | Proinflammatory  
Stimulates angiogenesis, reepithelialization, tissue remodeling  
Enhances collagen synthesis |
| IL-4, IL-6       | Leukocytes  
Fibroblasts, endothelial cells, macrophages, keratinocytes  
Keratinocytes, fibroblasts | Stimulates inflammation, angiogenesis, reepithelialization, collagen deposition, tissue remodeling  
Stimulates granulation tissue formation, keratinocyte differentiation, reepithelialization  
Stimulates angiogenesis  
Stimulates inflammation, angiogenesis, collagen deposition |
| Activin          | Keratinocytes, fibroblasts                                                       | |
| Angiopoietin-1/-2CX3CL1 | Endothelial cells  
Macrophages, endothelial cells                                               | Stimulates inflammation, angiogenesis, collagen deposition |
| Granulocyte-macrophage colony-stimulating factor | Macrophage/monocytes, endothelial cells, fibroblasts                  | Stimulates macrophage differentiation/proliferation |

CX3CL1 = chemokine (C-X3-C motif) ligand; EGF = epidermal growth factor; FGF = fibroblast growth factor; HGF = hepatocyte growth factor; IL = interleukin; PDGF = platelet-derived growth factor; TGF = transforming growth factor.
affected patients manifest genetic defects encoding \( \alpha \)-chains of collagen type V, causing it to be either quantitatively or structurally defective. These changes lead to “classic” EDS with phenotypic findings that include thin, friable skin with prominent veins, easy bruising, poor wound healing, atrophic scar formation, recurrent hernias, and hyperextensible joints. Gastrointestinal problems include bleeding, hiatal hernia, intestinal diverticulae, and rectal prolapse. Small blood vessels are fragile, making suturing difficult during surgery. Large vessels may develop aneurysms, varicosities, or arteriovenous fistulas or may spontaneously rupture.\textsuperscript{31-33} Table 9-3 presents a description of EDS subtypes, including a recently recognized autosomal recessive form characterized by tenasin-X deficiency. The defect is a quantitative loss of protein, resulting in phenotypic changes similar to those observed in other types of EDS.

EDS must be considered in every child with recurrent hernias and coagulopathy, especially when accompanied by platelet abnormalities and low coagulation factor levels. Inguinal hernias in these children resemble those seen in adults. Great care should be taken to avoid tearing the skin and fascia. The transversalis fascia is thin, and the internal ring is greatly dilated. Like adults, hernia repair in these patients with the use of mesh or felt may result in a lower incidence of recurrence.\textsuperscript{34}

Closing wounds in patients with EDS might represent a major challenge to the surgeon. Dermal wounds should be closed in two layers, approximated with the sutures under tension, and the stitches should be left in place twice as long as usual. In addition, external fixation with adhesive tape can help reinforce the scar and prevent stretching.\textsuperscript{35}

### Table 9-3

<table>
<thead>
<tr>
<th>TYPE</th>
<th>CLINICAL FEATURES</th>
<th>INHERITANCE</th>
<th>BIOCHEMICAL DEFECT</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Skin: soft, hyperextensible, easy bruising, fragile, atrophic scars; hypermobile joints; varicose veins; premature births</td>
<td>AD</td>
<td>Not known</td>
</tr>
<tr>
<td>II</td>
<td>Similar to type I, except less severe</td>
<td>AD</td>
<td>Not known</td>
</tr>
<tr>
<td>III</td>
<td>Skin: soft, not hyperextensible, normal scars; small and large joint hypermobility</td>
<td>AD</td>
<td>Not known</td>
</tr>
<tr>
<td>IV</td>
<td>Skin: thin, translucent, visible veins, normal scarring, no hyperextensibility; no joint hypermobility; arterial, bowel, and uterine rupture</td>
<td>AD</td>
<td>Type III collagen defect</td>
</tr>
<tr>
<td>V</td>
<td>Similar to type II</td>
<td>XLR</td>
<td>Not known</td>
</tr>
<tr>
<td>VI</td>
<td>Skin: hyperextensible, fragile, easy bruising; hypermobile joints; hypotonia; kyphoscoliosis</td>
<td>AR</td>
<td>Lysyl hydroxylase deficiency</td>
</tr>
<tr>
<td>VII</td>
<td>Skin: soft, mild hyperextensibility, no increased fragility; extremely lax joints with dislocations</td>
<td>AD</td>
<td>Type I collagen gene defect</td>
</tr>
<tr>
<td>VIII</td>
<td>Skin: soft, hyperextensible, easy bruising, abnormal scars with purple discoloration; hypermobile joints; generalized periodontitis</td>
<td>AD</td>
<td>Not known</td>
</tr>
<tr>
<td>IX</td>
<td>Skin: soft, lax; bladder diverticula and rupture; limited pronation and supination; broad clavicle; occipital horns</td>
<td>XLR</td>
<td>Lysyl oxidase defect with abnormal copper use</td>
</tr>
<tr>
<td>X</td>
<td>Similar to type II with abnormal clotting studies</td>
<td>AR</td>
<td>Fibronectin defect</td>
</tr>
<tr>
<td>TNx</td>
<td>Hypermobile joints, skin fragility</td>
<td>AR</td>
<td>Absence of tenasin X protein</td>
</tr>
</tbody>
</table>

AD = autosomal dominant; AR = autosomal recessive; XLR = X-linked recessive.


---

**Marfan’s Syndrome.** Patients with Marfan’s syndrome have tall stature, arachnodactyly, lax ligaments, myopia, scoliosis, pectus excavatum, and aneurysm of the ascending aorta. Patients who suffer from this syndrome also are prone to hernias. Skin may be hyperextensible but shows no delay in wound healing.\textsuperscript{36,37}

The genetic defect associated with Marfan’s syndrome is a mutation in the \( FBN1 \) gene, which encodes for fibrillin. Previously, it was thought that structural alteration of the microfibrillar system was responsible for the phenotypic changes seen with the disease. However, recent research indicates an intricate role that \( FBN1 \) gene products play in TGF-β signaling. These extracellular matrix molecules normally bind and regulate TGF-β signaling; abnormal \( FBN1 \) gene function may cause an increase in TGF-β signaling, particularly in the aortic wall.\textsuperscript{38}

**Osteogenesis Imperfecta**

Patients with osteogenesis imperfecta (OI) have brittle bones, osteopenia, low muscle mass, hernias, and ligament and joint laxity. OI is a result of a mutation in type I collagen. Mutations in prolidase, an enzyme responsible for cleaving c-terminal proline and hydroxyproline, may have a role in the disease. There are four major OI subtypes with mild to lethal manifestations. Patients experience dermal thinning and increased bruisability. Scarring is normal, and the skin is not hyperextensible. Surgery can be successful but difficult in these patients, as the bones fracture easily under minimal stress.\textsuperscript{31,34} Table 9-4 lists the various features associated with the clinical subtypes of OI.
Epidermolysis Bullosa

Epidermolysis bullosa (EB) is classified into four major subtypes: EB simplex, junctional EB, dystrophic EB, and Kindler’s syndrome. The first three are determined by location in various skin layers; the last can present as multiple blisters throughout different layers of skin. There are identified genetic defects for each subtype, but the overall phenotype is remarkably similar. The disease manifestations include impairment in tissue adhesion within the epidermis, basement membrane, or dermis, resulting in tissue separation and blistering with minimal trauma. Characteristic features of EB are blistering and ulceration. The recessively inherited dystrophic type is characterized by defects in the COL7A1 gene, encoding type 7 collagen, important for connecting the epidermis to the dermis, and therefore phenotypically resulting in blistering. Management of nonhealing wounds in patients with EB is a challenge, as their nutritional status is compromised because of oral erosions and esophageal obstruction. Surgical interventions include esophageal dilatation and gastrostomy tube placement. Dermal incisions must be meticulously placed to avoid further trauma to skin. The skin requires nonadhesive pads covered by a “bulky” dressing to avoid blistering.

Acrodermatitis Enteropathica

Acrodermatitis enteropathica (AE) is an autosomal recessive disease of children that causes an inability to absorb sufficient zinc from breast milk or food. The AE mutation affects zinc transport in the intestine by preventing zinc from binding to the transporter, with impaired granulation tissue formation, as zinc is a necessary cofactor for DNA polymerase and reverse transcriptase, as well as different responses to cytokines and growth factors among these different fibroblast populations. Ultimate anastomotic strength is not always related to the absolute amount of collagen, and the structure and arrangement of the collagen matrix may be more important.

Table 9-4

<table>
<thead>
<tr>
<th>TYPE</th>
<th>CLINICAL FEATURES</th>
<th>INHERITANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Mild bone fragility, blue sclera</td>
<td>Dominant</td>
</tr>
<tr>
<td>II</td>
<td>“Prenatal lethal”; crumpled long bones, thin ribs, dark blue sclera</td>
<td>Dominant</td>
</tr>
<tr>
<td>III</td>
<td>Progressively deforming; multiple fractures; early loss of ambulation</td>
<td>Dominant/ recessive</td>
</tr>
<tr>
<td>IV</td>
<td>Mild to moderate bone fragility; normal or gray sclera; mild short stature</td>
<td>Dominant</td>
</tr>
</tbody>
</table>


HEALING IN SPECIFIC TISSUES

Gastrointestinal Tract

Repair and healing of the gastrointestinal tract is essential for the normal functions of the GI tract such as absorptive, barrier, and motor functions. Full-thickness GI injury healing remains an unresolved clinical issue. Healing of full-thickness GI wounds begins with a surgical or mechanical reapposition of the bowel ends, which is most often the initial step in the repair process. Sutures or staples are principally used, although various other means such as buttons, plastic tubes, and various wrappings have been attempted with variable success. Failure of healing results in dehiscence, leaks, and fistulas, which carry significant morbidity and mortality. Conversely, excessive healing can be just as troublesome, resulting in stricture formation and stenosis of the lumen.

The gross anatomic features of the GI tract are remarkably constant throughout most of its length. Within the lumen, the epithelium is supported by the lamina propria and underlying muscularis mucosa. The submucosa lies radially and circumferentially outside of these layers, is comprised of abundant collagenous and elastic fibers, and supports neural and vascular structures. Further toward the peritoneal surface of the bowel are the inner and outer muscle layers and ultimately a peritoneal extension, the serosa. The submucosa is the layer that imparts the greatest tensile strength and greatest suture-holding capacity, a characteristic that should be kept in mind during surgical repair of the GI tract. Additionally, serosal healing is essential for quickly achieving a watertight seal from the luminal side of the bowel. The importance of the serosa is underscored by the significantly higher rates of anastomotic failure observed clinically in segments of bowel that are extraperitoneal and lack serosa (i.e., the esophagus and rectum).

Injuries to all parts of the GI tract undergo the same sequence of healing as cutaneous wounds. However, there are some significant differences (Table 9-5). Mesothelial (serosal) and mucosal healing can occur without scarring. The early integrity of the anastomosis is dependent on formation of a fibrin seal on the serosal side, which achieves water tightness, and on the suture-holding capacity of the intestinal wall, particularly the submucosal layer. There is a significant decrease in marginal strength during the first week due to an early and marked collagenolysis. The lysis of collagen is carried out by collagenase derived from neutrophils, macrophages, and intraluminal bacteria. Recently, it has been shown that strains of Pseudomonas aeruginosa undergo phenotypic shifts characterized by higher collagenase secretion in an injured/anastomosed bowel environment. Collagenase activity occurs early in the healing process, and during the first 3 to 5 days, collagen breakdown far exceeds collagen synthesis. The integrity of the anastomosis represents equilibrium between collagen lysis, which occurs early, and collagen synthesis, which takes a few days to initiate (Fig. 9-5). Collagenase is expressed following injury in all segments of the GI tract, but it is much more marked in the colon compared to the small bowel. Collagen synthesis in the GI tract is carried out by both fibroblasts and smooth muscle cells. Colon fibroblasts produce greater amounts of collagen than skin fibroblasts, reflecting different phenotypic features, as well as different responses to cytokines and growth factors among these different fibroblast populations.
Technical Considerations. Traditional teaching holds that in order for an anastomosis to heal without complications it must be tension free, have an adequate blood supply, receive adequate nutrition, and be free of sepsis. Although sound principles for all wound healing, there are several considerations unique to GI anastomotic healing. From a technical viewpoint, the ideal method of suturing two ends of bowel together has not yet been identified. Although debate exists concerning methods of creating an anastomosis, clinically there has been no convincing evidence that any given technique has any advantage over another (i.e., hand-sutured vs. stapled, continuous vs. interrupted sutures, absorbable vs. nonabsorbable sutures, or single- vs. two-layer closure). A recent meta-analysis revealed that stapled ileocolic anastomoses have fewer leak rates than hand-constructed ones, but this might not apply to colo-colic or small bowel anastomoses.46

Fluid third spacing, tissue edema, and increased intraabdominal pressure secondary to overzealous fluid administration can result in blood flow compromise in small vessels at the edge of anastomosis and thus interfere with GI healing.47,48

Bone
Following any type of injury to bone, several changes take place at the site of injury to restore structural and functional integrity. Most of the phases of healing resemble those observed in dermal healing, but some notable individual characteristics apply to

Table 9-5. Comparison of wound healing in the gastrointestinal tract and skin

<table>
<thead>
<tr>
<th></th>
<th>GI TRACT</th>
<th>SKIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wound environment</td>
<td>pH Varies throughout GI tract in accordance with local exocrine secretions</td>
<td>Usually constant except during sepsis or local infection</td>
</tr>
<tr>
<td>Microorganisms</td>
<td>Aerobic and anaerobic, especially in the colon and rectum; problematic if they contaminate the peritoneal cavity</td>
<td>Skin commensals rarely cause problems; infection usually results from exogenous contamination or hematogenous spread</td>
</tr>
<tr>
<td>Shear stress</td>
<td>Intraluminal bulk transit and peristalsis exert distracting forces on the anastomosis</td>
<td>Skeletal movements may stress the suture line but pain usually acts as a protective mechanism preventing excess movement</td>
</tr>
<tr>
<td>Tissue oxygenation</td>
<td>Dependent on intact vascular supply and neocapillary formation</td>
<td>Circulatory transport of oxygen as well as diffusion</td>
</tr>
<tr>
<td>Collagen synthesis</td>
<td>Cell type Fibroblasts and smooth muscle cells</td>
<td>Fibroblasts</td>
</tr>
<tr>
<td>Lathyrogenes</td>
<td>d-Penicillamine has no effect on collagen cross-linking</td>
<td>Significant inhibition of cross-linking with decreased wound strength</td>
</tr>
<tr>
<td>Steroids</td>
<td>Contradictory evidence exists concerning their negative effect on GI healing; increased abscess in the anastomotic line may play a significant role</td>
<td>Significant decrease in collagen accumulation</td>
</tr>
<tr>
<td>Collagenase activity</td>
<td>Increased presence throughout GI tract after transection and reanastomosis; during sepsis, excess enzyme may promote dehiscence by decreasing suture-holding capacity of tissue</td>
<td>Not as significant a role in cutaneous wounds</td>
</tr>
<tr>
<td>Wound strength</td>
<td>Rapid recovery to preoperative level.</td>
<td>Less rapid than GI tissue</td>
</tr>
<tr>
<td>Scar formation</td>
<td>Definite scarring seen in fetal wound sites</td>
<td>Usually heals without scar formation in the fetus</td>
</tr>
</tbody>
</table>

Figure 9-5. Diagrammatic representation of the concept of GI wound healing as a fine balance between collagen synthesis and collagenolysis. The “weak” period when collagenolysis exceeds collagen synthesis can be prolonged or exacerbated by any factors that upset the equilibrium. (Reproduced with permission from Dunphy JE: Fundamentals of Wound Management in Surgery. New York, NY: Chirurgecom, Inc.; 1976.)
bone injuries. The initial stage of hematoma formation consists of an accumulation of blood at the fracture site, which also contains devitalized soft tissue, dead bone, and necrotic marrow. The next stage accomplishes the liquefaction and degradation of nonviable products at the fracture site. The normal bone adjacent to the injury site can then undergo revascularization, with new blood vessels growing into the fracture site. This is similar to the formation of granulation in soft tissue. The symptoms associated with this stage are characteristic of inflammation, with clinical evidence of swelling and erythema.

Three to 4 days following injury, soft tissue forms a bridge between the fractured bone segments in the next stage (soft callus stage). The soft tissue is deposited where neovascularization has taken place and serves as an internal splint, preventing damage to the newly laid blood vessels and achieving a fibrocartilaginous union. The soft callus is formed externally along the bone shaft and internally within the marrow cavity. Clinically, this phase of healing is characterized by the cessation of pain and inflammatory signs.

The next phase consists of mineralization of the soft callus and conversion to bone (hard callus stage). This may take up to 2 to 3 months and leads to complete bony union. The bone is now considered strong enough to allow weight bearing and will appear healed on radiographs. Then remodeling phase follows, in which the excessive callus is reabsorbed and the marrow cavity is recanalized. Remodeling allows for the correct transmission of forces and restores the contours of the bone.

As in dermal healing, the process of osseous union is mediated by soluble growth factors and cytokines. The most extensively studied group is the bone morphogenetic proteins (BMPs), which belong to the TGF-β superfamily. By stimulating the differentiation of mesenchymal cells into chondroblasts and osteoblasts, BMPs directly affect bone and cartilage repair. Other growth factors such as PDGF, TGF-β, TNF-α, and bFGF also participate in bony repair by mediating the inflammatory and proliferative phases of healing.

**Cartilage**

Cartilage consists of cells (chondrocytes) surrounded by an extracellular matrix made up of several proteoglycans, collagen fibers, and water. Unlike bone, cartilage is very avascular and depends on diffusion for transmittal of nutrients across the matrix. Additionally, the hypovascular perichondrium contributes substantially to the nutrition of the cartilage. Therefore, injuries to cartilage may be associated with permanent defects due to tenuous blood supply.

The healing response of cartilage depends on the depth of injury. In a superficial injury, there is disruption of the proteoglycan matrix and injury to the chondrocytes. There is no inflammatory response, but an increase in synthesis of proteoglycan and collagen dependent entirely on the chondrocyte. The healing power of cartilage is often inadequate, and overall regeneration is incomplete. Therefore, superficial cartilage injuries are slow to heal and often result in persistent structural defects.

In contrast to superficial injuries, deep injuries involve the underlying bone and soft tissue. This leads to the exposure of vascular channels of the surrounding damaged tissue that may help in the formation of granulation tissue. Hemorrhage allows for the initiation of the inflammatory response and the subsequent mediator activation of cellular function for repair. As the granulation tissue is laid down, fibroblasts migrate toward the wound and synthesize fibrous tissue that undergoes chondrification. Gradually, hyaline cartilage is formed, which restores the structural and functional integrity of the injured site.

**Tendon**

Tendons and ligaments are specialized structures that link muscle and bone, and bone and bone, respectively. They consist of parallel bundles of collagen interspersed with spindle cells. Tendons and ligaments can be subjected to a variety of injuries, such as laceration, rupture, and contusion. Due to the mobility of the underlying bone or muscles, the damaged ends usually separate. Tendon and ligament healing progresses in a similar fashion as in other areas of the body (i.e., through hematoma formation, organization, laying down of reparative tissue, and scar formation). Matrix is characterized by accumulation of types I and III collagen along with increased water, DNA, and glycosaminoglycan content. As the collagen fibers are organized, transmission of forces across the damaged portion can occur. Restoration of the mechanical integrity may never be equal to that of the undamaged tendon.

Tendon vasculature has a clear effect on healing. Hypovascular tendons tend to heal with less motion and more scar formation than tendons with better blood supply. The specialized cells, tenocytes, are metabolically very active and retain a large regenerative potential, even in the absence of vascularity. Cells on the tendon surface are identical to those within the sheath and play a role in tendon healing as well.

**Nerve**

Nerve injuries are very common, with an estimated 200,000 repairs performed every year in the United States. Peripheral nerves are a complex arrangement of axons, nonneuronal cells, and extracellular elements. There are three types of nerve injuries: neurapraxia (focal demyelination), axonotmesis (interruption of axonal continuity but preservation of Schwann cell basal lamina), and neurotmesis (complete transection). Following all types of injury, the nerve ends progress through a predictable pattern of changes involving three crucial steps: (a) survival of axonal cell bodies; (b) regeneration of axons that grow across the transected nerve to reach the distal stump; and (c) migration and connection of the regenerating nerve ends to the appropriate nerve ends or organ targets.

Phagocytes remove the degenerating axons and myelin sheath from the distal stump (Wallerian degeneration). Regenerating axonal sprouts extend from the proximal stump and probe the distal stump and the surrounding tissues. Schwann cells envelope and help in remyelinating the regenerating axons. Functional units are formed when the regenerating axons connect with the appropriate end targets. Several factors play a role in nerve healing, such as growth factors, cell adhesion molecules, and nonneuronal cells and receptors. Growth factors include nerve growth factor, brain-derived neurotrophic factor, basic and acidic fibroblastic growth factors, and neuroleukin. Cell adhesion molecules involved in nerve healing include nerve adhesion molecule, neuron-glia adhesion molecule, myelin adhesion glycoprotein, and N-cadherin. This complex interplay of growth factors and adhesion molecules helps in nerve regeneration.

**Fetal Wound Healing**

The main characteristic that distinguishes the healing of fetal wounds from that of adult wounds is the lack of scar formation. Understanding how fetal wounds achieve integrity without...
Evidence of scarring holds promise for the possible manipulation of unwanted fibrosis or excessive scar formation in adults. Although early fetal healing is characterized by the absence of scarring and resembles tissue regeneration, there is a phase of transition during gestational life when a more adult-like healing pattern emerges. This so-called “transition wound” occurs at the beginning of the third trimester, and during this period, there is scarless healing; however, there is a loss of the ability to regenerate skin appendages. Eventually a classic, adult-patterned healing with scar formation occurs exclusively, although overall healing continues to be faster than in adults.

There are a number of characteristics that may influence the differences between fetal and adult wounds. These include wound environment, inflammatory responses, differential growth factor profiles, and wound matrix.

**Wound Environment.** The fetus is bathed in a sterile, temperature-stable fluid environment, although this alone does not explain the observed differences. Experiments have demonstrated that scarless healing may occur outside of the amniotic fluid environment, and conversely, scars can form in utero.

**Inflammation.** The extent and robustness of the inflammatory response correlates directly with the amount of scar formation in all healing wounds. Reduced fetal inflammation due to the immaturity of the fetal immune system may partially explain the lack of scarring observed. Not only is the fetus neutropenic, but fetal wounds also contain lower numbers of PMNs and macrophages.

**Growth Factors.** Fetal wounds are notable for the absence of TGF-β, which may have a significant role in scarring. Blocking TGF-β1 or TGF-β2 using neutralizing antibodies considerably reduces scar formation in adult wounds. Exogenous application of TGF-β3 downregulates TGF-β1 and TGF-β2 levels at the wound site with a resultant reduction in scarring. Thus, the balance between the concentration and/or activity of TGF-β isoforms may be important for regulating scar production.

**Wound Matrix.** The fetal wound is characterized by excessive and extended hyaluronic acid production, a high-molecular-weight glycosaminoglycan that is produced primarily by fibroblasts. Although adult wounds also produce hyaluronic acid, its synthesis is sustained only in the fetal wound. Components of amniotic fluid, most specifically fetal urine, have a unique ability to stimulate hyaluronic acid production. Fetal fibroblasts produce more collagen than adult fibroblasts, and the increased level of hyaluronic acid may aid in the orderly organization of collagen. As a result of these findings, hyaluronic acid is used topically to enhance healing and to inhibit postoperative adhesion formation. The collagen pattern of fetal wounds is reticular in nature and resembles surrounding tissue, while adult patterns express large bundles of parallel collagen fibrils oriented perpendicular to the surface.

**CLASSIFICATION OF WOUNDS**

Wounds are classified as either acute or chronic. By definition, an acute wound becomes chronic if healing is not achieved after 4 weeks of treatment. Acute wounds heal in a predictable manner and time frame as previously mentioned. The process occurs with few, if any, complications, and the end result is a well-healed wound. Surgical wounds can heal in several ways. An incised wound that is clean and closed by sutures is said to heal by primary intention. Often, because of bacterial contamination or tissue loss, a wound will be left open to heal by granulation tissue formation and contraction; this constitutes healing by secondary intention. Delayed primary closure, or healing by tertiary intention, represents a combination of the first two, consisting of the placement of sutures, allowing the wound to stay open for a few days, and the subsequent closure of the sutures (Fig. 9-6).

The healing spectrum of acute wounds is broad (Fig. 9-7). In examining the acquisition of mechanical integrity and strength during healing, the normal process is characterized by a constant and continual increase that reaches a plateau at some point after injury. In regular wounds, the maximal wound strength is reached after about 6 weeks of healing. A fully healed wound achieves only 75% to 80% of a normal tissue. Wounds with delayed healing are characterized by decreased wound-breaking strength in comparison to wounds that heal at a normal rate; however, they eventually achieve the same integrity and strength as wounds that heal normally. Conditions such as nutritional deficiencies, infections, or severe trauma cause delayed healing, which reverts to normal with correction of the underlying pathophysiology. Impaired healing is characterized by a failure to achieve mechanical strength equivalent to normally healed wounds. Patients with compromised immune systems such as those with diabetes, chronic steroid usage, or tissues damaged by radiotherapy are prone to this type of impaired healing. The surgeon must be aware of these situations and exercise great care in the placement of incision and suture selection, postoperative care, and adjunctive therapy to maximize the chances of healing without supervening complications.

In general, wounds heal by a combination of mechanisms, including connective tissue deposition, contraction, and...
epithelialization, depending on wound type. Surgically closed wounds need mostly epithelialization for healing, while open wounds require a combination of tissue contraction, connective tissue deposition, and epithelialization to a lesser extent. Chronic ulcers heal by secondary intention similar to open wounds.

Normal healing is affected by both systemic and local factors (Table 9-6). The clinician must be familiar with these factors and should attempt to counteract their deleterious effects. Complications occurring in wounds with higher risk can lead to failure of healing or the development of chronic, nonhealing wounds.

### Factors Affecting Wound Healing

#### Advanced Age

Most surgeons believe that aging produces intrinsic physiologic changes that result in delayed or impaired wound healing. Clinical experience with elderly patients tends to support this belief. Studies of hospitalized surgical patients show a direct correlation between older age and poor wound healing outcomes such as dehiscence and incisional hernia. However, these statistics fail to take into account underlying illnesses or diseases as a possible source of impaired wound healing in the elderly. The increased incidence of cardiovascular disease, metabolic diseases (diabetes mellitus, malnutrition, and vitamin deficiencies), and cancer, and the widespread use of drugs that impair wound healing may all contribute to the higher incidence of wound problems in the elderly. However, more recent clinical experience suggests that major operative interventions can be accomplished safely in the elderly.

The results of animal studies regarding the effects of aging on wound healing have yielded contradictory results. In healthy human volunteers, there was a significant delay of 1.9 days in the epithelialization of superficial skin defects in those older than 70 years of age when compared to younger volunteers. In the same volunteers, using a micro-model of fibroplasia, no difference in DNA or hydroxyproline wound accumulation could be demonstrated between the young and elderly groups; however, the young volunteers had a significantly higher amount of total α-amino nitrogen in their wounds, a reflection of total protein content of the wound. Thus, although wound collagen synthesis does not seem to be impaired with advanced age, noncollagenous protein accumulation at wounded sites is decreased with aging, which may impair the mechanical properties of scarring in elderly patients. Generally, in a relatively healthy person age will cause a delay in healing rather than nonhealing.

### Hypoxia, Anemia, and Hypoperfusion

Low oxygen tension has a profoundly deleterious effect on all aspects of wound healing. Fibroplasia, although stimulated initially by the hypoxic wound environment, is significantly impaired by local hypoxia. Optimal collagen synthesis requires oxygen as a cofactor, particularly for the hydroxylation steps. Increasing subcutaneous oxygen tension levels by increasing the fraction of inspired oxygen ($\text{FiO}_2$) of inspired air for brief periods during and immediately following surgery results in enhanced collagen deposition and in decreased rates of wound infection after elective surgery.

Major factors affecting local oxygen delivery include hypoperfusion either for systemic reasons (low volume or cardiac failure) or due to local causes (arterial insufficiency, local vasoconstriction, or excessive tension on tissues). The level of vasoconstriction of the subcutaneous capillary bed is exquisitely responsive to fluid status, temperature, and hyperactive sympathetic tone as is often induced by postoperative pain. Correction of these factors can have a remarkable influence on wound outcome, particularly on decreasing wound infection rates. Mild to moderate normovolemic anemia does not appear to adversely affect wound oxygen tension and collagen synthesis. However, profound anemia with 15% less hematocrit can interfere with wound healing.

### Steroids and Chemotherapeutic Drugs

Large doses or chronic usage of glucocorticoids reduce collagen synthesis and wound strength. The major effect of steroids is to inhibit the inflammatory phase of wound healing (angiogenesis, neutrophil and macrophage migration, and fibroblast proliferation) and the release of lysosomal enzymes. The stronger the anti-inflammatory effect of the steroid compound used, the greater the inhibitory effect on wound healing. Steroids used after the first 3 to 4 days after injury do not affect wound healing as severely as when they are used in the immediate postoperative phase.

---

**Table 9-6**

<table>
<thead>
<tr>
<th><strong>Factors affecting wound healing</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Systemic</strong></td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Nutrition</td>
</tr>
<tr>
<td>Trauma</td>
</tr>
<tr>
<td>Metabolic diseases</td>
</tr>
<tr>
<td>Immunosuppression</td>
</tr>
<tr>
<td>Connective tissue disorders</td>
</tr>
<tr>
<td>Smoking</td>
</tr>
<tr>
<td><strong>Local</strong></td>
</tr>
<tr>
<td>Mechanical injury</td>
</tr>
<tr>
<td>Infection</td>
</tr>
<tr>
<td>Edema</td>
</tr>
<tr>
<td>Ischemia/necrotic tissue</td>
</tr>
<tr>
<td>Topical agents</td>
</tr>
<tr>
<td>Ionizing radiation</td>
</tr>
<tr>
<td>Low oxygen tension</td>
</tr>
<tr>
<td>Foreign bodies</td>
</tr>
</tbody>
</table>
period. Therefore, if possible, their use should be delayed, or alternatively, forms with lesser anti-inflammatory effects should be administered.

In addition to their effect on collagen synthesis, steroids also inhibit epithelialization and contraction and contribute to increased rates of wound infection, regardless of the time of administration. Steroid-delayed healing of cutaneous wounds can be stimulated to epithelialize by topical application of vitamin A. Collagen synthesis of steroid-treated wounds also can be stimulated by vitamin A.

All chemotherapeutic antimitabolite drugs adversely affect wound healing by inhibiting early cell proliferation, wound DNA and protein synthesis, attenuation of the inflammatory phase, decrease fibrin deposition, and delay wound contraction, all of which are critical to successful healing. The effect is worse if these agents are given preoperatively; so a delay in the use of such drugs for about 2 weeks after injury appears to lessen wound healing impairment. Extravasation of most chemotherapeutic agents is associated with tissue necrosis, marked ulceration, and protracted healing at the affected site.

**Metabolic Disorders.** Diabetes mellitus is the best known of the metabolic disorders contributing to increased rates of wound infection and failure. Uncontrolled diabetes results in reduced inflammation, angiogenesis, and collagen synthesis. Additionally, the large- and small-vessel disease that is the hallmark of advanced diabetes contributes to hypoperfusion and local hypoxemia. Defects in granulocyte function, capillary ingrowth, and fibroblast proliferation all have been described in diabetes. Obesity, insulin resistance, hyperglycemia, and diabetic renal failure contribute significantly and independently to the impaired wound healing observed in diabetics. In wound studies on experimental diabetic animals, insulin restores collagen synthesis and granulation tissue formation to normal levels if given during the early phases of healing. In clean, noninfected, and well-perfused experimental wounds in human diabetic volunteers, type I diabetes mellitus was noted to decrease wound collagen accumulation in the wound, independent of the degree of glycemic control. Type 2 diabetic patients showed no effect on collagen accumulation when compared to healthy, age-matched controls. Furthermore, the diabetic wound appears to be lacking in sufficient growth factor levels, which signal normal healing. It remains unclear whether decreased collagen synthesis or an increased breakdown due to an abnormally high proteolytic wound environment is responsible.

Careful preoperative correction of blood sugar levels improves the outcome of wounds in diabetic patients. Increasing the inspired oxygen tension, judicious use of antibiotics, and correction of other coexisting metabolic abnormalities all can result in improved wound healing. Additionally, revascularization to local areas with chronic ulcers will aid in the healing process.

Uremia also has been associated with disordered wound healing and impairs defenses for infection. Experimentally, uremic animals demonstrate decreased wound collagen synthesis and breaking strength, causing delayed healing of intestinal anastomosis and abdominal wounds. The contribution of uremia alone to this impairment, rather than that of associated malnutrition, is difficult to assess. The clinical use of dialysis to correct the metabolic abnormalities and nutritional restoration should impact greatly on the wound outcome of such patients. In some uremic patients, wounds might be associated with abnormal deposition of calcium and phosphate in the tissue leading to uremic gangrene syndrome (calciphylaxis). In such patients, the wounds are extremely painful and difficult to heal.

Obesity is the largest growing public health problem in the United States and the world. Over 60% of Americans are overweight or obese. Uncomplicated obesity (i.e., in the absence of comorbid conditions such as cardiovascular disease, diabetes, or respiratory insufficiency) has by itself significant deleterious effects on wound healing. Visceral adiposity is active metabolically and immunologically and, through generation of proinflammatory cytokines and adipokines, leads to the development of the metabolic syndrome. Many of these molecules have effects on cells participating in the healing response. In nondiabetic obese rodents, wounds are mechanically weaker, and there is less dermal and reparative scar collagen. Preadipocytes infiltrate the dermis, and although they can evolve into fibroblasts, their regulatory mechanisms appear different from those of dermal or wound fibroblasts. Many studies indicate that obese patients have higher rates of perioperative complications, with estimates as high as 30% for wound dehiscence, 17% for surgical site infections, 30% for incisional hernias, 19% for seromas, 13% for hematomas, and 10% for fat necrosis. Increased subcutaneous fat was associated with a tenfold increased risk of surgery-related complications including anastomotic leaks, abdominal collection, and wound infections. In many studies, obesity is a constant and major risk factor for hernia formation and recurrence after repair. The mechanism by which obesity impairs wound healing awaits complete delineation.

**Nutrition.** The importance of nutrition in the recovery from traumatic or surgical injury has been recognized by clinicians since the time of Hippocrates. Poor nutritional intake or lack of individual nutrients significantly alters many aspects of wound healing. The clinician must pay close attention to the nutritional status of patients with wounds, since wound failure or wound infections may be no more than a reflection of poor nutrition. Although the full interaction of nutrition and wound healing is still not fully understood, efforts are being made to develop wound-specific nutritional interventions and institute the pharmacologic use of individual nutrients as modulators of wound outcomes.

Experimental rodents fed either a 0% or 4% protein diet have impaired collagen deposition with a secondary decrease in skin and fascial wound-breaking strength and increased wound infection rates. Induction of energy-deficient states by providing only 50% of the normal caloric requirement leads to decreased granulation tissue formation and matrix protein deposition in rats. Acute fasting in rats markedly impairs collagen synthesis while decreasing procollagen mRNA.

Clinically, it is extremely rare to encounter pure energy or protein malnutrition, and the vast majority of patients exhibit combined protein-energy malnutrition. Such patients have diminished hydroxyproline accumulation (an index of collagen deposition) into subcutaneously implanted polystyrene tubes when compared to normally nourished patients (Fig. 9-8). Furthermore, malnutrition correlates clinically with enhanced rates of wound complications and increased wound failure following diverse surgical procedures. This reflects impaired healing response as well as reduced cell-mediated immunity, phagocytosis, and intracellular killing of bacteria by macrophages and neutrophils during protein-calorie malnutrition.

Two additional nutrition-related factors warrant discussion. First, the degree of nutritional impairment need not be
long-standing in humans, as opposed to the experimental situation. Thus, patients with brief preoperative illnesses or reduced nutrient intake in the period immediately preceding the injury or operative intervention will demonstrate impaired fibroplasia.\(^7\)\(^7\)\(^8\) Second, brief and not necessarily intensive nutritional intervention, either via the parenteral or enteral route, can reverse or prevent the decreased collagen deposition noted with malnutrition or with postoperative starvation.\(^7\)\(^9\)

The possible role of single amino acids in enhanced wound healing has been studied for the last several decades. Arginine appears most active in terms of enhancing wound fibroplasia. Arginine deficiency results in decreased wound-breaking strength and wound-collagen accumulation in chow-fed rats. Rats that are given 1% arginine HCl supplementation, and therefore are not arginine-deficient, have enhanced wound-breaking strength and collagen synthesis when compared to chow-fed controls.\(^8\)\(^0\) Studies have been carried out in healthy human volunteers to examine the effect of arginine supplementation on collagen accumulation. Young, healthy, human volunteers (age 25–35 years) were found to have significantly increased wound-collagen deposition following oral supplementation with either 30 g of arginine aspartate (17 g of free arginine) or 30 g of arginine Hall (24.8 g of free arginine) daily for 14 days.\(^8\)\(^1\) In a study of healthy older humans (age 67–82 years), daily supplements of 30 g of arginine aspartate for 14 days resulted in significantly enhanced collagen and total protein deposition at the wound site when compared to controls given placebos. There was no enhanced DNA synthesis present in the wounds of the arginine-supplemented subjects, suggesting that the effect of arginine is not mediated by an inflammatory mode of action.\(^8\)\(^2\) In this and later studies, arginine supplementation, whether administered orally or parenterally, had no effect on the rate of epithelialization of a superficial skin defect. This further suggests that the main effect of arginine on wound healing is to enhance wound collagen deposition. Recently, a dietary supplemental regimen of arginine, β-hydroxy-β-methylbutyrate, and glutamine was found to significantly and specifically enhance collagen deposition in elderly, healthy human volunteers when compared to an isocaloric, isonitrogenous supplement (Fig. 9-9).\(^3\)\(^3\) As increases in breaking strength during the first weeks of healing are directly related to new collagen synthesis, arginine supplementation may result in an improvement in wound strength as a consequence of enhanced collagen deposition.

The vitamins most closely involved with wound healing are vitamin C and vitamin A. Scurvy or vitamin C deficiency leads to a defect in wound healing, particularly via a failure in collagen synthesis and cross-linking. Biochemically, vitamin C is required for the conversion of proline and lysine to hydroxyproline and hydroxylysine, respectively. Vitamin C deficiency has also been associated with an increased incidence of wound infection, and if wound infection does occur, it tends to be more severe. These effects are believed to be due to an associated impairment in neutrophil function, decreased complement activity, and decreased walling-off of bacteria secondary to insufficient collagen deposition. The recommended dietary allowance is 60 mg daily. This provides a considerable safety margin for most healthy nonsmokers. In severely injured or extensively burned patients, this requirement may increase to as high as 2 g daily. There is no evidence that excess vitamin C is toxic; however, there is no evidence that supratherapeutic doses of vitamin C are of any benefit.\(^8\)\(^4\)

Vitamin A deficiency impairs wound healing, while supplemental vitamin A benefits wound healing in nondeficient humans and animals. Vitamin A increases the inflammatory
response in wound healing, probably by increasing the lability of lysosomal membranes. There is an increased influx of macrophages, with an increase in their activation and increased collagen synthesis. Vitamin A directly increases collagen production and epidermal growth factor receptors when it is added in vitro to cultured fibroblasts. As mentioned before, supplemental vitamin A can reverse the inhibitory effects of corticosteroids on wound healing. Vitamin A also can restore wound healing that has been impaired by diabetes, tumor formation, cyclophosphamide, and radiation. Serious injury or stress leads to increased vitamin A requirements. In the severely injured patient, supplemental doses of vitamin A have been recommended. Doses ranging from 25,000 to 100,000 IU per day have been advocated.

The connections between specific minerals and trace elements and deficits in wound healing are complex. Frequently, deficiencies are multiple and include macronutrient deficiencies. As with some of the vitamins described earlier, the specific trace element may function as a cofactor or part of an enzyme that is essential for homeostasis and wound healing. Clinically, preventing deficiencies is often easier to accomplish than diagnosing them.

Zinc is the most well-known element in wound healing and has been used empirically in dermatologic conditions for centuries. It is essential for wound healing in animals and humans. There are over 150 known enzymes for which zinc is either an integral part or an essential cofactor, and many of these enzymes are critical to wound healing. With zinc deficiency, there is decreased fibroblast proliferation, decreased collagen synthesis, impaired overall wound strength, and delayed epithelialization. These defects are reversed by zinc supplementation. To date, no study has shown improved wound healing with zinc supplementation in patients who are not zinc deficient.

Infections. Wound infections continue to represent a major medical problem, both in terms of how they affect the outcome of surgical procedures (surgical site infections), and for their impact on the length of hospital stay and medical costs. Many otherwise successful surgical operations fail because of the development of wound infections. The occurrence of infections is of major concern when implants are used, and their occurrence may lead to the removal of the prosthetic material, thus subjecting the patient to further operations and severe risk of morbidity and mortality. Infections can weaken an abdominal closure or hernia repair and result in wound dehiscence or recurrence of the hernia. Cosmetically, infections can lead to disfiguring, unsightly, or delayed closures.

Exhaustive studies have been undertaken that examine the appropriate prophylactic treatment of operative wounds. Bacterial contaminants normally present on skin are prevented from entry into deep tissues by intact epithelium. Surgery breaches the intact epithelium, allowing bacteria access to these tissues and the bloodstream. Antibiotic prophylaxis is most effective when adequate concentrations of antibiotic are present in the tissues at the time of incision, and assurance of adequate preoperative antibiotic dosing and timing has become a significant hospital performance measure. The addition of antibiotics after operative contamination has occurred clearly is ineffective in preventing postoperative wound infections.

Studies that compare operations performed with and without antibiotic prophylaxis demonstrate that classes II, III, and IV procedures (see below) treated with appropriate prophylactic antibiotics have only one third the wound infection rate of previously reported untreated series. More recently, repeat dosing of antibiotics has been shown to be essential in decreasing postoperative wound infections in operations with durations exceeding the biochemical half-life (t1/2) of the antibiotic or in which there is large-volume blood loss and fluid replacement. In lengthy cases, those in which prosthetic implants are used, or when unexpected contamination is encountered, additional doses of antibiotic may be administered for 24 hours postoperatively.

Selection of antibiotics for use in prophylaxis should be tailored to the type of surgery to be performed, operative contaminants that might be encountered during the procedure, and the profile of resistant organisms present at the institution where the surgery is performed. The continuing widespread appearance of methicillin-resistant Staphylococcus aureus (MRSA), and vancomycin-resistant enterococci (VRE) has significantly restricted the selection of these agents for routine use. Surgery-specific treatment guidelines are provided in Table 9-7.

Patients with prosthetic heart valves or any implanted vascular or orthopedic prostheses should receive antibiotic prophylaxis prior to any procedure in which significant contaminants are encountered. Dental procedures require prophylaxis with broad-spectrum penicillins or amoxicillin, while urologic instrumentation should be pretreated with a second-generation cephalosporin. Patients with prostheses who undergo gastrointestinal surgery should receive anaerobic coverage combined with a cephalosporin. Nasal screening and decolonization for S. aureus carriers are recommended for selected procedures (i.e., cardiac, orthopedic, neurosurgical procedures with implants).

The incidence of wound infection is about 5% to 10% nationwide and has not changed during the last few decades. Quantitatively, it has been shown that if the wound is contaminated with >10^5 microorganisms per gram of tissue, the risk of wound infection is markedly increased, but this threshold may be much lower in the presence of foreign materials. The source of pathogens for the infection is usually the endogenous flora of the patient's skin, mucous membranes, or from hollow organs. The most common organisms responsible for wound infections in order of frequency are Staphylococcus species, coagulase-negative Streptococcus, enterococci, and Escherichia coli. The incidence of wound infection bears a direct relationship to the degree of contamination that occurs during the operation from the disease process itself (clean—class I, clean contaminated—class II, contaminated—class III, and dirty—class IV). Many factors contribute to the development of postoperative wound infections. Most surgical wound infections become apparent within 7 to 10 days postoperatively, although a small number manifest years after the original operative intervention. With the hospital stay becoming shorter and shorter, many infections are detected in the outpatient setting, leading to underreporting of the true incidence of wound infections absent intensive surveillance. There has been much debate about the actual definition of wound infection. The narrowest definition would include wounds that drain purulent material with bacteria identified on culture. The broader definition would include all wounds draining pus, whether or not the bacteriologic studies are positive; wounds that are opened by the surgeon; and wounds that the surgeon considers infected.

Anatomically, wound infections can be classified as superficial incisional, deep incisional, and organ/space wound infections, involving fascia, muscle, or the abdominal cavity.
### Table 9-7
Antimicrobial prophylaxis for surgery

<table>
<thead>
<tr>
<th>NATURE OF OPERATION</th>
<th>COMMON PATHOGENS</th>
<th>RECOMMENDED ANTIMICROBIALS</th>
<th>ADULT DOSAGE BEFORE SURGERY²¹</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cardiac</strong></td>
<td><em>Staphylococcus aureus, S. epidermidis</em></td>
<td>Cefazolin or Cefuroxime or Vancomycin⁴</td>
<td>1–2 g IV²,³ 1.5 g IV³ 1 g IV</td>
</tr>
<tr>
<td><strong>Gastrointestinal</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Esophageal/gastroduodenal</td>
<td>Enteric gram-negative bacilli, gram-positive cocci</td>
<td>High risk⁵ only: cefazolin⁶</td>
<td>1–2 g IV²</td>
</tr>
<tr>
<td>Biliary tract</td>
<td>Enteric gram-negative bacilli, enterococci, clostridia</td>
<td>High risk⁷ only: cefazolin⁶,⁸</td>
<td>1–2 g IV²</td>
</tr>
<tr>
<td>Colorectal</td>
<td>Enteric gram-negative bacilli, anaerobes, enterococci</td>
<td>Oral: neomycin + erythromycin base⁹ or metronidazole⁹ Parenteral: cefoxitin⁶ or Cefotetan⁶ or Cefazolin + Metronidazole⁶ or Ampicillin/sulbactam</td>
<td>see note 9</td>
</tr>
<tr>
<td>Appendectomy, nonperforated¹¹</td>
<td>Same as for colorectal</td>
<td>Cefoxitin⁶ or cefotetan⁶ or Cefazolin⁶ + Metronidazole</td>
<td>1–2 g IV 1–2 g IV² 0.5 g IV 0.5 g IV</td>
</tr>
<tr>
<td><strong>Genitourinary</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cystoscopy alone</td>
<td>Enteric gram-negative bacilli, enterococci</td>
<td>High risk only¹²: ciprofloxacin¹⁰ or Trimethoprim-sulfamethoxazole</td>
<td>500 mg PO or 400 mg IV 1 DS tablet</td>
</tr>
<tr>
<td>Cystoscopy with manipulation or upper tract instrumentation¹³</td>
<td>Enteric gram-negative bacilli, enterococci</td>
<td>Ciprofloxacin¹⁰ or Trimethoprim-sulfamethoxazole</td>
<td>500 mg PO or 400 mg IV 1 DS tablet</td>
</tr>
<tr>
<td>Open or laparoscopic surgery¹⁴</td>
<td>Enteric gram-negative bacilli, enterococci</td>
<td>Cefazolin⁶</td>
<td>1–2 g IV²</td>
</tr>
<tr>
<td><strong>Gynecologic and obstetric</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vaginal, abdominal, or laparoscopic hysterectomy</td>
<td>Enteric gram-negative bacilli, anaerobes, group B streptococci, enterococci</td>
<td>Cefazolin⁶ or cefoxitin⁶ or cefotetan⁶ or Ampicillin/sulbactam⁶,¹⁰</td>
<td>1–2 g IV² 3 g IV</td>
</tr>
<tr>
<td>Cesarean section</td>
<td>Same as for hysterectomy</td>
<td>Cefazolin⁶</td>
<td>1–2 g IV²</td>
</tr>
<tr>
<td>Abortion, surgical</td>
<td>Same as for hysterectomy</td>
<td>Doxycycline</td>
<td>300 mg PO¹⁵</td>
</tr>
<tr>
<td><strong>Head and neck surgery</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incisions through oral or pharyngeal mucosa</td>
<td>Anaerobes, enteric gram-negative bacilli, <em>S. aureus</em></td>
<td>Clindamycin or Cefazolin + Metronidazole or Ampicillin/sulbactam¹⁰</td>
<td>600–900 mg IV 1–2 g IV² 0.5 g IV 3 g IV</td>
</tr>
<tr>
<td><strong>Neurosurgery</strong></td>
<td><em>S. aureus, S. epidermidis</em></td>
<td>Cefazolin</td>
<td>1–2 g IV²</td>
</tr>
<tr>
<td><strong>Ophthalmic</strong></td>
<td><em>S. epidermidis</em>, <em>S. aureus</em>, streptococci, enteric gram-negative bacilli, <em>Pseudomonas</em> spp.</td>
<td>Gentamicin, tobramycin, ciprofloxacin, gatifloxacin, levofloxacin, moxifloxacin, ofloxacin or neomycin-gramicidin-polyoxin B OR cefazolin</td>
<td>Multiple drops topically over 2 to 24 hours 100 mg subconjunctivally</td>
</tr>
<tr>
<td><strong>Orthopedic</strong></td>
<td><em>S. aureus, S. epidermidis</em></td>
<td>Cefazolin¹⁶ or Vancomycin¹⁶</td>
<td>1–2 g IV² 1 g IV</td>
</tr>
<tr>
<td><strong>Thoracic (noncardiac)</strong></td>
<td><em>S. aureus, S. epidermidis</em>, streptococci, enteric gram-negative bacilli</td>
<td>Cefazolin or Ampicillin/sulbactam¹⁰ or Vancomycin⁴</td>
<td>1–2 g IV² 3 g IV 1 g IV</td>
</tr>
</tbody>
</table>

(Continued)
About three-fourths of all wound infections are superficial, involving skin and subcutaneous tissue only. Clinical diagnosis is easy when a postoperative wound looks edematous and erythematous and is tender. Often the presentation is more subtle, and the development of postoperative fever, usually low-grade; the development of a mild and unexplained leukocytosis; or the presence of undue incisional pain should direct attention to the wound. Inspection of the wound is most useful in detecting subcutaneous and skin layers to the full extent of the infected pocket. Samples should be taken for aerobic and anaerobic cultures, with very few patients requiring antibiotic therapy. Patients who are immunosuppressed (diabetics and those on steroids or chemotherapeutic agents), who have evidence of tissue penetration or systemic toxicity, or who have had prosthetic devices inserted (vascular grafts, heart valves, artificial joints, or mesh) should be treated with systemic antibiotics.92

Deep wound infections arise immediately adjacent to the fascia, either above or below it, and often have an intra-abdominal component. Most intra-abdominal infections do not, however, communicate with the wound. Deep infections present with fever and leukocytosis. The incision may drain pus spontaneously, or the intra-abdominal extension may be recognized following the drainage of what was thought to be a superficial wound infection, but pus draining between the fascial sutures will be noted. Sometimes wound dehiscence will occur.

The most dangerous of the deep infections is necrotizing fasciitis. It results in high mortality, particularly in the elderly. This is an invasive process that involves the fascia and leads to secondary skin necrosis. Pathophysiologically, it is a septic thrombosis of the vessels between the skin and the deep layers.
The skin demonstrates hemorrhagic bullae and subsequent frank necrosis, with surrounding areas of inflammation and edema. The fascial necrosis is usually wider than the skin involvement or than the surgeon estimates on clinical grounds. The patient is toxic and has high fever, tachycardia, and marked hypovolemia, which if uncorrected, progresses to cardiovascular collapse. Bacteriologically, this is a mixed infection, and samples should be obtained for Gram stain smears and cultures to aid in diagnosis and treatment. As soon as bacteriologic studies have been obtained, high-dose penicillin treatment needs to be started (20–40 million U/d intravenously) due to concern over the presence of Clostridia perfringens and other related species; broad-spectrum antibiotics should be added and the regimen modified based on culture results. Cardiovascular resuscitation with electrolyte solutions, blood, and/or plasma is carried out as expeditiously as possible prior to induction of anesthesia. The aim of surgical treatment is thorough removal of all necrosed skin and fascia. If viable skin overlies necrotic fascia, multiple longitudinal skin incisions can be made to allow for excision of the devitalized fascia. Although removal of all necrotic tissue is the goal of the first surgical intervention, the distinction between necrotic and simply edematous tissue often is difficult. Careful inspection every 12 to 24 hours will reveal any new necrotic areas, and these need further debridement and excision. When all necrotic tissue has been removed and the infection has been controlled, the wounds may be covered with homografts or xenografts until definitive reconstruction and autografting can take place.

The mere presence of bacteria in an open wound, either acute or chronic, does not constitute an infection, because large numbers of bacteria can be present in the normal situation. In addition, the bacteria identified by cultures may not be representative of the bacteria causing the actual wound infection. There seems to be confusion as to what exactly constitutes wound infection. For purposes of clarity, we have to differentiate between contamination, colonization, and infection. Contamination is the presence of bacteria without multiplication, colonization is multiplication without host response, and infection is the presence of host response in reaction to deposition and multiplication of bacteria. The presence of a host response helps to differentiate between infection and colonization as seen in chronic wounds. The host response that helps in diagnosing wound infection comprises celluitis, abnormal discharge, delayed healing, change in pain, abnormal granulation tissue, bridging, and abnormal color and odor.

As discussed previously, neutrophils play a major role in preventing wound infections. Chronic granulomatous disease (CGD) comprises a genetically heterogeneous group of diseases in which the reduced nicotinamide adenine dinucleotide phosphate (NADPH)-dependent oxidase enzyme is deficient. This defect impairs the intracellular killing of microorganisms, leaving the patient liable to infection by bacteria and fungi. Afflicted patients have recurrent infections and form granulomas, which can lead to obstruction of the gastric antrum and genitourinary tracts and poor wound healing. Surgeons become involved when the patient develops infectious or obstructive complications.

The nitro blue tetrazolium (NBT) reduction test is used to diagnose CGD. Normal neutrophils can reduce this compound, while neutrophils from affected patients do not, facilitating the diagnosis via a colorimetric test. Clinically, patients develop recurrent infections such as pneumonia, lymphadenitis, hepatic abscess, and osteomyelitis. Organisms most commonly responsible are S. aureus, Aspergillus, Klebsiella, Serratia, or Candida. When CGD patients require surgery, a preoperative pulmonary function test should be considered since they are predisposed to obstructive and restrictive lung disease. Wound complications, mainly infection, are common. Sutures should be removed as late as possible since the wounds heal slowly. Abscess drains should be left in place for a prolonged period until the infection is completely resolved.

Hyperglycemia has been shown to be a significant risk factor of postoperative infections. Tight blood glucose control, beginning preoperatively and continued into the operating room and beyond, has been associated with significant reduction in infectious complications, in particular following cardiac surgery. Too tight of a glycemic control (80–100 mg/dL) appears to be associated with more complications and is as effective, if not less than, moderate control (120–180 mg/dL).

Another host factor that has been implicated in the development of superficial surgical site infection relates to the state of the subcutaneous capillary bed. Thomas K. Hunt had shown through several decades of work that this capillary bed is exquisitely sensitive to hypovolemia, hypothermia, and stress, leading to rapid vasoconstriction with secondary impaired oxygen delivery and increased rates of infection. Maintenance of euovolemia, core temperature above 36°C to 36.5°C, and pain control have all been shown singly and additively to reduce rates of wound infections. Another suggestion has been to increase inspired FiO₂ to 0.8 for the duration of the operation and in the immediate postoperative period, as a means of increasing subcutaneous tissue oxygen delivery. Although successful in most studies, there have also been negative results from such a single approach; this suggests that addressing volume, temperature, pain control, and oxygen delivery in concert may be the more fruitful approach to reduce surgical wound infections.

**Chronic Wounds**

Chronic wounds are defined as wounds that have failed to proceed through the orderly process that produces satisfactory anatomic and functional integrity or that have proceeded through the repair process without producing an adequate anatomic and functional result. The majority of wounds that have not healed in 3 months are considered chronic, although a duration as low as 4 weeks has been used to indicate chronicity. Skin ulcers, which usually occur in traumatized or vascular compromised soft tissue, are also considered chronic in nature, and proportionately are the major component of chronic wounds. In addition to the factors discussed earlier that can delay wound healing, other causative mechanisms may also play a role in the etiology of chronic wounds. Repeated trauma, poor perfusion or oxygenation, and/or excessive inflammation contribute to the causation and the perpetuation of the chronicity of wounds.

Unresponsiveness to normal regulatory signals also has been implicated as a predictive factor of chronic wounds. This may come about as a failure of normal growth factor synthesis, and thus an increased breakdown of growth factors within a wound environment that is markedly proteolytic because of overexpression of protease activity or a failure of the normal antiprotease inhibitor mechanisms. Fibroblasts from chronic wounds also have been found to have decreased proliferative potential, perhaps because of senescence or decreased expression of growth factor receptors. Chronic wounds occur due to various etiologic factors, and several of the most common are discussed later.
Malignant transformation of chronic ulcers can occur in any long-standing wound (Marjolin’s ulcer). Any wound that does not heal for a prolonged period of time is prone to malignant transformation. Malignant wounds are differentiated clinically from nonmalignant wounds by the presence of overturned wound edges (Fig. 9-10). In patients with suspected malignant transformations, biopsy of the wound edges must be performed to rule out malignancy. Cancers arising de novo in chronic wounds include both squamous and basal cell carcinomas.

Ischemic Arterial Ulcers. These wounds occur due to a lack of blood supply and are typically extremely painful in patients with pure ischemic ulcers. They usually are associated with other symptoms of peripheral vascular disease, such as history of intermittent claudication, rest pain, and color or trophic changes. These wounds commonly are present at the most distal portions of the extremities such as the interdigital clefts, although more proximal locations are also encountered. On examination, there may be diminished or absent pulses with decreased ankle-brachial index and poor formation of granulation tissue. Other signs of peripheral ischemia, such as dryness of skin, hair loss, scaling, and pallor can be present. The wound itself usually is shallow with smooth margins, and a pale base and surrounding skin may be present. The management of these wounds is two-pronged and includes revascularization and wound care. Nonhealing of these wounds is the norm unless successful revascularization is performed. In patients with ischemia and bed confinement, prevention of ulcer development is extremely important. Removal of restrictive stockings (in patients with critical ischemia), frequent repositioning, and surveillance are vital to preventing these ulcers.

Venous Stasis Ulcers. Although there is unanimous agreement that venous ulcers are due to venous stasis and increased venous pressure, there is less consensus as to what are the exact pathophysiologic pathways that lead to ulceration and impaired healing. On the microvascular level, there is alteration and distention of the dermal capillaries with leakage of fibrinogen into the tissues; polymerization of fibrinogen into fibrin cuffs leads to perivascular cuffing that can impede oxygen exchange, thus contributing to ulceration. These same fibrin cuffs and the leakage of macromolecules such as fibrinogen and α₂-macroglobulin trap growth factors and impede wound healing. Another hypothesis suggests that neutrophils adhere to the capillary endothelium and cause plugging with diminished dermal blood flow. Venous hypertension and capillary damage lead to extravasation of hemoglobin. The products of this breakdown are irritating and cause pruritus and skin damage. The resulting brownish pigmentation of skin combined with the loss of subcutaneous fat produces characteristic changes called lipodermatosclerosis. Regardless of the pathophysiologic mechanisms, the clinically characteristic picture is that of an ulcer that fails to reepithelialize despite the presence of adequate granulation tissue in a patient with skin color changes in the area of ulceration and signs of venous hypertension.

Venous stasis occurs due to increased venous hypertension caused by either venous insufficiency or venous outflow obstruction. Venous insufficiency can be due to any combination of deep, superficial, and perforator vein valvular reflux. The higher the ambulatory venous hypertension, the higher the chance of ulceration. Chronic venous ulcers are commonly painless. Stasis ulcers tend to occur at the sites of incompetent
perforators, the most common being above the medial malleolus, over Cockett’s perforator. Upon examination, the typical location combined with a history of venous incompetence and other skin changes is diagnostic. The wound usually is shallow with irregular margins and pigmented surrounding skin.

The cornerstone of treatment of venous ulcers is compression therapy, although the best method to achieve it remains controversial. Compression can be accomplished via rigid or flexible means. The most commonly used method is the rigid, zinc oxide–impregnated, nonelastic bandage. Others have proposed a four-layered bandage approach as a more optimal method of obtaining graduated compression. Wound care in these patients focuses on maintaining a moist wound environment, which can be achieved with hydrocolloids. Other, more modern approaches include use of vasoactive substances and growth factor application, as well as the use of skin substitutes. Recently, sprayed allogeneic keratinocytes and fibroblasts plus four-layer bandages have been shown to hasten healing when compared to compression alone. Addressing the causes of venous hypertension aids in the healing of venous ulcers. Unfortunately, recurrences are frequent despite preventative measures, largely because of patients’ lack of compliance.

Diabetic Wounds. Ten percent to 25% of diabetic patients run the risk of developing ulcers. There are approximately 50,000 to 60,000 amputations performed in diabetic patients each year in the United States. The major contributors to the formation of diabetic ulcers include neuropathy, foot deformity, and ischemia. It is estimated that 60% to 70% of diabetic ulcers are due to neuropathy. 15% to 20% are due to ischemia, and another 15% to 20% are due to a combination of both. The neuropathy is both sensory and motor and is secondary to persistently elevated glucose levels. The loss of sensory function allows unrecognized injury to occur from ill-fitting shoes, foreign bodies, or other trauma. The motor neuropathy or Charcot’s foot leads to collapse or dislocation of the interphalangeal or metatarsophalangeal joints, causing pressure on areas with little protection. There is also severe micro- and macrovascular circulatory impairment.

Once ulceration occurs, the chances of healing are poor and the chances of recurrent ulceration are high. The treatment of diabetic wounds involves local and systemic measures. Achievement of adequate blood sugar levels is very important. Most diabetic wounds are infected, and eradication of the infectious source is paramount to the success of healing. Treatment should address the possible presence of osteomyelitis and should employ antibiotics that achieve adequate levels both in soft tissue and bone. Wide debridement of all necrotic or infected tissue is another cornerstone of treatment. Off-loading of the ulcerated area by using specialized orthotic shoes or casts allows for ambulation while protecting the fragile wound environment. Topical application of PDGF and granulocyte-macrophage colony-stimulating factor has met with limited but significant success in achieving closure. The application of engineered skin allograft substitutes, although expensive, also has shown some significant success. Prevention and specifically foot care play an important role in the management of diabetics.

Decubitus or Pressure Ulcers. The incidence of pressure ulcers ranges from 2.7% to 9% in the acute care setting, in comparison to 2.4% to 23% in long-term care facilities. A pressure ulcer is a localized area of tissue necrosis that develops when soft tissue is compressed between a bony prominence and an external surface. Excessive pressure causes capillary collapse and impedes the delivery of nutrients to body tissues. Pressure ulcer formation is accelerated in the presence of friction, shear forces, and moisture. Other contributory factors in the pathogenesis of pressure ulcers include immobility, altered activity levels, altered mental status, chronic conditions, and altered nutritional status. The four stages of pressure ulcer formation are as follows: stage I, no blanching erythema of intact skin; stage II, partial-thickness skin loss involving epidermis or dermis or both; stage III, full-thickness skin loss, but not through the fascia; and stage IV, full-thickness skin loss with extensive involvement of muscle and bone.

The treatment of established pressure ulcers is most successful when carried out in a multidisciplinary manner by involving wound care teams consisting of physicians, nurses, dietitians, physical therapists, and nutritionists. Care of the ulcer itself comprises debridement of all necrotic tissue, maintenance of a favorable moist wound environment that will facilitate healing, relief of pressure, and addressing host issues such as nutritional, metabolic, and circulatory status. Debridement is most efficiently carried out surgically, but enzymatic proteolytic preparations and hydrotherapy also are used. The wound bed should be kept moist by employing dressings that absorb secretions but do not desiccate the wound. Operative repair, usually involving flap rotation, has been found to be useful in obtaining closure. Unfortunately, recurrence rates are extremely high, owing to the population at risk and the inability to fully address the causative mechanisms.

EXCESS HEALING

Clinically, excess healing can be as significant as delayed or nonhealing. It is likely that more operative interventions are required for correction of the morbidity associated with excessive healing than are required for wound failure. The clinical manifestations of exuberant healing are protean and differ in the skin (mutilating or debilitating scars, burn contractions), tendons (frozen repairs), the GI tract (strictures or stenoses), solid organs (cirrhosis, pulmonary fibrosis), or the peritoneal cavity (adhesive disease). Hypertrophic scars (HTSs) and keloids represent an overabundance of fibroplasia in the dermal healing process. HTSs rise above the skin level but stay within the confines of the original wound and often regress over time. Keloids rise above the skin level as well, but they extend beyond the border of the original wound and rarely regress spontaneously. Both HTSs and keloids occur after trauma to the skin and may be tender, pruritic, and cause a burning sensation. Keloids are 15 times more common in darker-pigmented ethnicities, with individuals of African, Spanish, and Asian ethnicities being especially susceptible. Men and women are equally affected. Genetically, the predilection to keloid formation appears to be autosomal dominant with incomplete penetration and variable expression.

HTSs usually develop within 4 weeks after trauma. The risk of HTS increases if epithelialization takes longer than 21 days, independent of site, age, and race. Rarely elevated more than 4 mm above the skin level, HTSs stay within the boundaries of the wound. They usually occur across areas of tension and flexor surfaces, which tend to be at right angles to joints or skin creases. The lesions are initially erythematous and raised and over time may evolve into pale, flatter scars.
Keloids can result from surgery, burns, skin inflammation, acne, chickenpox, zoster, folliculitis, lacerations, abrasions, tattoos, vaccinations, injections, insect bites, or ear piercing, or may arise spontaneously. Keloids tend to occur 3 months to years after the initial insult, and even minor injuries can result in large lesions. They vary in size from a few millimeters to large, pedunculated lesions with a soft to rubbery or hard consistency. While they project above surrounding skin, they rarely extend into underlying subcutaneous tissues. Certain body sites have a higher incidence of keloid formation, including the skin of the earlobe as well as the deltoid, presternal, and upper back regions. They rarely occur on eyelids, genitalia, palms, soles, or across joints. Keloids rarely involute spontaneously, and surgical intervention can lead to recurrence, often with a worse result (Table 9-8).

Histologically, both HTSs and keloids demonstrate increased thickness of the epidermis with an absence of rete ridges. There is an abundance of collagen and glycoprotein deposition. Normal skin has distinct collagen bundles, mostly parallel to the epithelial surface, with random connections between bundles by fine fibrillar strands of collagen. In HTS, the collagen bundles are flatter and more random, and the fibers are in a wavy pattern. In keloids, the collagen bundles are virtually nonexistent, and the fibers are connected haphazardly in loose sheets with a random orientation to the epithelium. The collagen fibers are larger and thicker, and myofibroblasts are generally absent. Keloidal fibroblasts have normal proliferation parameters but synthesize collagen at a rate 20 times greater than that observed in normal dermal fibroblasts, and 3 times higher than fibroblasts derived from HTS. Abnormal amounts of extracellular matrix such as fibronectin, elastin, and proteoglycans also are produced. The synthesis of fibronectin, which promotes clot generation, granulation tissue formation, and reepithelialization, decreases during the normal healing process; however, production continues at high levels for months to years in HTSs and keloids. This perturbed synthetic activity is mediated by altered growth factor expression. TGF-β expression is higher in HTS, and both HTS- and keloid-derived fibroblasts respond to lower concentrations of TGF-β than do normal dermal fibroblasts. HTSs also express increased levels of insulin-like growth factor-1, which reduces collagenase mRNA activity and increases mRNA for types I and II procollagen. Keloid fibroblasts have enhanced expression of TGF-β1 and TGF-β2, VEGF, and plasminogen activator inhibitor-1 and an increased number of PDGF receptors; they also have upregulated antiapoptotic gene expression, which can be differentially expressed within different areas of the same scar.

The underlying mechanisms that cause HTSs and keloids are not known. The immune system appears to be involved in the formation of both HTSs and keloids, although the exact relationship is unknown. Much is inferred from the presence of various immune cells in HTSs and keloids. For example, in both HTSs and keloids, keratinocytes express human leukocyte antigen (HLA)-2 and ICAM-1 receptors, which are absent.

### Table 9-8

<table>
<thead>
<tr>
<th></th>
<th>KELOID</th>
<th>HYPERTROPHIC SCAR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Incidence</strong></td>
<td>Rare</td>
<td>Frequent</td>
</tr>
<tr>
<td><strong>Ethnic groups</strong></td>
<td>African American, Asian, Hispanic</td>
<td>No predilection</td>
</tr>
<tr>
<td><strong>Prior injury</strong></td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Site predilection</strong></td>
<td>Neck, chest, ear lobes, shoulders, upper back</td>
<td>Anywhere</td>
</tr>
<tr>
<td><strong>Genetics</strong></td>
<td>Autosomal dominant with incomplete penetration</td>
<td>No</td>
</tr>
<tr>
<td><strong>Timing</strong></td>
<td>Symptom-free interval; may appear years after injury</td>
<td>4–6 weeks post injury</td>
</tr>
<tr>
<td><strong>Symptoms</strong></td>
<td>Pain, pruritus, hyperesthesia, growth beyond wound margins</td>
<td>Raised, some pruritus, respects wound confines</td>
</tr>
<tr>
<td><strong>Regression</strong></td>
<td>No</td>
<td>Frequent</td>
</tr>
<tr>
<td><strong>Contracture</strong></td>
<td>Rare</td>
<td>Frequent</td>
</tr>
<tr>
<td><strong>Histology</strong></td>
<td>Hypocellular, thick, wavy collagen fibers in random orientation</td>
<td>Parallel orientation of collagen fibers</td>
</tr>
</tbody>
</table>
in normal scar keratinocytes. Keloids also have increased deposition of immunoglobulins IgG, IgA, and IgM, and their formation correlates with serum levels of IgE. Antinuclear antibodies against fibroblasts, epithelial cells, and endothelial cells are found in keloids, but not HTSs. HTSs have higher T lymphocyte and Langerhans cell contents. There is also a larger number of mast cells present in both HTSs and keloids compared to normal scars. Another recently described cell population is the fibrocyte, a leukocyte subpopulation derived from peripheral mononuclear cells. Present in large numbers at the site of excess scarring, fibrocytes can stimulate fibroblast numbers and collagen synthesis. They also generate large numbers of cytokines, growth factors, and extracellular matrix proteins, which are characteristically upregulated in keloid tissue. Other mechanisms that may cause abnormal scarring include mechanical tension (although keloids often occur in areas of minimal tension) and prolonged irritation and/or inflammation that may lead to the generation of abnormal concentrations of profibrotic cytokines.

Treatment goals for keloid and HTS include restoration of function to the area, relief of symptoms, and prevention of recurrence. Many patients seek intervention due to cosmetic concerns. Because the underlying mechanisms causing keloids and HTSs remain unknown, many different modalities of treatment have been used without consistent success.  

Excision alone of keloids is subject to a high recurrence rate, ranging from 45% to 100%. Inclusion of the dermal advancing edge that characterizes keloids, use of incisions in skin tension lines, and tension-free closure all have been proposed to decrease recurrence rates. There are fewer recurrences when surgical excision is combined with other modalities such as intralesional corticosteroid injection, topical application of silicone sheets, or the use of radiation or pressure. Surgery is recommended for debulking large lesions or as second-line therapy when other modalities have failed. Silicone application is relatively painless and should be maintained for 24 hours a day for about 3 months to prevent rebound hypertrophy. It may be secured with tape or worn beneath a pressure garment. The mechanism of action is not understood, but increased hydration of the skin, which decreases capillary activity, inflammation, hyperemia, and collagen deposition, may be involved. Silicone is more effective than other occlusive dressings and is an especially good treatment for children and others who cannot tolerate the pain involved in other modalities.

Intralesional corticosteroid injections decrease fibroblast proliferation, collagen and glycosaminoglycan synthesis, the inflammatory process, and TGF-β levels. When used alone, however, there is a variable rate of response and recurrence; therefore, steroids are recommended as first-line treatment for keloids and second-line treatment for HTSs if topical therapies have failed. Intralesional injections are more effective on younger scars. They may soften, flatten, and give symptomatic relief to keloids, but they cannot make the lesions disappear, and they cannot narrow wide HTSs. Success is enhanced when used in combination with surgical excision. Serial injections every 2 to 3 weeks are required. Complications include skin atrophy, hypopigmentation, telangiectasias, necrosis, and ulceration.

Although radiation destroys fibroblasts, it has variable, unreliable results and produces poor results, with 10% to 100% recurrence when used alone. It is more effective when combined with surgical excision. The timing, duration, and dosage for radiation therapy are still controversial, but doses ranging from 1500 to 2000 rads appear effective. Given the risks of hyperpigmentation, pruritus, erythema, paresthesias, pain, and possible secondary malignancies, radiation should be reserved for adults with scars resistant to other modalities.

Pressure aids collagen maturation, flattens scars, and improves thinning and pliability. It reduces the number of cells in a given area, possibly by creating ischemia, which decreases tissue metabolism and increases collagenase activity. External compression is used to treat HTSs, especially after burns. Therapy must begin early, and a pressure between 24 and 30 mmHg must be achieved in order to exceed capillary pressure, yet preserve peripheral blood circulation. Garments should be worn for 23 to 24 hours a day for up to 1 or more years to avoid rebound hypertrophy. Scars older than 6 to 12 months respond poorly.

Topical retinoids also have been used as treatment for both HTSs and keloids, with reported responses of 50% to 100%. Intralesional injections of IFN-γ, a cytokine released by T lymphocytes, reduce collagen types I, II, and III by decreasing mRNA and possibly by reducing levels of TGF-β. As monotherapy, IFN-γ has failed because of high recurrence rates due to resistance to repeated injections. More recently, imiquimod, an immunomodulator that induces IFN-γ and other cytokines at the site of application, has been recommended following excision. Intralesional injections of chemotherapeutic agents such as 5-fluorouracil have been used both alone and in combination with steroids. The use of bleomycin or mitomycin C has been reported to achieve some success in older scars resistant to steroids.

Peritoneal Scarring. Peritoneal adhesions are fibrous bands of tissues formed between organs that are normally separated and/or between organs and the internal body wall. Most intra-abdominal adhesions are a result of peritoneal injury, either by a prior surgical procedure or due to intra-abdominal infection. Postmortem examinations demonstrate adhesions in 67% of patients with prior surgical procedures and in 28% with a history of intra-abdominal infection. Intra-abdominal adhesions are the most common cause (65%–75%) of small bowel obstruction, especially in the ileum. Operations in the lower abdomen have a higher chance of producing small bowel obstruction. Following rectal surgery, left colectomy, or total colectomy, there is an 11% chance of developing small bowel obstruction within 1 year, and this rate increases to 30% by 10 years. Adhesions also are a leading cause of secondary infertility in women and can cause substantial abdominal and pelvic pain. Adhesions account for 2% of all surgical admissions and 3% of all laparotomies in general surgery.

Adhesions form when the peritoneal surface is damaged due to surgery, thermal or ischemic injury, inflammation, or foreign body reaction. The injury disrupts the protective mesothelial cell layer lining the peritoneal cavity and the underlying connective tissue. The injury elicits an inflammatory response consisting of hyperemia, fluid exudation, release, and activation of white blood cells and platelets in the peritoneal cavity, activation of inflammatory cytokines, and the onset of the coagulation and complement cascades. Fibrin deposition occurs between the damaged but opposed serosal surfaces. These filmy adhesions often are transient and degraded by proteases of the fibrinolytic system, with restoration of the normal peritoneal surface. If insufficient fibrinolytic activity is present, permanent fibrous adhesions will form by collagen deposition within 1 week of the injury (Fig. 9-12).
There are two major strategies for adhesion prevention or reduction. Surgical trauma is minimized within the peritoneum by careful tissue handling, avoiding desiccation and ischemia, and spare use of cautery, laser, and retractors. Fewer adhesions form with laparoscopic surgical techniques due to reduced tissue trauma. The second major advance in adhesion prevention has been the introduction of barrier membranes and gels, which separate and create barriers between damaged mesothelial surfaces, allowing for adhesion-free healing. Currently, only three products are Food and Drug Administration (FDA) approved for reducing adhesion formation: Interceed (oxidized regenerated cellulose, indicated only in pelvic surgery), Seprafilm (a film composed of hyaluronic acid and carboxymethylcel lulose) that is usually applied below the incision, and Adept (4% icodextrin, a corn starch derivative in electrolyte solution, also for use mainly in pelvic surgery). However, use of these substances directly over bowel anastomoses is contraindicated due to an elevated risk of leak. There have been innumerable studies investigating different molecules in hopes of preventing adhesion formation, but most of the success is limited to animal models, and clinically significant results in humans have yet to be achieved.

TREATMENT OF WOUNDS

Local Care (Fig. 9-13)

Management of acute wounds begins with obtaining a careful history of the events surrounding the injury. The history is followed by a meticulous examination of the wound. Examination should assess the depth and configuration of the wound, the extent of nonviable tissue, and the presence of foreign bodies and other contaminants. The wound/ulcer should be described based on location, dimensions, presence of infection, drainage and type of drainage, base characteristics, presence or absence of necrosis, presence of pain, and description of edges. Possible etiology should be mentioned, and the presence of systemic factors and circulation should be evaluated.

After completion of the history, examination, and administration of tetanus prophylaxis, the wound should be meticulously anesthesitized. Lidocaine (0.5%–1%) or bupivacaine (0.25%–0.5%) combined with a 1:100,000 to 1:200,000 dilution of epinephrine provides satisfactory anesthesia and hemostasis.
Epinephrine should not be used in wounds of the fingers, toes, ears, nose, or penis, due to the risk of tissue necrosis secondary to terminal arteriole vasospasm in these structures.

Irrigation to visualize all areas of the wound and remove foreign material is best accomplished with normal saline (without additives). High-pressure wound irrigation is more effective in achieving complete debridement of foreign material and nonviable tissues. Iodine, povidone-iodine, hydrogen peroxide, and organically based antibacterial preparations have all been shown to impair wound healing due to injury to wound neutrophils and macrophages, and thus should not be used. All hematomas present within wounds should be carefully evacuated and any remaining bleeding sources controlled with ligature or cautery. If the injury has resulted in the formation of a marginally viable flap of skin or tissue, this should be resected or revascularized prior to further wound repair and closure.

After the wound has been anesthetized, explored, irrigated, and debrided, the area surrounding the wound should be cleaned, inspected, and the surrounding hair clipped. The area surrounding the wound should be prepared with povidone iodine, chlorhexidine, or similar bacteriostatic solutions and draped with sterile towels. Having ensured hemostasis and adequate debridement of nonviable tissues and removal of any remaining foreign bodies, irregular, macerated, or beveled wound edges should be debrided in order to provide a fresh edge for reapproximation. Although plastic surgical techniques such as W- or Z-plasty are seldom recommended for acute wounds, great care must be taken to realign wound edges properly. This is particularly important for wounds that cross the vermilion border, eyebrow, or hairline. Initial sutures that realign the edges of these different tissue types will speed and greatly enhance the aesthetic outcome of the wound repair.

In general, the smallest suture required to hold the various layers of the wound in approximation should be selected in order to minimize suture-related inflammation. Nonabsorbable or slowly absorbing monofilament sutures are most suitable for approximating deep fascial layers, particularly in the abdominal wall. Subcutaneous tissues should be closed with braided absorbable sutures, with care to avoid placement of sutures in fat. Multiple layer closure of the abdominal wall is preferable. However, additional layers with sutures might increase the risk of wound infection.

In areas of significant tissue loss, rotation of adjacent musculocutaneous flaps or free flaps may be required to provide sufficient tissue mass for closure. In cases of significant superficial tissue loss, split-thickness skin grafting (placed in a delayed manner to assure an adequate tissue bed) may be required and will speed formation of an intact epithelial barrier to fluid loss and infection. In acute, contaminated wounds with skin loss, use of porcine skin xenografts or skin cadaveric allografts might be needed to avoid infection.

After closing deep tissues and replacing significant tissue deficits, skin edges should be reapproximated for cosmesis and to aid in rapid wound healing. Skin edges may be quickly reapproximated with stainless steel staples or nonabsorbable monofilament sutures. Care must be taken to remove these from the wound prior to epithelialization of the skin tracts where sutures or staples penetrate the dermal layer. Failure to remove the sutures or staples prior to 7 to 10 days after repair will result in a cosmetically inferior wound. Where wound cosmesis is important, the aforementioned problems may be avoided by placement of buried dermal sutures using absorbable braided sutures. This method of wound closure allows for a precise reapproximation of wound edges and may be enhanced by application of wound closure tapes to the surface of the wound. Intradermal absorbable sutures do not require removal.

The development of octyl-cyanoacrylate tissue glue has shown good results for the management of simple, linear wounds with viable skin edges. These new glues are less prone to brittleness and have superior burst-strength characteristics. Studies have shown them to be suitable for use in contaminated situations without significant risk of infection. When used in the previously mentioned types of wounds, these glues appear to provide superb cosmetic results and result in significantly less trauma than sutured repair, particularly when used in pediatric patients.

**Antibiotics**

Antibiotics should be used only when there is an obvious wound infection. Most wounds are contaminated or colonized with bacteria. The presence of a host response constitutes an infection and justifies the use of antibiotics. Signs of infection to look for include erythema, cellulitis, swelling, and purulent discharge. Indiscriminate use of antibiotics should be avoided to prevent emergence of multidrug-resistant bacteria.

Antibiotic treatment of acute wounds must be based on organisms suspected to be found within the infected wound and the patient’s overall immune status. When a single specific organism is suspected, treatment may be commenced using a single antibiotic. Conversely, when multiple organisms are suspected, as with enteric contamination or when a patient’s immune function is impaired by diabetes, chronic disease, or medication, treatment should commence with a broad-spectrum antibiotic or several agents in combination. Antibiotics also can be delivered topically as part of irrigations or dressings, although their efficacy is questionable.

**Dressings**

The main purpose of wound dressings is to provide the ideal environment for wound healing. The dressing should facilitate the major changes taking place during healing to produce an optimally healed wound. Although the ideal dressing still is not a clinical reality, technological advances are promising (Table 9-9).

Dressings should take into account the type of wound and the comorbid conditions and associated factors such as edema and circulation. In patients with edema, dressing should compress the edematous area to aid in healing. In patients with significant circulation compromise, a compressing dressing should be avoided. The dressing should maintain

<table>
<thead>
<tr>
<th>Table 9-9 Desired characteristics of wound dressings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Promote wound healing (maintain moist environment)</td>
</tr>
<tr>
<td>Conformability</td>
</tr>
<tr>
<td>Pain control</td>
</tr>
<tr>
<td>Odor control</td>
</tr>
<tr>
<td>Nonallergenic and nonirritating</td>
</tr>
<tr>
<td>Permeability to gas</td>
</tr>
<tr>
<td>Safety</td>
</tr>
<tr>
<td>Nontraumatic removal</td>
</tr>
<tr>
<td>Cost-effectiveness</td>
</tr>
<tr>
<td>Convenience</td>
</tr>
</tbody>
</table>
an adequate moist environment but should help absorption excessive drainage. Occlusion of a wound with dressing material helps healing by controlling the level of hydration and oxygen tension within the wound. It also allows transfer of gases and water vapor from the wound surface to the atmosphere. Occlusion affects both the dermis and epidermis, and it has been shown that exposed wounds are more inflamed and develop more necrosis than covered wounds. Occlusion also helps in dermal collagen synthesis and epithelial cell migration and limits tissue desiccation. Since it may enhance bacterial growth, occlusion is contraindicated in infected and/or highly exudative wounds.

Dressings can be classified as primary or secondary. A primary dressing is placed directly on the wound and may provide absorption of fluids and prevent desiccation, infection, and adhesion of a secondary dressing. A secondary dressing is one that is placed on the primary dressing for further protection, absorption, compression, and occlusion. Although the ideal dressing does not exist, many types of dressings help achieve certain goals, so knowledge of the wound and the dressing function is essential to make it possible to choose the appropriate dressing.

Absorbent Dressings. This type of dressing helps control exudate without soaking through the dressing, which can increase infection potential.

Nonadherent Dressings. Nonadherent dressings are impregnated with paraffin, petroleum jelly, or water-soluble jelly for use as nonadherent coverage. A secondary dressing must be placed on top to seal the edges and prevent desiccation and infection.

Occlusive and Semiocclusive Dressings. Occlusive and semiocclusive dressings provide a good environment for clean, minimally exudative wounds. These film dressings are waterproof and impervious to microbes but permeable to water vapor and oxygen.

Hydrophilic and Hydrophobic Dressings. These dressings are components of a composite dressing. Hydrophilic dressing aids in absorption, whereas a hydrophobic dressing is waterproof and prevents absorption.

Hydrocolloid and Hydrogel Dressings. Hydrocolloid and hydrogel dressings attempt to combine the benefits of occlusion and absorbency. Hydrocolloids and hydrogels form complex structures with water, and fluid absorption occurs with particle swelling, which aids in atraumatic removal of the dressing. Absorption of exudates by the hydrocolloid dressing leaves a yellowish-brown gelatinous mass after dressing removal that can be washed off. Hydrogel is a cross-linked polymer that has high water content. Hydrogels allow a high rate of evaporation without compromising wound hydration, which makes them useful in burn wound treatment.

Alginites. Alginites are derived from brown algae and contain long chains of polysaccharides containing mannuronic and glucuronic acid. The ratios of these sugars vary with the species of algae used, as well as the season of harvest. Processed as the calcium forms, alginites turn into soluble sodium alginate through ion exchange in the presence of wound exudates. The polymers gel, swell, and absorb a great deal of fluid. Alginites are being used when there is skin loss, in open surgical wounds with medium exudation, and on full-thickness chronic wounds. Alginate widely used as primary dressing and can be reinforced with other forms of dressing such as compression dressing.

Absorbable Materials. Absorbable materials are mainly used within wounds as hemostats and include collagen, gelatin, oxidized cellulose, and oxidized regenerated cellulose.

Medicated Dressings. Medicated dressings have long been used as a drug-delivery system. Agents delivered in the dressings include benzylic peroxide, zinc oxide, neomycin, and bacitracin-zinc. These agents have been shown to increase epithelialization by 28%.

The type of dressing to be used depends on the amount of wound drainage. A nondraining wound can be covered with semiocclusive dressing. Drainage of less than 1 to 2 mL/d may require a semiocclusive or absorbent nonadherent dressing. Moderately draining wounds (3–5 mL/d) can be dressed with a nonadherent primary layer plus an absorbent secondary layer plus an occlusive dressing to protect normal tissue. Heavily draining wounds (>5 mL/d) require a similar dressing as moderately draining wounds, but with the addition of a highly absorbent secondary layer.

Mechanical Devices. Mechanical therapy augments and improves on certain functions of dressings, in particular the absorption of exudates and control of odor. The negative pressure dressing systems assist in wound closure by applying localized negative pressure to the surface and margins of the wound. The negative-pressure therapy is applied to a special foam dressing cut to the dimensions of the wound and positioned in the wound cavity or over a flap or graft. The continuous negative pressure is very effective in removing exudates from the wound. This form of therapy has been found to be effective for chronic open wounds (diabetic ulcers and stages III and IV pressure ulcers), acute and traumatic wounds,125 flaps and grafts, and subacutent wounds (i.e., dehisced incisions), although more randomized trials need to be carried out to confirm efficacy. Different types of sponges are available to use on wounds with negative pressure systems.

Skin Replacements

All wounds require coverage in order to prevent evaporative losses and infection and to provide an environment that promotes healing. Both acute and chronic wounds may demand use of skin replacement, and several options are available.

Conventional Skin Grafts. Skin grafts have long been used to treat both acute and chronic wounds. Split-(partial)-thickness grafts consist of the epidermis plus part of the dermis, whereas full-thickness grafts retain the entire epidermis and dermis. Autologous grafts (autografts) are transplants from one site on the body to another; allogeneic grafts (allografts, homografts) are transplants from a living nonidentical donor or cadaver to the host; and xenogeneic grafts (heterografts) are taken from another species (e.g., porcine). Split-thickness grafts require less blood supply to restore skin function. The dermal component of full-thickness grafts lends mechanical strength and resists wound contraction better, resulting in improved cosmesis. Allogeneic and xenogeneic grafts require the availability of tissue, are subject to rejection, and may contain pathogens.

The use of skin grafts or bioengineered skin substitutes and other innovative treatments (e.g., topically applied growth factors, systemic agents, and gene therapy) cannot be effective unless the wound bed is adequately prepared. This may include debridement to remove necrotic or fibrinous tissue, control of edema, revascularization of the wound bed, reduction in the bacterial burden, with minimal exudate. Temporary
Cellular and Tissue-Based Products in Chronic Wound and Ulcer Management

Wound management and ulcer healing are among the most challenging problems in medical practices. The lack of large clinical trials, the heterogeneity of wound causes, and mechanisms of chronicity add to the complexity of wound management. In most cases, management is based on the experience of the providers and the availability of the treatment modalities within the health facilities. In spite of advances in wound care, the basic principles of wound management remain to be essential to healing. Additional measures and products might accelerate the rate of healing but will not replace basic wound care. The basic principles include achieving optimal blood flow, control of infection, removal of debris, proper dressing, offloading the injured area, and compression therapy in certain cases. Once these basic principles have been achieved, advanced treatment modalities such as cellular and tissue-based products (CTP) can be considered for enhanced healing. CTPs presumably act by altering the biology of wounds and ulcers or by preparing the wound/ulcer bed for healing and other potential procedures. CTPs are divided into two categories: dermoinductive and democonductive. As the name indicates, dermoinductive products help provide cells and factors that will activate healing within the wound by inducing tissue growth or inducing granulation within the wound. Such products include Apligraf (Organogenesis, Canton, MA), Theraskin (Soluble systems, LLC, Newport News, VA), and Dermagraft (Organogenesis Canton, MA). None of these products have achieved healing in all wounds, but generally they have improved the percentage of healing, or achieved healing over a shorter period of time. In a pivotal Apligraf study, for example, it was found that patients who received Apligraf achieved 56% healing over 65 days compared to 38% healing over 90 days in those who received saline dressing. On the other hand, the dermoconductive products provide scaffolding to a wound ending in a neodermis by allowing migration of surrounding tissues across the wound, and this helps healing. An example of such products is Integra, which is composed of type I bovine collagen, shark chondroitin 6-sulfate, and a silicone layer and helps to prepare the wound bed for subsequent skin grafting. In our experience, we use Integra for wound bed preparation in superficial wounds but also, in some cases, in deeper wounds. In addition, we also use it to cover exposed bone, especially smaller areas and tendons; however, it might require more than one application in some cases to achieve optimal wound bed preparation. In a systematic review of the literature, the authors reviewed 15 randomized trials, one prospective comparative study, and five systematic reviews. The authors concluded that living cell-based skin substitutes have a convincing supportive body of evidence in wound healing with some promise for acellular skin substitutes. In their review, the authors indicated that the evidence they based their review on was of low quality. Stem cell–based therapy is gaining more traction in the management of difficult wounds. Although it was initially used as an attempt to achieve scarless healing, stem cell therapy has gained more popularity in recent years as a means for enhanced healing and skin regeneration, in addition to reduction of scar formation. Similar to other fields, stem cells in wound healing produce growth factors and chemokines that can differentiate into different cells and tissues. Stem cells in wound healing are mostly derived from human amniotic membrane of placental tissue. Examples of such products include Epithex (MiMedx Group Inc, Marietta CA) and Grafix (Osirix Therapeutics Inc, Columbia, MD), among others. It is believed that the growth factors in the amniotic membranes are preserved through different processes, thus assisting with wound healing. Multiple growth factors derived from the amniotic membrane are thought to contribute to wound healing, including vascular endothelia growth factor, platelet-derived growth factor, epidermal growth factor, transforming growth factor, and others. A significant drawback of such products is that they are extremely expensive. Furthermore, it is recommended that these products should be used weekly to achieve optimal healing, which adds to the cost of using such products.

Originally, these products were devised to provide coverage of extensive wounds because of the limited availability of autografts, which remains the ideal tissue coverage. However, CTPs have now gained acceptance as natural dressings. Manufactured by tissue engineering, they combine novel materials with living cells to provide functional skin substitutes, providing a bridge between dressings and skin grafts.

Skin substitutes have the theoretical advantages of being readily available and not requiring painful harvests like in skin grafts, and they may be applied freely or with surgical suturing. In addition, they promote healing, either by stimulating host cytokine generation or by providing cells that may also produce growth factors locally. Their disadvantages include limited survival, high cost, and the need for multiple applications (Table 9-10). Allografting, albeit with a very thin graft, may at times be required to accomplish complete coverage.

A variety of skin substitutes are available, each with its own set of advantages and disadvantages; however, the ideal skin substitute has yet to be developed (Table 9-11). The development of the newer composite substitutes, which provide both the dermal and epidermal components essential for permanent skin replacement, may represent an advance toward that goal.

Cultured epithelial autografts (CEAs) represent expanded autologous or homologous keratinocytes. CEAs are expanded from a biopsy of the patient’s own skin, will not be rejected, and can stimulate reepithelialization as well as the growth of underlying connective tissue. Keratinocytes harvested from a biopsy roughly the size of a postage stamp are cultured with fibroblasts and growth factors and grown into sheets that can cover large areas and give the appearance of normal skin. Until the epithelial sheets are sufficiently expanded, the wound must be covered with an occlusive dressing or a temporary allograft or xenograft. The dermis regenerates very slowly, if at all, for full-thickness wounds because the sheets are very fragile, are difficult to work with, are susceptible to infection, and do not resist contracture well, leading to poor cosmetic results.

<table>
<thead>
<tr>
<th>Desired features of tissue-engineered skin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rapid reestablishment of functional skin (epidermis/dermis)</td>
</tr>
<tr>
<td>Receptive to body’s own cells (e.g., rapid “take” and integration)</td>
</tr>
<tr>
<td>Graftable by a single, simple procedure</td>
</tr>
<tr>
<td>Graftable on chronic or acute wounds</td>
</tr>
<tr>
<td>Engraftment without use of extraordinary clinical intervention (i.e., immunosuppression)</td>
</tr>
</tbody>
</table>
CEAs are available from cadavers, unrelated adult donors, or neonatal foreskins. Fresh or cryopreserved cultured allogeneic keratinocytes can be left in place long enough to be superseded by multiplying endogenous skin cells because, unlike allografts containing epidermal Langerhans cells, they do not express major histocompatibility antigens. Cryopreserved CEAs are readily available “off the shelf,” and provide growth factors that may aid healing. However, like autologous keratinocyte sheets, the grafts lack the strength provided by a dermal component and pose a risk of disease transmission.

Viable fibroblasts can be grown on bioabsorbable or nonbioabsorbable meshes to yield living dermal tissue that can act as a scaffold for epidermal growth. Fibroblasts stimulated by growth factors can produce type I collagen and glycosaminoglycans (e.g., chondroitin sulfates), which adhere to the wound surface to permit epithelial cell migration, as well as adhesive ligands (e.g., the matrix protein fibronectin), which promote cell adhesion. This approach has the virtue of being less time-consuming and expensive than culturing keratinocyte sheets. There are a number of commercially available, bioengineered dermal replacements approved for use in burn wound treatment as well as other indications.

Bioengineered skin substitutes have evolved from keratinocyte monolayers to dermal equivalents to split-thickness products with a pseudo-epidermis, and most recently, to products containing both epidermal and dermal components that resemble the three-dimensional structure and function of normal skin (see Table 9-11). Indicated for use with standard compression therapy in the treatment of venous insufficiency ulcers and for the treatment of neuropathic diabetic foot ulcers, these bilayered skin equivalents also are being used in a variety of wound care settings.

Growth Factor Therapy. As discussed previously, it is believed that nonhealing wounds result from insufficient or inadequate growth factors in the wound environment. A simplistic solution would be to flood the wound with single or multiple growth factors in order to “jump-start” healing and reepithelialization. Although there is a large body of work demonstrating the effects of growth factors in animals, translation of these data into clinical practice has met with limited success. Growth factors for clinical use may be either recombinant or homologous/autoologous. Autologous growth factors are harvested from the patient’s own platelets, yielding an unpredictable combination and concentration of factors, which are then applied to the wound. This approach allows treatment with patient-specific factors at an apparently physiologic ratio of growth factor concentrations. Disappointingly, a recent meta-analysis failed to demonstrate any value for autologous platelet-rich plasma in the treatment of chronic wounds. Recombinant molecular biologic means permit the purification of high concentrations of individual growth factors. Current FDA-approved formulations, as well as those used experimentally, deliver concentrations approximately 10^3 times higher than those observed physiologically.

At present, only platelet-derived growth factor BB (PDGF-BB) is currently approved by the FDA for treatment of diabetic foot ulcers. Application of recombinant human PDGF-BB in a gel suspension to these wounds increases the incidence of total healing and decreases healing time. Several other growth factors have been tested clinically and show some promise, but currently none are approved for use. A great deal more needs to be discovered about the concentration, temporal release, and receptor cell population before growth factor therapy is to make a consistent impact on wound healing.

Gene or Cell Therapy. Given the disappointing results from the application of purified growth factors onto wounds, the possible therapeutic potential of gene therapy has been recognized and studied. Direct access to the open wound bed, which characterizes almost all chronic wounds, has facilitated this therapy. Gene delivery to wounds includes traditional approaches such as viral vectors and plasmid delivery or, more recently, electroporation and microseeding.

Although a variety of genes expressing interleukin-8, PDGF, IGF-1, keratinocyte growth factor, and laminin-5 have been successfully delivered to wounds in both animal and human models, the effects have been modest and specific to unique wound situations. Delivering extra genes into the wound bed presents the challenge of expression of the necessary signals to turn the genes on and off at appropriate times so that dysregulated, hypertrophic, and abnormal healing does not occur. Elaborate systems have been created for topical use as on/off switches for genes. The more important question is which genes to express, in what temporal sequence, and in what regions of the wound bed, as it is unlikely that a single gene coding for one protein can significantly affect overall healing. There is growing consensus that delivery of genes is not going to represent the universal solution. Although gene therapy replaces missing or
defective genes, most acute wounds already have and express the necessary genes for successful healing and the wound environment produces signals adequate to the activation of these genes. What, if any, are the deficiencies in gene expression or activity in failed wounds is unknown.

Another approach is to deliver multiple genes coding for proteins that can act synergistically and even in a timed sequence, as would occur during normal healing. This would involve the use of activated cells that participate in the healing sequence that could be delivered in an activated state to the wound environment. Use of mesenchymal stem cells as a delivery vector for many genes simultaneously is the latest such approach. The feasibility of applying bone marrow-derived, umbilical cord-derived, adipose-derived, and epidermal stem cells that can differentiate into various cells that participate in the wound healing response also has been documented. These cells, as part of their differentiation and activation in the wound, have been shown to produce a variety of growth factors including VEGF, PDGF, bFGF, and MMP-9. The challenges remain how to maintain the viability and activity of the transplanted cells, how to document that the observed effects are due to the delivered cells, and what are the mechanisms necessary for regulating or ending their activity.

Oxygen Therapy in Wound Healing

Oxygen is required for almost all steps of wound healing and is also an important factor in the body’s defense against bacterial infection. In addition to its role in healing, oxygen plays an essential role in the production of reactive oxygen species such as superoxide that are angiogenesis stimulators and are bacteriostatic. Chronic wounds have a decreased oxygen supply, and for a long time lack of oxygen was recognized as a potential cause of delayed healing. To counteract this delay, supplemental oxygen therapy was used to improve healing, and both local oxygen therapy and systemic therapy were used for that purpose. Local therapy included oxygen dressings and topical oxygen therapy, while systemic therapy included supplemental inspired oxygen therapy and pressurized oxygen treatment.

Of the different methods of oxygen therapy, pressurized oxygen, also termed hyperbaric oxygen therapy (HBOT) is the most used modality and the most investigated. In HBOT, oxygen is delivered under pressure, more than atmospheric pressure, leading to a higher concentration of oxygen in tissues. The Undersea and Hyperbaric Medical Society (UHMS) defined HBOT as an intervention that involves breathing near 100% oxygen intermittently under high pressure achieved by a pressurized chamber to more than sea level pressure (sea level pressure = 1 atmosphere absolute [ATA]). The therapeutic pressure should be at least 1.4 ATA. Chambers can be single occupancy or multiple occupancy. Although there are numerous indications and potential indications for HBOT, there are 14 accepted indications by Undersea and Hyperbaric Medical Society, and the FDA. Indications related to wounds and ulcers include clostridial myonecrosis, crush injury, radiation-induced soft tissue and bone necrosis, necrotizing soft tissue infections, diabetic ulcers Wagner III or higher, refractory osteomyelitis, and thermal burns. Two systematic reviews were published on the effect of HBOT on wound healing. The first one was published in 2003 on all studies done up to 2001. Generally, the review showed beneficial effects of HBOT on different disease processes. HBOT was found to improve graft survival, complete healing of grafts, and lessen infection in patients with a graft. Patients with osteoradionecrosis showed improved bone changes. The effect on tissue radionecrosis was described as positive, but not all studies showed statistical significance. HBOT was found to have improved amputation rates in patients with gas gangrene with better healing. In addition, HBOT was found to reduce mortality rates from necrotizing fasciitis, and significantly decreased wound size in nonhealing diabetic wounds. In the second systematic review, a total of 29 studies published between 2001 and 2016 were included. A total of 14 studies were related to chronic wounds: 12 studies in acute wounds, 1 study on both acute and chronic wounds, and 2 experimental studies. Out of those 29 studies, 18 (62%) showed at least one positive outcome. One of the drawbacks of such studies in general is the lack of heterogeneity and the lack of accurate randomization. Perhaps, HBOT in diabetic foot ulcers attracted more attention than other forms of skin ulcers. In patients with diabetic foot ulcers, studies showed that HBOT led to significantly more healing wounds at 1 year, greater reduction in the wound bed, greater healing with less proximal amputation, and more complete wound healing at 1 month after HBOT.

Biofilm and Chronic Wound Healing

Traditionally, nonhealing of chronic wounds has been associated with numerous risk factors, including longer duration of ulcers, advanced age, increased body mass index, venous reflux, arterial and venous disease, nutritional deficiencies, diabetes mellitus, and smoking. Chronic bacterial infection is another factor that has been associated with nonhealing wounds.

Chronic wounds, in general, behave differently in relation to bacterial growth when compared to acute wounds. Chronic wounds develop bacterial growth that is resistant to invasion by antibiotics and is protected by the host immune defenses. Biofilm is the term used for the bacterial growth on a chronic wound that is encapsulated by a protective layer made up of the host and bacterial proteins. Bijnsholt et al have suggested a simplified definition of the biofilm as “an aggregate of bacteria tolerant to treatment and the host defense.” It has been found that more than 60% of chronic wounds have a biofilm. Biofilms lead to a chronic inflammatory process that will interfere with healing. Biofilms are formed in the wound environment, which provides a micro-environment that is ideal for bacterial growth. Once a biofilm colony forms, it will continuously shed bacteria to uncolonized areas, causing more biofilm colonies to form.

The presence of biofilms protects bacteria from the host defenses. Bacteria will then begin to exhibit phenotypic and genotypic plurality: the former allows bacteria to adapt to different growth conditions such as nutrient availability, pH, and oxidizing potential within the biofilm, while the latter allows for virulence and bacterial resistance to drugs. The genetic plurality is passed horizontally among bacteria in the wound, adding to the resistance to treatment and allowing the bacteria to avoid the defense mechanisms of the host such as bacterial phagocytosis,
neutrophil degranulation, and formation of reactive oxygen species. In some cases, the biofilm allows bacteria to become quiescent and thus become less sensitive to antimicrobials that typically affect dividing bacterial cells.

Biofilms can form and recover from debridement rapidly. Fully mature biofilm colonies can form within 2 to 4 days depending on the species and growth conditions. Common bacterial species such as *Staphylococcus*, *Streptococcus*, and *Pseudomonas* can attach to the surface of the wound in minutes, forming adherent microcolonies in 2 to 4 hours. Bacterial species develops the extracellular polymeric substances (EPS) and resistance to disinfectants, antiseptics, and antibiotics in 6 to 12 hours. The biofilm then develops into mature colonies in 2 to 4 days. Furthermore, it rapidly recovers from mechanical disruption, such as debridement, within 24 hours. The presence of biofilms leads to delayed wound healing by stimulating chronic inflammation in the wound. The host responds to the biofilms by mobilizing macrophages and neutrophils in the biofilm area, resulting in secretion of high levels of reactive oxygen species and proteases that can cause damage to the normal and healing tissue. The resulting increase in exudate production provides a source of nutrients to the biofilms, leading to more resistance to healing.

The best method of treating wounds with biofilms is not well defined. It is believed that frequent debridement to mechanically remove the biofilm remains to be the best method of treatment. The frequency of debridement is not well defined, although a weekly debridement has been suggested to improve healing. More than one method of debridement might be needed in wounds resistant to healing. A few products have been suggested to control and remove biofilms, such as surfactant and some wound cleansing formulas. In addition, topical broad-spectrum antimicrobials such as silver, iodine, honey, and others have been suggested to aid in the treatment of biofilms. Using a combination of debridement methods, cleansing agents, and antimicrobials may be needed in resistant chronic wounds. An effective therapy to combat biofilm should include effective removal of the biofilm, antimicrobial eradication of bacteria/organisms, and prevention of biofilm recurrence, and these treatment options may need to be repeated multiple times to achieve effective therapy. Many agents and devices thought to be effective in eradicating and treating biofilms are under evaluation or development; however, their long-term effect is yet to be proven.

**REFERENCES**

Entries highlighted in bright blue are key references.


As the population ages, oncology is becoming a larger portion of surgical practice. The surgeon often is responsible for the initial diagnosis and management of solid tumors. Knowledge of cancer epidemiology, etiology, staging, and natural history is required for initial patient assessment, as well as to determine the optimal surgical therapy.

Modern cancer therapy is multidisciplinary, involving the coordinated care of patients by surgeons, medical oncologists, radiation oncologists, reconstructive surgeons, pathologists, radiologists, and primary care physicians. **Primary (or definitive) surgical therapy** refers to en bloc resection of tumor with adequate margins of normal tissues and regional lymph nodes as necessary. **Adjuvant therapy** refers to radiation therapy and systemic therapies, including chemotherapy, immunotherapy, hormonal therapy, and, increasingly, biologic therapy. The primary goal of surgical and radiation therapy is local and regional control. On the other hand, the primary goal of adjuvant therapy is systemic control by treatment of distant foci of subclinical disease to prevent distant recurrence. Surgeons must be familiar with adjuvant therapies to coordinate multidisciplinary care and to determine the best sequence of therapy. They must also be aware of the potential for patients to receive effective systemic therapies prior to surgery as a means of reducing tumor volume.

Recent advances in molecular biology are revolutionizing medicine. New information is being translated rapidly into clinical use, with the development of new prognostic and predictive markers and new biologic therapies. Increasingly cancer therapy is becoming personalized, incorporating information about each patient’s tumor characteristics, patient’s own genome, as well as host immune responses and tumor microenvironment, into clinical decision-making. It is therefore essential that surgeons understand the principles of molecular oncology to appropriately interpret these new contributions and incorporate them into practice.
Key Points

1. Modern cancer therapy is multidisciplinary, involving coordinated care by surgeons, medical oncologists, radiation oncologists, reconstructive surgeons, pathologists, radiologists, and primary care physicians.

2. Understanding cancer biology is essential to successfully implement personalized cancer therapy.

3. The following alterations are critical for malignant cancer growth: self-sufficiency of growth signals, insensitivity to growth-inhibitory signals, evasion of apoptosis, potential for limitless replication, angiogenesis, invasion and metastasis, reprogramming of energy metabolism, and evading immune destruction.

Epidemiology

Basic Principles of Cancer Epidemiology

The term incidence refers to the number of new cases occurring. Incidence is usually expressed as the number of new cases per 100,000 persons per year. Mortality refers to the number of deaths occurring and is expressed as the number of deaths per 100,000 persons per year. Incidence and mortality data are usually available through cancer registries. Mortality data are also available as public records in many countries where deaths are registered as vital statistics, often with the cause of death. In areas where cancer registries do not exist, mortality data are used to extrapolate incidence rates. These numbers are likely to be less accurate than registry data, as the relationship between incidence and cause-specific death is likely to vary significantly among countries owing to the variation in health care delivery.

The incidence of cancer varies by geography. This is due in part to genetic differences and in part to differences in environmental and dietary exposures. Epidemiologic studies that monitor trends in cancer incidence and mortality have tremendously enhanced our understanding of the etiology of cancer. Furthermore, analysis of trends in cancer incidence and mortality allows us to monitor the effects of different preventive and screening measures, as well as the evolution of therapies for specific cancers.

The two types of epidemiologic studies that are conducted most often to investigate the etiology of cancer and the effect of prevention modalities are cohort studies and case-control studies. Cohort studies follow a group of people who initially do not have a disease over time and measure the rate of development of a disease. In cohort studies, a group that is exposed to a certain environmental factor or intervention usually is compared to a group that has not been exposed (e.g., smokers vs. nonsmokers). A case-control study compares a group of patients affected with a disease to a group of individuals without the disease and looks back retrospectively to compare how frequently the exposure to a risk factor is present in each group to determine the relationship between the risk factor and the disease. The results are expressed in terms of an odds ratio, or relative risk. A relative risk <1 indicates a protective effect of the exposure, whereas a relative risk >1 indicates an increased risk of developing the disease with exposure.

Cancer Incidence and Mortality in the United States

In the year 2017, it is estimated that 1.7 million new cancer cases will be diagnosed in the United States, excluding carcinoma in situ of any site except bladder, and excluding basal cell and squamous cell carcinomas of the skin. In addition, 63,410 cases of carcinoma in situ of the breast, and 74,680 of melanoma in situ are expected. It is estimated that in 2017 an estimated 600,920 people will die of cancer in the United States, corresponding to about 1650 deaths per day. The estimated new cancer cases and deaths by cancer type are shown in Table 10-1. The most common causes of cancer death in men are cancers of the lung and bronchus, colon and rectum, and prostate; in women, the most common cancers are of the lung and bronchus, breast, and colon and rectum. These four cancers account for almost half of total cancer deaths among men and women.

The annual age-adjusted cancer incidence rates among males and females for selected cancer types are shown in Fig. 10-1. Incidence rates are declining for most cancer sites (Fig. 10-2). Incidence rates for thyroid cancer have begun to stabilize recently, possibly due to changes in clinical practice guidelines that were initiated in 2009 and included more conservative indications for biopsy. The age-adjusted incidence rate of breast cancer started to decrease from 2001 to 2004. This decrease in breast cancer incidence has at least temporally been associated with the first report of the Women’s Health Initiative, which documented an increased risk of coronary artery disease and breast cancer with the use of hormone replacement therapy; this was followed by a drop in the use of hormone replacement therapy by postmenopausal women in the United States. Unfortunately, there was a slight increase in breast cancer incidence from 2004 to 2013. This was driven wholly by nonwhite women; rates increased by about 2% per year among women other than white or black and by 0.5% per year among black women. Thus, rates have risen slightly in women as a whole from 2013 to 2017, although rates remained stable in white women.

Declines in colorectal cancer incidence have been mainly attributed to increased screening that allows for removal of precancerous polyps. Prostate cancer rates rapidly increased and decreased between 1995 and 1998. These trends are thought to be attributable to increased use of prostate-specific antigen (PSA) screening. Due to growing concerns about overdiagnosis and overtreatment, a U.S. Preventive Services Task Force recommended against routine use of PSA testing to screen for prostate cancer. As a result, there was more than 10% annual reduction in prostate cancer incidence from 2010 to 2013. Differences in lung cancer incidence patterns between women and men are thought to reflect historical differences in tobacco use. Differences in smoking prevalence is also thought to contribute to regional differences in lung cancer incidence. Lung cancer incidence is fourfold higher in Kentucky, which has the highest smoking prevalence, compared with Utah, which has the lowest smoking prevalence.

The 5-year survival rates for selected cancers are listed in Table 10-2. From 2010 to 2014, cancer death rates decreased by 1.8% per year in men and by 1.4% per year in women. These declines in mortality have been consistent in the past decade and larger than what was observed in the previous decade. Over the
Table 10-1

<table>
<thead>
<tr>
<th>Category</th>
<th>Estimated New Cases</th>
<th>Estimated Deaths</th>
<th>Category</th>
<th>Estimated New Cases</th>
<th>Estimated Deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cancers</td>
<td>1,688,780</td>
<td>600,920</td>
<td>Genital system</td>
<td>279,800</td>
<td>59,100</td>
</tr>
<tr>
<td>Oral cavity and pharynx</td>
<td>49,670</td>
<td>9700</td>
<td>Uterine cervix</td>
<td>12,820</td>
<td>4210</td>
</tr>
<tr>
<td>Digestive system</td>
<td>310,440</td>
<td>157,700</td>
<td>Uterine corpus</td>
<td>61,380</td>
<td>10,920</td>
</tr>
<tr>
<td>Esophagus</td>
<td>16,940</td>
<td>15,690</td>
<td>Ovary</td>
<td>22,440</td>
<td>14,080</td>
</tr>
<tr>
<td>Stomach</td>
<td>28,000</td>
<td>10,960</td>
<td>Vulva</td>
<td>6020</td>
<td>1150</td>
</tr>
<tr>
<td>Small intestine</td>
<td>10,190</td>
<td>1390</td>
<td>Vagina and other genital, female</td>
<td>4810</td>
<td>1240</td>
</tr>
<tr>
<td>Colon and rectum</td>
<td>95,520 / 39,910</td>
<td>50,260</td>
<td>Prostate</td>
<td>161,360</td>
<td>26,730</td>
</tr>
<tr>
<td>Anus, anal canal, and anorectum</td>
<td>8200</td>
<td>1100</td>
<td>Testis</td>
<td>8850</td>
<td>410</td>
</tr>
<tr>
<td>Liver and intrahepatic bile duct</td>
<td>40,710</td>
<td>28,920</td>
<td>Penis and other genital, male</td>
<td>2120</td>
<td>360</td>
</tr>
<tr>
<td>Gallbladder and other biliary</td>
<td>11,740</td>
<td>3830</td>
<td>Urinary system</td>
<td>146,650</td>
<td>32,190</td>
</tr>
<tr>
<td>Pancreas</td>
<td>53,670</td>
<td>43,090</td>
<td>Urinary bladder</td>
<td>79,030</td>
<td>16,870</td>
</tr>
<tr>
<td>Other digestive organs</td>
<td>5560</td>
<td>2460</td>
<td>Kidney and renal pelvis</td>
<td>63,990</td>
<td>14,400</td>
</tr>
<tr>
<td>Respiratory system</td>
<td>243,170</td>
<td>160,420</td>
<td>Ureter and other urinary organs</td>
<td>3630</td>
<td>920</td>
</tr>
<tr>
<td>Larynx</td>
<td>13,360</td>
<td>3660</td>
<td>Eye and orbit</td>
<td>3130</td>
<td>330</td>
</tr>
<tr>
<td>Lung and bronchus</td>
<td>220,500</td>
<td>155,870</td>
<td>Brain and other nervous system</td>
<td>23,800</td>
<td>16,700</td>
</tr>
<tr>
<td>Other respiratory organs</td>
<td>7310</td>
<td>890</td>
<td>Endocrine system</td>
<td>59,250</td>
<td>3010</td>
</tr>
<tr>
<td>Bones and joints</td>
<td>3260</td>
<td>1550</td>
<td>Thyroid</td>
<td>56,870</td>
<td>2010</td>
</tr>
<tr>
<td>Soft tissue (including heart)</td>
<td>12,390</td>
<td>4990</td>
<td>Other endocrine</td>
<td>2380</td>
<td>1000</td>
</tr>
<tr>
<td>Skin (excluding basal and squamous)</td>
<td>95,360</td>
<td>13,590</td>
<td>Lymphoma</td>
<td>80,500</td>
<td>21,210</td>
</tr>
<tr>
<td>Melanoma</td>
<td>87,110</td>
<td>9730</td>
<td>Multiple myeloma</td>
<td>30,280</td>
<td>12,590</td>
</tr>
<tr>
<td>Other nonepithelial</td>
<td>8250</td>
<td>3860</td>
<td>Leukemia</td>
<td>62,130</td>
<td>24,500</td>
</tr>
<tr>
<td>Breast</td>
<td>255,180</td>
<td>41,070</td>
<td>Other and unspecified primary sites*</td>
<td>33,770</td>
<td>42,270</td>
</tr>
</tbody>
</table>

*Rounded to the nearest 10, cases exclude basal cell and squamous cell skin cancers and in situ carcinoma except bladder. About 63,410 cases of carcinoma in situ of the female breast and 74,680 cases of melanoma in situ were diagnosed in 2017.

Deaths for colon and rectum cancers are combined because a large number of deaths from rectal cancer are misclassified as colon.

More deaths than cases may reflect lack of specificity in recording underlying cause of death on death certificate and/or an undercount in the case estimate.


past 3 decades, the 5-year relative survival rate for all cancers combined has increased by 20% among whites and 24% among blacks. Improvements in survival for the most common cancers have been similar across both sexes but are more pronounced among patients age 50 to 64 years than among those older than 65 years. This difference may reflect reduced efficacy of new therapies in the elderly or perhaps lower utilization. Progress has been rapid for hematopoietic and lymphoid malignancies due to improved treatment protocols and the discovery of targeted therapies. The decrease in lung cancer death rates in men is thought to be due to a decrease in tobacco use, whereas the decreases in death rates from breast, colorectal cancer, and prostate cancer likely reflect advances in early detection and treatment. For instance, there is potential for lung cancer to be diagnosed at an earlier stage through the use of screening with low-dose computed tomography.

Global Statistics on Cancer Incidence and Mortality

The five most common cancers for men worldwide are lung, prostate, colorectal cancer, stomach, and liver and for women are breast, colorectal, cervix, lung, and stomach. Notably, for several cancer types there is wide geographical variability in cancer incidence (Fig. 10-3). The mortality rates for different cancers also vary significantly among countries. This is attributable not only to variations in incidence but also to variations in survival after a cancer diagnosis. The survival rates are influenced by treatment patterns as well as by variations in cancer screening practices, which affect the stage of cancer at diagnosis. For example, the 5-year survival rate for stomach cancer is much higher in Japan, where the cancer incidence is high enough to warrant mass screening, which is presumed to lead to earlier diagnosis. In the case of prostate cancer, on the other
hand, the mortality rates diverge much less than the incidence rates among countries. Survival rates for prostate cancer are much higher in North America than in developing countries. It is possible that the extensive screening practices in the United States allow discovery of cancers at an earlier, more curable stage; however, it is also possible that this screening leads to discovery of more latent, less biologically aggressive cancers, which may not have caused death even if they had not been identified.

In 2008 (the last date for which complete data are available), about 1 million new cases of stomach cancer were estimated to have occurred (988,000 cases, 7.8% of the total), making it the fourth most common malignancy in the world, behind cancers of the lung, breast, and colorectal cancer. The incidence of stomach cancer varies significantly among different regions of the world. The difference in risk by country is presumed to be primarily due to differences in dietary factors. The risk is increased by high consumption of preserved salted foods, such as meats and pickles, and decreased by high intake of fruits and vegetables. There is also some international variation in the incidence of infection with *Helicobacter pylori*, which is known to play a major role in gastric cancer development. Fortunately, a steady decline is being observed in the incidence and mortality rates of gastric cancer. This may be related to improvements in preservation and storage of foods as well as due to changes in the prevalence of *H pylori*. More than 70% of cases (713,000 cases) occur in developing countries, and half the cases in the world occur in Eastern Asia (mainly in China). Age-standardized incidence rates are about twice as high for men as for women, ranging from 3.9 in Northern Africa to 42.4 in Eastern Asia for men, and from 2.2 in Southern Africa to 18.3 in Eastern Asia for women. Stomach cancer is the second leading cause of cancer death in both sexes worldwide.

Overall, the incidence of breast cancer is rising in most countries. Incidence varies from 19.3 per 100,000 women in Eastern Africa to 89.7 per 100,000 women in Western Europe, and are high in developed regions of the world (except Japan) and low in most of the developing regions. Although breast cancer has been linked to cancer susceptibility genes, mutations in these genes account for only 5% to 10% of breast tumors, which suggests that the wide geographic variations in breast cancer incidence are not due to geographic variations in the prevalence of these genes. Most of the differences, therefore, are attributed to differences in reproductive factors, diet, alcohol, obesity, physical activity, and other environmental differences. Indeed, breast cancer risk increases significantly in females

---

**Figure 10-1.** Ten leading cancer types with the estimated new cancer cases and deaths by sex in the United States, 2013. *Excludes basal and squamous cell skin cancers and in situ carcinomas except those of the urinary bladder. Estimates are rounded to the nearest 10 (Modified with permission from Siegel RL, Miller KD, Jemal A: Cancer Statistics, 2017, CA Cancer J Clin. 2017 Jan;67(1):7-30.)

<table>
<thead>
<tr>
<th>Cancer Type</th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Estimated new cases</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prostate</td>
<td>161,360</td>
<td>19%</td>
</tr>
<tr>
<td>Lung &amp; bronchus</td>
<td>116,990</td>
<td>14%</td>
</tr>
<tr>
<td>Colon &amp; rectum</td>
<td>71,420</td>
<td>9%</td>
</tr>
<tr>
<td>Urinary bladder</td>
<td>60,490</td>
<td>7%</td>
</tr>
<tr>
<td>Melanoma of the skin</td>
<td>52,170</td>
<td>6%</td>
</tr>
<tr>
<td>Kidney &amp; renal pelvis</td>
<td>40,610</td>
<td>5%</td>
</tr>
<tr>
<td>Non-Hodgkin lymphoma</td>
<td>40,080</td>
<td>5%</td>
</tr>
<tr>
<td>Leukemia</td>
<td>36,290</td>
<td>4%</td>
</tr>
<tr>
<td>Oral cavity &amp; pharynx</td>
<td>35,720</td>
<td>4%</td>
</tr>
<tr>
<td>Liver &amp; intrahepatic bile duct</td>
<td>29,200</td>
<td>3%</td>
</tr>
<tr>
<td>All Sites</td>
<td><strong>836,150</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cancer Type</th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Estimated deaths</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lung &amp; bronchus</td>
<td>84,590</td>
<td>27%</td>
</tr>
<tr>
<td>Colon &amp; rectum</td>
<td>27,150</td>
<td>9%</td>
</tr>
<tr>
<td>Prostate</td>
<td>26,730</td>
<td>8%</td>
</tr>
<tr>
<td>Pancreas</td>
<td>22,300</td>
<td>7%</td>
</tr>
<tr>
<td>Liver &amp; intrahepatic bile duct</td>
<td>19,610</td>
<td>6%</td>
</tr>
<tr>
<td>Leukemia</td>
<td>14,300</td>
<td>4%</td>
</tr>
<tr>
<td>Esophagus</td>
<td>12,720</td>
<td>4%</td>
</tr>
<tr>
<td>Urinary bladder</td>
<td>12,240</td>
<td>4%</td>
</tr>
<tr>
<td>Non-Hodgkin lymphoma</td>
<td>11,450</td>
<td>4%</td>
</tr>
<tr>
<td>Brain &amp; other nervous system</td>
<td>9,620</td>
<td>3%</td>
</tr>
<tr>
<td>All Sites</td>
<td><strong>318,420</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>
who have migrated from Asia to America.\textsuperscript{5} The range of breast cancer mortality rates is much less (approximately 6 to 19 per 100,000) because of the more favorable survival of breast cancer in developed regions. As a result, breast cancer ranks as the fifth cause of death from cancer overall (458,000 deaths), but it is still the most frequent cause of cancer death in women in both developing (269,000 deaths, 12.7\% of total) and developed regions (estimated 189,000 deaths).\textsuperscript{4}

There is a 25-fold variation in colon cancer incidence worldwide.\textsuperscript{5} The incidence of colon and rectal cancer is higher in developed countries than in developing countries. The incidence rates are highest in North America, Australia and New Zealand, and Western Europe, and especially in Japanese men.\textsuperscript{5} In contrast, the incidence is relatively low in North Africa, South America, and Eastern, Southeastern, and Western Asia. These geographic differences are thought to reflect environmental exposures and are presumed to be related mainly to dietary differences in consumption of animal fat, meat, and fiber.\textsuperscript{5}

Worldwide liver cancer is the fifth most common cancer in men (523,000 cases, 7.9\% of the total) and the seventh in women (226,000 cases, 6.5\% of the total). Almost 85\% of liver cancer cases occur in developing countries, and particularly in men.\textsuperscript{4} The overall sex ratio of male to female is 2:4. The regions of high incidence are Eastern and Southeastern Asia, Middle and Western Africa, as well as Melanesia and Micronesia/Polynesia (particularly in men). Low rates are estimated in developed regions, with the exception of Southern Europe. There were an estimated 694,000 deaths from liver cancer in 2008 (477,000 in men, 217,000 in women), and because of its high fatality (overall ratio of mortality to incidence of 0.93), liver cancer is the third most common cause of death from cancer worldwide. The geographical distribution of the mortality rates is similar to that observed for incidence. Worldwide, the major risk factors for liver cancer are infection with hepatitis B and C viruses and consumption of foods contaminated with aflatoxin. Hepatitis B immunization in children has recently been shown to reduce the incidence of liver cancer.\textsuperscript{5}

In summary, the incidence rates of many common cancers vary widely by geography. This is due in part to genetic differences, including racial and ethnic differences. It is due also in part to differences in environmental and dietary exposures, factors that can potentially be altered. Therefore, establishment of regional and international databases is critical to improving our understanding of the etiology of cancer and will ultimately assist in the initiation of targeted strategies for global cancer prevention. Furthermore, the monitoring of cancer mortality rates and 5-year cancer-specific survival rates will identify regions where there are inequities of health care, so that access to health care can be facilitated and guidelines for treatment can be established.

### Table 10-2

Five-year relative survival rates adjusted to normal life expectancy by year of diagnosis, United States, 1975–2008

<table>
<thead>
<tr>
<th>CANCER TYPE</th>
<th>RELATIVE 5-YEAR SURVIVAL RATES (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cancers</td>
<td>49</td>
</tr>
<tr>
<td>Brain</td>
<td>22</td>
</tr>
<tr>
<td>Breast (female)</td>
<td>75</td>
</tr>
<tr>
<td>Uterine cervix</td>
<td>69</td>
</tr>
<tr>
<td>Colon</td>
<td>51</td>
</tr>
<tr>
<td>Uterine corpus</td>
<td>87</td>
</tr>
<tr>
<td>Esophagus</td>
<td>5</td>
</tr>
<tr>
<td>Hodgkin’s disease</td>
<td>72</td>
</tr>
<tr>
<td>Kidney</td>
<td>50</td>
</tr>
<tr>
<td>Larynx</td>
<td>66</td>
</tr>
<tr>
<td>Leukemia</td>
<td>34</td>
</tr>
<tr>
<td>Liver</td>
<td>3</td>
</tr>
<tr>
<td>Lung and bronchus</td>
<td>12</td>
</tr>
<tr>
<td>Melanoma of the skin</td>
<td>82</td>
</tr>
<tr>
<td>Multiple myeloma</td>
<td>25</td>
</tr>
<tr>
<td>Non-Hodgkin’s lymphoma</td>
<td>47</td>
</tr>
<tr>
<td>Oral cavity</td>
<td>53</td>
</tr>
<tr>
<td>Ovary</td>
<td>36</td>
</tr>
<tr>
<td>Pancreas</td>
<td>2</td>
</tr>
<tr>
<td>Prostate</td>
<td>68</td>
</tr>
<tr>
<td>Rectum</td>
<td>48</td>
</tr>
<tr>
<td>Stomach</td>
<td>15</td>
</tr>
<tr>
<td>Testis</td>
<td>83</td>
</tr>
<tr>
<td>Thyroid</td>
<td>92</td>
</tr>
<tr>
<td>Urinary bladder</td>
<td>73</td>
</tr>
</tbody>
</table>


### CANCER BIOLOGY

#### Hallmarks of Cancer

Although there are >100 types of cancer, it has been proposed that there are six essential alterations in cell physiology that dictate malignant growth: self-sufficiency of growth signals, insensitivity to growth-inhibitory signals, evasion of apoptosis (programmed cell death), potential for limitless replication, angiogenesis, and invasion and metastasis. Recently two additional hallmarks have emerged—reprogramming of energy metabolism and evading immune destruction. These hallmarks of cancer are being pursued as targets for cancer therapy (Fig. 10-4).

#### Cell Proliferation and Transformation

In normal cells, cell growth and proliferation are under strict control. In cancer cells, cells become unresponsive to normal growth controls, which leads to uncontrolled cell division. Human cells require several genetic changes for neoplastic transformation. Cell type–specific differences also exist in the process by which a normal cell is transformed into a cancerous one. Abnormally proliferating, transformed cells outgrow normal cells in the culture dish (i.e., in vitro) and commonly display several abnormal characteristics. These include loss of contact inhibition (i.e., cells continue to proliferate after a confluent monolayer is formed); an altered appearance and poor adherence to other cells or to the substratum; loss of anchorage dependence for growth; immortalization; and gain of tumorigenicity (i.e., the ability to give rise to tumors when injected into an appropriate host).

#### Cancer Initiation

Tumorigenesis is proposed to have three steps: initiation, promotion, and progression. Initiating events such as gain of function of genes known as oncogenes or loss of function of genes known as tumor-suppressor genes may lead a single cell to acquire a distinct growth advantage. Although tumors usually arise from a single cell or clone, it is thought that sometimes not a single cell but
rather a large number of cells in a target organ may have undergone the initiating genetic event. Thus, many normal-appearing cells may have an increased malignant potential. This is referred to as a field effect. The initiating events are usually genetic and occur as deletions of tumor-suppressor genes or amplification or mutation of oncogenes. Subsequent events can lead to accumulations of additional deleterious mutations in the clone.

Cancer is thought to be a disease of clonal progression as tumors arise from a single cell and accumulate mutations that confer on the tumor an increasingly aggressive behavior. Most tumors go through a progression from benign lesions to in situ tumors to invasive cancers (e.g., atypical ductal hyperplasia to ductal carcinoma in situ to invasive ductal carcinoma of the breast). Fearon and Vogelstein proposed the model for colorectal tumorigenesis presented in Fig. 10-5.9 Colorectal tumors arise from the mutational activation of oncogenes coupled with mutational inactivation of tumor-suppressor genes, the latter being the predominant change.9 Mutations in at least four or five genes are required for formation of a malignant tumor, while fewer changes suffice for a benign tumor. Although genetic mutations often occur in a preferred sequence, a tumor’s biologic properties are determined by the total accumulation of its genetic changes.

Gene expression is a multistep process that starts from transcription of a gene into messenger ribonucleic acid (mRNA) and then translation of this sequence into the functional protein. There are several controls at each level. In addition to alterations at the genome level (e.g., amplifications of a gene), alterations at the transcription level (e.g., methylation of the DNA leading to transcriptional silencing) or at the level of mRNA processing, mRNA stability, mRNA translation, or protein stability, all can alter the levels of critical proteins and thus contribute to tumorigenesis. Alternatively, changes in the genomic sequence can lead to a mutated product with altered function.

**Cell-Cycle Dysregulation in Cancer**

The proliferative advantage of tumor cells is a result of their ability to bypass a quiescent state. Cancer cells often show alterations in signal transduction pathways that lead to proliferation in response to external signals. Mutations or alterations in the expression of cell-cycle proteins, growth factors, growth factor receptors, intracellular signal transduction proteins, and nuclear transcription factors all can lead to disturbance of the basic regulatory mechanisms that control the cell cycle, allowing unregulated cell growth and proliferation.

The cell cycle is divided into four phases (Fig. 10-6).10 During the synthetic or S phase, the cell generates a single copy of its genetic material, whereas in the mitotic or M phase, the cellular components (including copies of DNA) are partitioned between two daughter cells. The G₁ and G₂ phases represent gap phases, during which the cells prepare themselves for
completion of the S and M phases, respectively. When cells cease proliferation, they exit the cell cycle and enter the quiescent state referred to as G0. The two key classes of regulatory molecules that regulate cellular progress through the cell cycle are the cyclins and the cyclin-dependent kinases (CDKs), which associate to form an activated heterodimer. CDKs are expressed constitutively and have a catalytic activity (phosphorylation of downstream proteins), whereas the cyclins serve a regulatory function and are synthesized at specific times during the cell cycle. Two families of genes, the cip/kip (CDK interacting protein/Kinase inhibitory protein) family and the INK4a/ARF (Inhibitor of Kinase 4/Alternative Reading Frame) family

---

**Figure 10-4.** Hallmarks of cancer and their therapeutic implications. Drugs that interfere with each of the acquired capabilities necessary for tumor growth and progression are in clinical trials and in some cases approved for clinical use in treating forms of human cancer. The drugs listed are illustrative examples. (Reproduced with permission from Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation, CCell. 2011 Mar 4;144(5):646-674.)

---

**Figure 10-5.** The adenoma-carcinoma model of human colorectal carcinogenesis. The neoplastic process is initiated by mutations in the adenomatous polyposis coli (APC) or β-catenin genes. Tumor progression results from mutations in other genes (e.g., K-ras, Smad 4 and p53) and the development of genomic instability. Patients with familial adenomatous polyposis inherit mutations in the APC gene and develop multiple aberrant crypt foci. Some of these may progress to cancer as they acquire other genetic mutations. (Reproduced with permission from Li C-J, Zhang, X, Fan G-W. Updates in colorectal cancer stem cell research, J Cancer Res Ther. 2014 Dec;10 Suppl:233-239.)
Normal cellular genes that contribute to cancer when abnormal are called oncogenes. The normal counterpart of such a gene is referred to as a proto-oncogene. Oncogenes are usually designated by three-letter abbreviations, such as myc or ras. Oncogenes are further designated by the prefix “v-“ for virus or “c-“ for cell or chromosome, corresponding to the origin of the oncogene when it was first detected. Proto-oncogenes can be activated (show increased activity) or overexpressed (expressed at increased protein levels) by translocation (e.g., abl), promoter insertion (e.g., c-myc), mutation (e.g., ras), or amplification (e.g., HER2/neu). More than 100 oncogenes have been identified.

 Oncogenes may be growth factors (e.g., platelet-derived growth factor), growth factor receptors (e.g., HER2), intracellular signal transduction molecules (e.g., ras), nuclear transcription factors (e.g., c-myc), or other molecules involved in the regulation of cell growth and proliferation. Growth factors are ubiquitous proteins that are produced and secreted by cells locally and that stimulate cell proliferation by binding specific cell-surface receptors on the same cells (autocrine stimulation) or on neighboring cells (paracrine stimulation). Persistent overexpression of growth factors can lead to uncontrolled autostimulation and neoplastic transformation. Alternatively, growth factor receptors can be aberrantly activated (turned on) through mutations or overexpression (continually presenting cells with growth-stimulatory signals, even in the absence of growth factors), which leads cells to respond as if growth factor levels are altered. The growth-stimulating effect of growth factors and other mitogens is mediated through postreceptor signal transduction molecules. These molecules mediate the passage of growth signals from the outside to the inside of the cell and then to the cell nucleus, initiating the cell cycle and DNA transcription. Aberrant activation or expression of cell-signaling molecules, cell-cycle molecules, or transcription factors may play an important role in neoplastic transformation. Protein tyrosine kinases account for a large portion of known oncogenes. One of the best-studied oncogenes, HER2 is discussed as an example later.

HER2, also known as neu or c-erbB-2, is a member of the epidermal growth factor receptor (EGFR) family and is one of the best-characterized tyrosine kinases. Unlike other receptor tyrosine kinases, HER2/neu does not have a direct soluble ligand. It plays a key role in signaling, however, because it is the preferred partner in heterodimer formation with all the other EGFR family members (EGFR/c-erbB-1, HER2/c-erbB-3, and HER3/c-erbB-4), which bind at least 30 ligands, including epidermal growth factor (EGF), transforming growth factor α (TGFα), heparin-binding EGF-like growth factor, amphiregulin, and heregulin. Heterodimerization with HER2 potentiates recycling of receptors rather than degradation, enhances signal potency and duration, increases affinity for ligands, and increases catalytic activity.

HER2 can interact with different members of the HER family and activate mitogenic and antiapoptotic pathways (Fig. 10-7). The specificity and potency of the intracellular signals are affected by the identity of the ligand, the composition of the receptors, and the phosphotyrosine-binding proteins associated with the erbB molecules. The Ras- and Shc-activated mitogen-activated protein kinase (MAPK) pathway is a target of all erbB ligands, which increase the transcriptional activity of early response genes such as c-myc, c-fos, and c-jun. MAPK-independent pathways such as the phosphoinosito-3 kinase (PI3K) pathway also are activated by most erbB dimers, although the potency and kinetics of activation may differ. Stimulation of the PI3K pathway through HER2 signaling also can lead to activation of survival molecule Akt, which suppresses apoptosis through multiple mechanisms.

The mutant rat neu gene was first recognized as an oncogene in neuroblastomas from carcinogen-treated rats. The HER2 gene is frequently amplified and the protein overexpressed in many cancers, including breast, ovarian, lung, gastric, and oral cancers. Overexpression of HER2 results in ligand-independent activation of HER2 kinase, which leads to mitogenic signaling. HER2 overexpression is associated with increased cell proliferation and anchorage-independent growth as well as resistance to proapoptotic stimuli. Further, overexpression of HER2 increases cell migration and upregulates the activities of matrix metalloproteinases (MMPs) and in vitro invasiveness. In animal models, HER2 increases tumorigenicity, angiogenesis, and metastasis. These results all suggest that HER2 plays a key role in cancer biology. More recently, HER2 mutations have also been reported in human cancer, including 3% of patients with lung cancer. A phase 2 study of the irreversible kinase inhibitor neratinib showed it to have efficacy in HER2-mutated breast cancer lacking HER amplification.

The critical role of HER2 in cancer biology has been leveraged for therapeutics, leading to several HER2-targeted drugs with different mechanism of action approved by the Food and Drug Administration (FDA): monoclonal antibodies trastuzumab and pertuzumab, small molecule inhibitor lapatinib, and antibody-drug conjugate ado-trastuzumab emtansine. Anti-HER2 agents first showed efficacy in the metastatic setting but
are now routinely used as adjuvant therapy of breast cancer and also as neoadjuvant treatments (“up-front chemotherapy”).

**Alterations in Apoptosis in Cancer Cells**

Apoptosis is a genetically regulated program to dispose of cells. Cancer cells must avoid apoptosis if tumors are to arise. The growth of a tumor mass is dependent not only on an increase in proliferation of tumor cells but also on a decrease in their apoptotic rate. Apoptosis is distinguished from necrosis because it leads to several characteristic changes. Soon after undergoing apoptosis, membrane phosphatidylserine translocates from the inner face of the plasma membrane to the cell surface where it can be detected via the use of a fluorescent conjugate of Annexin V, a protein that exhibits a high affinity for phosphatidylserine. Late in apoptosis there are characteristic changes in nuclear morphology, such as chromatin condensation, nuclear fragmentation, and DNA laddering, as well as membrane blebbing. Apoptotic cells are then engulfed and degraded by phagocytic cells. The effectors of apoptosis are a family of proteases called caspases (cysteine-dependent and aspartate-directed proteases). The initiator caspases (e.g., 8, 9, and 10), which are upstream, cleave the downstream executioner caspases (e.g., 3, 6, and 7) that carry out the destructive functions of apoptosis.

Two principal molecular pathways signal apoptosis by cleaving the initiator caspases with the potential for crosstalk: the mitochondrial pathway and the death receptor pathway. In the mitochondrial (or intrinsic) pathway, death results from the release of cytochrome c from the mitochondria. Cytochrome c, procaspase 9, and apoptotic protease activating factor 1 (Apaf-1) form an enzyme complex, referred to as the **apoptosome**, that activates the effector caspases. In addition to these proteins, the mitochondria contain other proapoptotic proteins such as Smac/DIABLO (second mitochondria-derived activator of caspase/direct inhibitor of apoptosis-binding protein with low pI). The mitochondrial pathway can be stimulated by many factors, including DNA damage, reactive oxygen species, or the withdrawal of survival factors. The permeability of the mitochondrial membrane determines whether the apoptotic pathway will proceed. The Bcl-2 family of regulatory proteins includes both proapoptotic proteins (e.g., Bax, BAD, and Bak) and antiapoptotic proteins (e.g., Bcl-2 and Bcl-xL). The activity of the Bcl-2 proteins is centered on the mitochondria, where they regulate...
membrane permeability. Growth factors promote survival signaling through the PI3K/Akt pathway, which phosphorylates and inactivates proapoptotic BAD. In contrast, growth factor withdrawal may promote apoptosis through signaling by unphosphorylated BAD. The heat shock proteins, including Hsp70 and Hsp27, are also involved in inhibition of downstream apoptotic pathways by blocking formation of the apoptosome complex and inhibiting release of cytochrome c from the mitochondria.19

The second principal apoptotic pathway is the death receptor pathway, sometimes referred to as the extrinsic pathway. Cell-surface death receptors include Fas/APO1/CD95, tumor necrosis factor receptor 1, and KILL-ER/DR5, which bind their ligands Fas-L, tumor necrosis factor (TNF), and TNF-related apoptosis-inducing ligand (TRAIL), respectively. When the receptors are bound by their ligands, they form a death-inducing signaling complex (DISC). At the DISC, procaspase 8 and procaspase 10 are cleaved, yielding active initiator caspases.20 The death receptor pathway may be regulated at the cell surface by the expression of “decoy” receptors for Fas (DcR3) and TRAIL (TRID and TRUNDD). The decoy receptors are closely related to the death receptors but lack a functional death domain; therefore, they bind death ligands but do not transmit a death signal. Another regulatory group is the FADD-like interleukin-1 protease-inhibitory proteins (FLIPs). FLIPs have homology to caspase 8; they bind to the DISC and inhibit the activation of caspase 8. Finally, inhibitors of apoptosis proteins (IAPs) block caspase 3 activation and have the ability to regulate both the death receptor and the mitochondrial pathway.

In human cancers, aberrations in the apoptotic program include increased expression of Fas and TRAIL decoy receptors; increased expression of antiapoptotic Bcl-2; increased expression of the IAP-related protein survivin; increased expression of c-FLIP; mutations or downregulation of proapoptotic Bax, caspase 8, APAF1, XAF1, and death receptors CD95, TRAIL-R1, and TRAIL-R2; alterations of the p53 pathway; overexpression of growth factors and growth factor receptors; and activation of the PI3K/Akt survival pathway.20

Autophagy in Cancer Cells

Autophagy (self-eating) is a major cellular pathway for protein and organelle turnover. The autophagic pathway is a mechanism for the delivery of cellular materials to lysosomes for degradation. This process leads to the basal turnover of cell components and provides energy and macromolecular precursors. This process helps maintain a balance between anabolism and catabolism for normal cell growth and development. Inability to activate autophagy in response to nutrient deprivation, or constitutive activation of autophagy in response to stress, can lead to cell death; thus, autophagy is sometimes referred to as a second form of programmed cell death. Autophagy plays an essential role during starvation, cellular differentiation, cell death, and aging. Autophagy is also involved in the elimination of cancer cells by triggering a nonapoptotic cell death program, which suggests a negative role in tumor development. Mouse models that are heterozygotes for the beclin 1 gene, an important gene for autophagy, have altered autophagic response and show a high incidence of spontaneous tumors, which establishes a role for autophagy in tumor suppression.21 This also suggests that mutations in other genes operating in this pathway may contribute to tumor formation through deregulation of autophagy. However, autophagy also acts as a stress response mechanism to protect cancer cells from low nutrient supply or therapeutic insults. Thus, in cancer, autophagy can have opposing and context-dependent roles, and interventions to both stimulate and inhibit autophagy have been proposed as possible anticancer treatments. Studies on the molecular controls of autophagy are ongoing and are expected to generate novel therapeutic strategies. Chloroquin is an antimalarial drug that acts as an autophagic inhibitor by blocking the autophagosome and has been tested for its anticancer properties.

Cancer Invasion

A feature of malignant cells is their ability to invade the surrounding normal tissue. Tumors in which the malignant cells appear to lie exclusively above the basement membrane are referred to as in situ cancer, whereas tumors in which the malignant cells are demonstrated to breach the basement membrane, penetrating into surrounding stroma, are termed invasive cancer. The ability to invade involves changes in adhesion, initiation of motility, and proteolysis of the extracellular matrix (ECM).

Cell-to-cell adhesion in normal cells involves interactions between cell-surface proteins. Calcium adhesion molecules of the cadherin family (E-cadherin, P-cadherin, and N-cadherin) are thought to enhance the cells’ ability to bind to one another and suppress invasion. Migration occurs when cancer cells penetrate and attach to the basal matrix of the tissue being invaded; this allows the cancer cell to pull itself forward within the tissue. Attachment to glycoproteins of the ECM such as fibronectin, laminin, and collagen is mediated by tumor cell integrin receptors. Integrins are a family of glycoproteins that form heterodimeric receptors for ECM molecules. The integrins can form at least 25 distinct pairings of their α and β subunits, and each pairing is specific for a unique set of ligands. In addition to regulating cell adhesion to the ECM, integrins relay molecular signals regarding the cellular environment that influence shape, survival, proliferation, gene transcription, and migration.

Factors that are thought to play a role in cancer cell motility include autocrine motility factor, autotaxin, scatter factor (also known as hepatocyte growth factor), TGFα, EGF, and insulin-like growth factors. Also, serine, cysteine, and aspartic proteinases and MMPs have all been implicated in cancer invasion. Urokinase and tissue plasminogen activators (uPA and tPA) are serine proteinases that convert plasminogen into plasmin. Plasmin, in turn, can degrade several ECM components. Plasmin also may activate MMPs. uPA has been more closely correlated with tissue invasion and metastasis than tPA. Plasminogen activator inhibitors 1 and 2 (PAI-1 and PAI-2) are produced in tissues and counteract the activity of plasminogen activators.

MMPs comprise a family of metal-dependent endopeptidases. Upon activation, MMPs degrade a variety of ECM components. Although MMPs often are referred to by their common names, which reflect the ECM component for which they have specificity, a sequential numbering system has been adopted for standardization. For example, collagenase-1 is now referred to as MMP-1. The MMPs are further classified as secreted and membrane-type MMPs. Most of the MMPs are synthesized as inactivezymogens (pro-MMP) and are activated by proteolytic removal of the propeptide domain outside the cell by other active MMPs or serine proteinases.

MMPs are upregulated in almost every type of cancer. Some of the MMPs are expressed by cancer cells, whereas others are expressed by the tumor stromal cells. Experimental models have demonstrated that MMPs promote cancer progression by
increasing cancer cell growth, migration, invasion, angiogenesis, and metastasis. MMPs exert these effects by cleaving not only structural components of the ECM but also growth factor–binding proteins, growth factor precursors, cell adhesion molecules, and other proteases to provide a growth advantage. The activity of MMPs is regulated by their endogenous inhibitors and tissue inhibitors of MMPs (TIMP-1, TIMP-2, TIMP-3, and TIMP-4).

**Angiogenesis**

Angiogenesis is the establishment of new blood vessels from a preexisting vascular bed. This neovascularization is essential for tumor growth and metastasis. Tumors develop an angiogenic phenotype as a result of accumulated genetic alterations and in response to local selection pressures such as hypoxia. Many of the common oncogenes and tumor-suppressor genes have been shown to play a role in the induction of angiogenesis.

In response to the angiogenic switch, pericytes retract and the endothelium secretes several growth factors such as basic fibroblast growth factor, platelet-derived growth factor (PDGF), and insulin-like growth factor. The basement membrane and stroma around the capillary are proteolytically degraded, a process that is mediated in most part by uPA. The endothelium then migrates through the degraded matrix, initially as a solid cord and later forming lumina. Finally, sprouting tips anastomose to form a vascular network surrounded by a basement membrane.

Angiogenesis is mediated by factors produced by various cells, including tumor cells, endothelial cells, stromal cells, and inflammatory cells. The first proangiogenic factor was identified by Folkman and colleagues in 1971. Since then, several other factors have been shown to be proangiogenic or antiangiogenic. Of the angiogenic stimulators, the best studied are the vascular endothelial growth factors (VEGFs). The VEGF family consists of six growth factors (VEGF-A, VEGF-B, VEGF-C, VEGF-D, VEGF-E, and placental growth factor) and three receptors (VEGFR1 or Flt-1, VEGFR2 or KDR/FLK-1, and VEGFR3 or Flt-4). Neuropilin 1 and 2 also may act as receptors for VEGF. VEGF is induced by hypoxia and by different growth factors and cytokines, including EGF, PDGF, TNF-α, TGFβ, and interleukin-1β. VEGF has various functions, including increasing vascular permeability, inducing endothelial cell proliferation and tube formation, and inducing endothelial cell synthesis of proteolytic enzymes such as uPA, PAI-1, urokinase plasminogen activator receptor, and MMP-1. Furthermore, VEGF may mediate blood flow by its effects on the vasodilator nitric oxide and act as an endothelial survival factor, thus protecting the integrity of the vasculature. The proliferation of new lymphatic vessels, lymphangiogenesis, is also thought to be controlled by the VEGF family. Signaling in lymphatic cells is thought to be modulated by VEGFR3. Experimental studies with VEGF-C and VEGF-D have shown that they can induce tumor lymphangiogenesis and direct metastasis via the lymphatic vessels and lymph nodes.

PDGFs A, B, C, and D also play important roles in angiogenesis. PDGFs are produced by tumor cells as well as supporting cells in the tumor microenvironment. PDGFs can enhance endothelial cell proliferation directly, and they can also upregulate VEGF expression in vascular smooth muscle cells, promoting endothelial cell survival via a paracrine effect. The angiopoietins, angiopoietin-1 and angiopoietin-2 (Ang-1 and Ang-2), are thought to regulate blood vessel maturation. Ang-1 and Ang-2 both bind to the angiopoietin-1 receptor (also known as tyrosine-protein kinase receptor TIE-2), but only the binding of Ang-1 activates signal transduction; thus Ang-2 is an Ang-1 antagonist. Ang-1, via the Tie-2 receptor, induces remodeling and stabilization of blood vessels. Therefore, the balance between these factors determines the angiogenetic capacity of a tumor.

Tumor angiogenesis is regulated by several factors in a coordinated fashion. In addition to upregulation of proangiogenic molecules, angiogenesis also can be encouraged by suppression of naturally occurring inhibitors. Such inhibitors of angiogenesis include thrombospondin 1 and angiostatin. Angiogenesis is a prerequisite not only for primary tumor growth but also for metastasis. Angiogenesis in the primary tumor, as determined by microvessel density, has been demonstrated to be an independent predictor of distant metastatic disease and survival in several cancers. Expression of angiogenic factors such as VEGFs has had prognostic value in many studies. These findings further emphasize the importance of angiogenesis in cancer biology.

**Metastasis**

Metastases arise from the spread of cancer cells from the primary site and the formation of new tumors in distant sites. The metastatic process consists of a series of steps that need to be completed successfully (Fig. 10-8). First, the primary cancer must develop access to the circulation through either the blood circulatory system or the lymphatic system. After the cancer cells are shed into the circulation, they must survive. Next, the circulating cells lodge in a new organ and extravasate into the new tissue. Next, the cells need to initiate growth in the new tissue and eventually establish vascularization to sustain the new tumor. Overall, metastasis is an inefficient process, although the initial steps of hematogenous metastasis (the arrest of tumor cells in the organ and extravasation) are believed to be performed efficiently. Only a small subset of cancer cells is then able to initiate micrometastases, and an even smaller portion goes on to grow into macrometastases.

Metastases can sometimes arise several years after the treatment of primary tumors. For example, although most breast cancer recurrences occur within the first 10 years after the initial treatment and recurrences are rare after 20 years, breast cancer recurrences have been reported decades after the original tumor. This phenomenon is referred to as dormancy, and it remains one of the biggest challenges in cancer biology. Persistence of solitary cancer cells in a secondary site such as the liver or bone marrow is one possible contributor to dormancy. Another explanation of dormancy is that cells remain viable in a quiescent state and then become reactivated by a physiologically perturbing event. Interestingly, primary tumor removal has been proposed to be a potentially perturbing factor. An alternate explanation is that cells establish preangiogenic metastases in which they continue to proliferate but that the proliferative rate is balanced by the apoptotic rate. Therefore, when these small metastases acquire the ability to become vascularized, substantial tumor growth can be achieved at the metastatic site, leading to clinical detection. More recently, it has been proposed that dormancy may be the result of the host losing immunologic control of subclinical metastatic foci of disease either through loss of immune cell populations with antigen-specific capabilities or via the mutation of tumor cells such that their immunogenicity is altered.

Several types of tumors metastasize in an organ-specific pattern. One explanation for this is mechanical and is based on
When different tumor types and their preferred metastasis sites were compared, 66% of organ-specific metastases were explained on the basis of blood flow alone. The other explanation for preferential metastasis is what is referred to as the “seed and soil” theory, the dependence of the seed (the cancer cell) on the soil (the secondary organ). According to this theory, once cells have reached a secondary organ, their growth efficiency in that organ is based on the compatibility of the cancer cell’s biology with its new microenvironment. For example, breast cancer cells may grow more efficiently in bone than in some other organs because of favorable molecular interactions that occur in the bone microenvironment. The ability of cancer cells to grow in a specific site likely depends on features inherent to the cancer cell, features inherent to the organ, and the interplay between the cancer cell and its microenvironment.

Many of the oncogenes discovered to date, such as HER2 and ras, are thought to potentiate not only malignant transformation but also one or more of the steps required in the metastatic process. Experimental models have suggested a role for several molecules, including RhoC, osteopontin and interleukin-11, and Twist, in tumor metastasis. Metastasis also may involve the loss of metastasis-suppressor genes. Laboratory work involving cancer cell lines that have been selected to have a higher metastatic potential have led to the realization that these more highly metastatic cells have a different gene expression profile than their less metastatic parental counterparts. This in turn has led to the currently held belief that the ability of a primary tumor to metastasize may be predictable by analysis of its gene expression profile. Indeed, several studies have focused on identifying a gene expression profile or a molecular signature that is associated with metastasis. It has been shown that such a gene expression profile can be used to predict the probability that the patient will remain free of distant metastasis. This suggests that the metastatic potential of a tumor is already predetermined by the genetic alterations that the cancer cells acquire early in tumorigenesis. Notably, this hypothesis differs from the multistep tumorigenesis theory in that the ability to metastasize is considered an inherent quality of the tumor from the beginning. It is assumed that metastasis develops not from a few rare cells in the primary tumor that acquire the ability to metastasize but that all cells in tumors with such molecular signatures develop the ability to metastasize. The reality probably lies in between since some early genetic changes detectable in the entire tumor can give tumors an advantage in the metastatic process, whereas additional genetic changes can give a clone of cells additional advantages, thus allowing them to succeed in metastasis.

**Epithelial-Mesenchymal Transition**

A regulatory program referred to as epithelial-mesenchymal transition (EMT) is a fundamental event in morphogenesis. During EMT, epithelial cells are converted to migratory and
invasive mesenchymal cells. EMT has also been implicated as the mechanism through which epithelial cells acquire the ability to migrate, invade, resist apoptosis, and metastasize in the setting of cancer. EMT is a developmental process, and a set of pleiotropically acting transcriptional factors, including Snail, Twist, Slug, and Zeb1/2 orchestrate EMT. Several of these transcription factors can directly repress E-cadherin gene expression reducing levels of this key suppressor of motility and invasiveness in cancer cells. This process also entails the induction of mesenchymal markers such as vimentin. It has been proposed that the process of invasion and metastases requires significant plasticity, suggesting that EMT is required for invasion, intravasation, and extravasation. Thus, suppression of EMT regulators (and consequently EMT reversion, or MET) can be important for metastatic outgrowth.

Cancer Stem Cells

Stem cells are cells that have the ability to perpetuate themselves through self-renewal and to generate mature cells of a particular tissue through differentiation. It was first documented for leukemia and multiple myeloma that only a small subset of cancer cells is capable of extensive proliferation. It has subsequently also been shown for many solid cancers that only a small proportion of cells is clonogenic in culture and in vivo. Similarly, in many solid tumor types only a small proportion of cells is clonogenic in culture and in vivo. If indeed tumor growth and metastasis are driven by a small population of cancer stem cells, this may alter our current approaches to cancer therapy. Currently available drugs can shrink metastatic tumors but often cannot eradicate them. The failure of these treatments usually is attributed to the acquisition of drug resistance by the cancer cells; however, the cancer stem cell hypothesis raises the possibility that existing therapies may simply fail to kill cancer stem cells effectively. Therapeutic approaches targeting stem cells specifically are under study.

Figure 10-9. Accumulation of somatic mutations acquired by the cancer cell. Mutations may be acquired while the cell lineage is phenotypically normal, reflecting intrinsic mutations acquired during normal cell division as well as the effects of exogenous mutagens. Other processes such as example DNA repair defects may contribute to the mutational burden. Passenger mutations do not have any effect on the cancer cell, but driver mutations cause clonal expansion. Relapse after chemotherapy can be associated with resistance mutations that may predate the initiation of treatment. (Reproduced with permission from Stratton MR, Campbell PJ, Futreal PA. The cancer genome, Nature. 2009 Apr 9;458(7239):719–724.)

CANCER ETIOLOGY

Cancer Genomics

One widely held opinion is that cancer is a genetic disease that arises from an accumulation of genomic alterations that leads to the selection of cells with increasingly aggressive behavior. These alterations may lead either to a gain of function by oncogenes or to a loss of function by tumor-suppressor genes. These acquired gene alterations present within the tumor are termed somatic mutations to distinguish them from germline mutations that are inherited from parents and transmitted to offspring. Somatic mutations in a cancer genome may consist of several classes of DNA sequence changes. These include substitutions of one base by another; insertions or deletions of small or large segments of DNA; rearrangements, in which the DNA sequence has been broken and then rejoined to another DNA segment; copy number losses that may result in complete absence of a DNA sequence; and copy number gains over and above the two copies present in the normal diploid genome.

Somatic mutations in a cancer cell genome have accumulated over the lifetime of the patient (Fig. 10-9). DNA in normal cells is continuously damaged by internal and external mutagens. Most of this damage is repaired; however, a small fraction may remain as fixed mutations. Mutation rates increase in the presence of substantial exogenous mutagenic exposures, such as tobacco carcinogens or various forms of radiation, including ultraviolet light. These exposures are associated with increased rates of lung and skin cancer, respectively, and somatic mutations within such cancers often exhibit the distinctive mutational signatures known to be associated with the mutagen. The rates of somatic mutations are also increased in several rare inherited diseases, such as Fanconi anemia, ataxia telangiectasia, and xeroderma pigmentosum, which are associated with increased risks of cancer. The rest of the somatic mutations in a cancer cell have been acquired after the cancer cell already shows phenotypic evidence of neoplastic change. Whether the somatic
mutation rate is always higher during this part of the lineage is controversial. This is clearly the case for some cancers. For instance, colorectal and endometrial cancers with defective DNA mismatch repair due to abnormalities in genes such as MLH1 and MSH2, exhibit increased rates of single nucleotide changes and small insertions/deletions at repetitive noncoding polynucleotide tracts known as microsatellites. It has been proposed that one early step in tumor progression is the development of a “mutator phenotype” that is the result of mutations in genes that normally function in the maintenance of genetic stability. This hypothesis was formulated in order to account for the disparity between the low frequency of spontaneous mutations in normal cells as compared to the large number of mutations seen in human tumors.

To date about 300 genes that have been reported to be mutated and causally implicated in cancer development. Ninety percent of cancer genes are mutated at the somatic or tumor level, 20% show germline mutations, and 10% show both. The most common class of genomic alterations among the known cancer genes is a chromosomal translocation that creates a chimeric gene. Many more cancer genes have been found in leukemias, lymphomas, and sarcomas than in other types of cancer; and these genes are usually altered by chromosomal translocation. The most common cancer genes are protein kinases. Several domains that are involved in DNA binding and transcriptional regulation are also common in proteins encoded by cancer genes. Somatic mutations in a cancer genome may be classified according to its consequences for cancer development. “Driver” mutations confer a growth advantage to the cells carrying them and have been positively selected during the evolution of the cancer. The remainder of mutations are “bystanders” or “passengers” that do not confer growth advantage. It is likely that most somatic mutations are passenger mutations. Each tumor may have dozens to hundreds of genomic alterations, making it critical to determine which alterations are indeed drivers, and potentially better therapeutic targets.

There have been many recent advancements in large-scale databases and tools to catalogue and interpret genomic variants in cancer patient populations. Currently, the NCI Genomic Data Commons provides a unified data repository, The Cancer Genome Atlas (TCGA), the NCI Center for Cancer Genomics (CCG), and the childhood cancer initiative entitled Therapeutically Applicable Research to Generate Effective Treatments (TARGET), as well as a suite of tools for users to interact with the GDC data and provide their own data. Other cancer genome repositories include the Catalogue of Somatic Mutations in Cancer (COSMIC) and the International Cancer Genome Consortium (ICGC). The Precision Medicine Initiative launched in 2016 that includes the All of Us Research Program, which will collect genetic data, biologic samples, and other clinical information from at least 1 million volunteer participants. To facilitate the clinical and biologic interpretation of genomic variants in cancer genomes, several open-access tools have been developed and expanded, including MuSiC, MutSigCV, and OncodriveFM. Other curated database resources used to annotate clinical phenotypes to variants observed in cancer genomes include ClinVAR, Clinical Interpretation of Variants in Cancer (CiVic), and the Precision Medicine Knowledgebase (PMKB). These resources and tools are being utilized to conduct pan-cancer analyses to characterize genomic variation and other molecular aberrations observed across tumors to define cancer drivers, clinically actionable targets, and prognostic and predictive signatures. This information is being integrated into clinical practice in many tumor types, such as lung cancer, where molecular drivers are being taken into consideration when selecting systemic therapies (Fig. 10-10). TCGA data was recently leveraged in a study to identify TRK fusions in tumors from six different cancer types, which led to the development of novel TRK inhibitor therapies.

**Tumor Heterogeneity and Molecular Evolution**

There is increasing recognition that tumors are heterogeneous; this represents an important challenge to utilizing genomic alterations to personalize cancer therapy (Fig. 10-11). First,
there may be significant heterogeneity between cancer patients, such that patients with tumors that seem similar histologically, may differ in genomic alterations and in malignant potential.46-48 Second, during cancer progression, subclones frequently arise, resulting in differences in the proportion and pattern of genomic alterations between the primary tumor and the metastases or local-regional recurrences.35 Third, there may also be significant heterogeneity within any one tumor deposit, with spatially separated heterogeneous somatic mutations and chromosomal imbalances.49 Such spatial heterogeneity of subclones within the primary tumor or metastases provides an additional challenge, as it has been proposed that sequencing of a biopsy specimen or only a portion of the tumor could miss therapeutically relevant genomic alterations. The genomic alterations found in a tumor can also change under the selective pressure of a targeted therapy, adding to the challenge of implementing genomically informed personalized therapy.

**Genes Associated With Hereditary Cancer Risk**

Most of our information on human cancer genes has been gained from hereditary cancers. In the case of hereditary cancers, the individual carries a particular germline mutation in every cell. To date, over 70 genes have been associated with hereditary cancers (Table 10-3).43 A few of these hereditary cancer genes are oncogenes, but most are tumor-suppressor genes. Although hereditary cancer syndromes are rare, somatic mutations that occur in sporadic cancer have been found to disrupt the cellular pathways altered in hereditary cancer syndromes, which suggests that these pathways are critical to normal cell growth, cell cycle, and proliferation. Recently, the results of a genome-wide association study of breast cancer in over 120,000 cases and 100,000 controls identified 65 new loci that are associated with overall breast cancer risk.50

The following factors may suggest the presence of a hereditary cancer51:

1. Tumor development at a much younger age than usual
2. Presence of bilateral disease
3. Presence of multiple primary malignancies
4. Presentation of a cancer in the less affected sex (e.g., male breast cancer)
5. Clustering of the same cancer type in relatives
6. Occurrence of cancer in association with other conditions such as mental retardation or pathognomonic skin lesions

It is crucial that all surgeons caring for cancer patients be aware of hereditary cancer syndromes, because a patient’s genetic background has significant implications for patient and family counseling, planning of surgical therapy, and cancer screening and prevention. Some of the more commonly encountered hereditary cancer syndromes are discussed here.

**rb1Gene.** The retinoblastoma gene rb1 was the first tumor suppressor to be cloned. The rb1 gene product, the Rb protein, is a regulator of transcription that controls the cell cycle, differentiation, and apoptosis in normal development.52 Retinoblastoma has long been known to occur in hereditary and nonhereditary forms. Interestingly, although most children with an affected parent develop bilateral retinoblastoma, some develop unilateral retinoblastoma. Furthermore, some children with an affected parent are not affected themselves but then have an affected child, which indicates that they are rb1 mutation carriers. These findings led to the theory that a single mutation is not sufficient for tumorigenesis. Alfred Knudson hypothesized that hereditary retinoblastoma involves two mutations, of which one is germ-line and one somatic, whereas nonhereditary retinoblastoma is due to two somatic mutations (Fig. 10-12).53 Thus, both hereditary and nonhereditary forms of retinoblastoma involve the same number of mutations, a hypothesis known as Knudson’s “two-hit” hypothesis. A “hit” may be a point mutation, a chromosomal deletion referred to as allelic loss, or a loss of heterozygosity, or silencing of an existing gene. Approximately 40% of retinoblastomas are hereditary and due to germline mutations in the RB1 gene. Children with hereditary RB are also at risk for developing a midline intracranial tumor, most commonly pineoblastoma.

**p53 and Li-Fraumeni Syndrome.** Li-Fraumeni syndrome (LFS) was first defined on the basis of observed clustering of malignancies, including early-onset breast cancer, soft tissue sarcomas, brain tumors, adrenocortical tumors, and leukemia.54 Criteria for classic LFS in an individual (the proband) include: (a) a bone or soft tissue sarcoma when younger than 45 years, (b) a first-degree relative with cancer before age 45 years, and (c) another first- or second-degree relative with either a sarcoma diagnosed at any age or any cancer diagnosed before age 45 years.55 Approximately 70% of LFS families have been shown to have germline mutations in the tumor-suppressor gene p53.56 Breast carcinoma, soft tissue sarcoma, osteosarcoma, brain tumors, adrenocortical carcinoma, Wilms’ tumor, and phyllodes tumor of the breast are strongly associated; pancreatic cancer is moderately associated; and leukemia and neuroblastoma are weakly associated with germline p53 mutations.57 Mutations of p53 have not been detected in approximately 30% of LFS families, and it is hypothesized that genetic alterations in other proteins interacting with p53 function may play a role in these families.

Of the known genes in human cancer, p53 is the most commonly mutated within cancer cells. The p53 protein regulates cell-cycle progression as well as apoptotic cell death as part of stress response pathways after exposure to ionizing or ultraviolet (UV) irradiation, chemotherapy, acidosis, growth factor deprivation, or hypoxia. When cells are exposed to stressors, p53 acts as a transcription factor for genes that induce cell-cycle arrest or apoptosis. A majority of p53 mutations are found within a central DNA recognition motif and disrupt DNA binding by p53. Families with germline missense mutations in the DNA-binding domain show a more highly penetrant phenotype than families with other p53 mutations.58 Furthermore, proband cancers are linked with significantly younger age at diagnosis in patients with missense mutations in the DNA-binding domain.59 It has become apparent that children and adults with LFS will benefit from intensive surveillance aimed at early detection of cancers and a modified version of the “Toronto protocol” that includes a combination of physical exams, blood tests, and imaging is recommended.

**BRCA1, BRCA2, and Hereditary Breast-Ovarian Cancer Syndrome.** It is estimated that 5% to 10% of breast cancers are hereditary. Of women with early-onset breast cancer (age 40 years or younger), nearly 10% have a germline mutation in one of the breast cancer genes BRCA1 or BRCA2.59 Mutation carriers are more prevalent among women who have a first- or second-degree relative with premenopausal breast cancer or ovarian cancer at any age. The likelihood of a BRCA mutation is higher in patients who belong to a population in which founder
| SYMBOL | NAME | TUMOR TYPES  
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ALK</td>
<td>anaplastic lymphoma kinase (Ki-1)</td>
<td>Neuroblastoma</td>
</tr>
<tr>
<td>APC</td>
<td>adenomatous polyposis of the colon gene</td>
<td>Colorectal, pancreatic, desmoid, hepatoblastoma, glioma, other CNS</td>
</tr>
<tr>
<td>ATM</td>
<td>ataxia telangiectasia mutated</td>
<td>Leukemia, lymphoma, medulloblastoma, glioma</td>
</tr>
<tr>
<td>BLM</td>
<td>Bloom syndrome</td>
<td>Leukemia, lymphoma, skin squamous cell, other cancers</td>
</tr>
<tr>
<td>BMPR1A</td>
<td>bone morphogenetic protein receptor, type IA</td>
<td>Gastrointestinal polyps</td>
</tr>
<tr>
<td>BRCA1</td>
<td>familial breast/ovarian cancer gene 1</td>
<td>Breast, ovarian</td>
</tr>
<tr>
<td>BRCA2</td>
<td>familial breast/ovarian cancer gene 2</td>
<td>Breast, ovarian, pancreatic</td>
</tr>
<tr>
<td>BRIP1</td>
<td>BRCA1 interacting protein C-terminal helicase 1</td>
<td>AML, leukemia, breast</td>
</tr>
<tr>
<td>BUB1B</td>
<td>BUB1 budding uninhibited by benzimidazoles 1 homolog beta (yeast)</td>
<td>Rhabdomyosarcoma</td>
</tr>
<tr>
<td>CDH1</td>
<td>cadherin 1, type 1, E-cadherin (epithelial) (ECAD)</td>
<td>Gastric, lobular cancer</td>
</tr>
<tr>
<td>CDK4</td>
<td>cyclin-dependent kinase 4</td>
<td>Melanoma</td>
</tr>
<tr>
<td>CDKN2A</td>
<td>cyclin-dependent kinase inhibitor 2A (p16[INK4a]) gene</td>
<td>Melanoma, pancreatic</td>
</tr>
<tr>
<td>CDKN2a(p14)</td>
<td>cyclin-dependent kinase inhibitor 2A–p14ARF protein</td>
<td>Melanoma, pancreatic</td>
</tr>
<tr>
<td>CHEK2</td>
<td>CHK2 checkpoint homolog (S pombe)</td>
<td>Breast</td>
</tr>
<tr>
<td>CYLD</td>
<td>familial cylindromatosis gene</td>
<td>Cylindroma</td>
</tr>
<tr>
<td>DDB2</td>
<td>damage-specific DNA binding protein 2</td>
<td>Skin basal cell, skin squamous cell, melanoma</td>
</tr>
<tr>
<td>DICER1</td>
<td>dicer 1, ribonuclease type III</td>
<td>Pleuropulmonary blastoma</td>
</tr>
<tr>
<td>EGFR</td>
<td>epidermal growth factor receptor (erythroblastic leukemia viral [v-erb-b] oncogene homolog, avian)</td>
<td>NSCLC</td>
</tr>
<tr>
<td>ERCC2, 3, 4, 5</td>
<td>excision repair cross-complementing rodent repair deficiency, complementation group</td>
<td>Skin basal cell, skin squamous cell, melanoma</td>
</tr>
<tr>
<td>EXT1</td>
<td>multiple exostoses type 1 gene</td>
<td>exostoses, osteosarcoma</td>
</tr>
<tr>
<td>FANCA, C, D2, E, F, G</td>
<td>Fanconi anemia, complementation group</td>
<td>AML, leukemia</td>
</tr>
<tr>
<td>FH</td>
<td>fumarate hydratase</td>
<td>leiomyomatas, renal</td>
</tr>
<tr>
<td>GPC3</td>
<td>glypican 3</td>
<td>Wilms’ tumor</td>
</tr>
<tr>
<td>HRAS</td>
<td>v-Ha-ras Harvey rat sarcoma viral oncogene homolog</td>
<td>v-Ha-ras Harvey rat sarcoma viral oncogene homolog</td>
</tr>
<tr>
<td>HRPT2</td>
<td>Hyperparathyroidism 2 (parafibromin)</td>
<td>parathyroid adenoma, multiple ossifying jaw fibroma</td>
</tr>
</tbody>
</table>

(Continued)
Table 10-3
Selected genes associated with hereditary cancer (Continued)

<table>
<thead>
<tr>
<th>SYMBOL</th>
<th>NAME</th>
<th>TUMOR TYPES (GERMLINE MUTATIONS)</th>
<th>CANCER SYNDROME</th>
</tr>
</thead>
<tbody>
<tr>
<td>KIT</td>
<td>v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog</td>
<td>GIST, epithelioma</td>
<td>Familial gastrointestinal stromal tumor</td>
</tr>
<tr>
<td>MADH4</td>
<td>Homolog of Drosophila Mothers Against Decapentaplegic 4 gene</td>
<td>Gastrointestinal polyps</td>
<td>Juvenile polyposis</td>
</tr>
<tr>
<td>MEN1</td>
<td>multiple endocrine neoplasia type 1 gene</td>
<td>Parathyroid adenoma, pituitary adenoma, pancreatic islet cell, carcinoid</td>
<td>Parathyroid adenoma, pituitary adenoma, pancreatic islet cell, carcinoid</td>
</tr>
<tr>
<td>MLH1</td>
<td><em>E. coli</em> MutL homolog gene</td>
<td>Colorectal, endometrial, ovarian</td>
<td>Hereditary nonpolyposis colorectal cancer, Turcot syndrome</td>
</tr>
<tr>
<td>MPL</td>
<td>myeloproliferative leukemia virus oncogene, thrombopoietin receptor</td>
<td>MPD</td>
<td>Familial essential thrombocytethmia</td>
</tr>
<tr>
<td>MSH2</td>
<td>mutS homolog 2 (<em>E. coli</em>)</td>
<td>colorectal, endometrial, ovarian</td>
<td>Hereditary non-polyposis colorectal cancer</td>
</tr>
<tr>
<td>MSH6</td>
<td>mutS homolog 6 (<em>E. coli</em>)</td>
<td>colorectal, endometrial, ovarian</td>
<td>Hereditary non-polyposis colorectal cancer</td>
</tr>
<tr>
<td>MUTYH</td>
<td>mutY homolog (<em>E. coli</em>)</td>
<td>Colorectal</td>
<td>Adenomatous polyposis coli</td>
</tr>
<tr>
<td>NBS1</td>
<td>Nijmegen breakage syndrome 1 (nibrin)</td>
<td>NHL, glioma, medulloblastoma, rhabdomyosarcoma</td>
<td>Nijmegen breakage syndrome</td>
</tr>
<tr>
<td>NF1</td>
<td>neurofibromatosis type 1 gene</td>
<td>Neurofibroma, glioma</td>
<td>Neurofibromatosis type 1</td>
</tr>
<tr>
<td>NF2</td>
<td>neurofibromatosis type 2 gene</td>
<td>Meningioma, acoustic neuroma</td>
<td>Neurofibromatosis type 2</td>
</tr>
<tr>
<td>PALB2</td>
<td>partner and localizer of BRCA2</td>
<td>Wilms tumor, medulloblastoma, AML, breast</td>
<td>Fanconi anemia N, breast cancer susceptibility</td>
</tr>
<tr>
<td>PHOX2B</td>
<td>paired-like homeobox 2b</td>
<td>Neuroblastoma</td>
<td>Familial neuroblastoma</td>
</tr>
<tr>
<td>PMS1</td>
<td>PMS1 postmeiotic segregation increased 1 (<em>S. cerevisiae</em>)</td>
<td>Colorectal, endometrial, ovarian</td>
<td>Hereditary non-polyposis colorectal cancer</td>
</tr>
<tr>
<td>PMS2</td>
<td>PMS2 postmeiotic segregation increased 2 (<em>S. cerevisiae</em>)</td>
<td>Colorectal, endometrial, ovarian, medulloblastoma, glioma</td>
<td>Hereditary nonpolyposis colorectal cancer, Turcot syndrome</td>
</tr>
<tr>
<td>PRKAR1A</td>
<td>protein kinase, cAMP-dependent, regulatory, type I, alpha (tissue specific extinguisher 1)</td>
<td>Myxoma, endocrine, papillary thyroid</td>
<td>Carney complex</td>
</tr>
<tr>
<td>PTCH</td>
<td>Homolog of Drosophila Patched gene</td>
<td>Skin basal cell, medulloblastoma</td>
<td>Neviod basal cell carcinoma syndrome</td>
</tr>
<tr>
<td>PTEN</td>
<td>phosphatase and tensin homolog gene</td>
<td>Hamartoma, glioma, prostate, endometrial</td>
<td>Cowden syndrome, Bannayan-Riley-Ruvalcaba syndrome</td>
</tr>
<tr>
<td>RB1</td>
<td>retinoblastoma gene</td>
<td>Retinoblastoma, sarcoma, breast, small cell lung</td>
<td>Familial retinoblastoma</td>
</tr>
<tr>
<td>RECQL4</td>
<td>RecQ protein-like 4</td>
<td>Osteosarcoma, skin basal and squamous cell</td>
<td>Rothmund-Thompson syndrome</td>
</tr>
<tr>
<td>RET</td>
<td>ret proto-oncogene</td>
<td>Medullary thyroid, papillary thyroid, pheochromocytoma</td>
<td>Multiple endocrine neoplasia 2A/2B</td>
</tr>
<tr>
<td>SBDS</td>
<td>Shwachman-Bodian-Diamond syndrome protein</td>
<td>AML, MDS</td>
<td>Schwachman-Diamond syndrome</td>
</tr>
</tbody>
</table>

(Continued)
mutations may be prevalent, such as in the Ashkenazi Jewish population. For a female BRCA1 mutation carrier, the cumulative risks of developing breast cancer and ovarian cancer by age 70 have been estimated to be 87% and 44%, respectively.60 The cumulative risks of breast cancer and ovarian cancer by age 70 in families with BRCA2 mutation have been estimated to be 84% and 27%, respectively.61 Although male breast cancer can occur with either BRCA1 or BRCA2 mutation, the majority of families (76%) with both male and female breast cancer have mutations in BRCA2.64 Besides breast and ovarian cancer, BRCA1 and BRCA2 mutations may be associated with increased risks for several other cancers. BRCA1 mutations confer a fourfold increased risk for colon cancer and threefold increased risk for prostate cancer.65 BRCA2 mutations confer a fivefold increased risk for prostate cancer, sevenfold in men younger than 65 years.62 Furthermore, BRCA2 mutations confer

<table>
<thead>
<tr>
<th>SYMBOL</th>
<th>NAME</th>
<th>TUMOR TYPES (GERMLINE MUTATIONS)</th>
<th>CANCER SYNDROME</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDH5</td>
<td>chromosome 11 open reading frame 79</td>
<td>Paraganglioma</td>
<td>Familial paraganglioma</td>
</tr>
<tr>
<td>SHD, B, D</td>
<td>succinate dehydrogenase complex</td>
<td>Paraganglioma, pheochromocytoma</td>
<td>Familial paraganglioma</td>
</tr>
<tr>
<td>SMARCB1</td>
<td>SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1</td>
<td>Malignant rhabdoid</td>
<td>Rhabdoid predisposition syndrome</td>
</tr>
<tr>
<td>STK11</td>
<td>serine/threonine kinase 11 gene (LKB1)</td>
<td>Jejunal hamartoma, ovarian, testicular, pancreatic</td>
<td>Peutz-Jeghers syndrome</td>
</tr>
<tr>
<td>SUFU</td>
<td>suppressor of fused homolog (Drosophila)</td>
<td>Medulloblastoma</td>
<td>Medulloblastoma predisposition</td>
</tr>
<tr>
<td>TCF1</td>
<td>transcription factor 1, hepatic (HNF1)</td>
<td>Hepatic adenoma, hepatocellular carcinoma</td>
<td>Familial Hepatic Adenoma</td>
</tr>
<tr>
<td>TP53</td>
<td>tumor protein p53</td>
<td>Breast, sarcoma, adrenocortical carcinoma, glioma, multiple other tumor types</td>
<td>Li-Fraumeni syndrome</td>
</tr>
<tr>
<td>TSC1</td>
<td>tuberous sclerosis 1 gene</td>
<td>Hamartoma, renal cell</td>
<td>Tuberous sclerosis 1</td>
</tr>
<tr>
<td>TSC2</td>
<td>tuberous sclerosis 2 gene</td>
<td>Hamartoma, renal cell</td>
<td>Tuberous sclerosis 2</td>
</tr>
<tr>
<td>TSHR</td>
<td>thyroid stimulating hormone receptor</td>
<td>Thyroid adenoma</td>
<td></td>
</tr>
<tr>
<td>VHL</td>
<td>von Hippel-Lindau syndrome gene</td>
<td>Renal, hemangioma, pheochromocytoma</td>
<td>von Hippel-Lindau syndrome</td>
</tr>
<tr>
<td>WRN</td>
<td>Werner syndrome (RECQL2)</td>
<td>Osteosarcoma, meningioma, others</td>
<td>Werner syndrome</td>
</tr>
<tr>
<td>WT1</td>
<td>Wilms’ tumor 1 gene</td>
<td>Wilms’</td>
<td>Denys-Drash syndrome, Frasier syndrome, Familial Wilms tumor</td>
</tr>
<tr>
<td>XPA, C</td>
<td>xeroderma pigmentosum, complementation group</td>
<td>Skin basal cell, skin squamous cell, melanoma</td>
<td>Xeroderma pigmentosum (A C)</td>
</tr>
</tbody>
</table>

A = amplification; AEL = acute eosinophilic leukemia; AL = acute leukemia; ALC = anaplastic large-cell lymphoma; ALL = acute lymphocytic leukemia; AML = acute myelogenous leukemia; ALCL = anaplastic large-cell lymphoma; ALL = acute lymphocytic leukemia; B-ALL = B-cell acute lymphocytic leukemia; B-CLL = B-cell lymphocytic leukemia; B-NHL = B-cell non-Hodgkin’s lymphoma; CLL = chronic lymphatic leukemia; CML = chronic myeloid leukemia; CMML = chronic myelomonocytic leukemia; CNS = central nervous system; D = large deletion; DFSP = dermatofibrosarcoma protuberans; DLBL = diffuse large B-cell lymphoma; DLCL = diffuse large-cell lymphoma; Dom = dominant; E = epithelial; F = frameshift; GIST = gastrointestinal stromal tumor; JMMCL = juvenile myelomonocytic leukemia; L = leukemia/lymphoma; M = mesenchymal; MALT = mucosa-associated lymphoid tissue lymphoma; MDS = myelodysplastic syndrome; Mis = Missense; MLCLS = mediastinal large cell lymphoma with sclerosis; MM = multiple myeloma; MPD = myeloproliferative disorder; N = nonsense; NHL = non-Hodgkin’s lymphoma; NK/T = natural killer T cell; NSCLC = non-small cell lung cancer; O = other; PMBL = primary mediastinal B-cell lymphoma; pre-B ALL = pre-B-cell acute lymphoblastic leukemia; Rec = recessive; S = splice site; T = translocation; T-ALL = T-cell acute lymphoblastic leukemia; T-CLL = T-cell chronic lymphocytic leukemia; TGCT = testicular germ cell tumor; T-PLL = T-cell prolymphocytic leukemia.

BASIC CONSIDERATIONS

PART I

demonstrated that they have distinct functions.64,65 In fact, breast cancers arising from \( BRCA1 \) or \( BRCA2 \) mutations are different at the molecular level and have been found to have distinct gene expression profiles.66 \( BRCA1 \)-associated tumors are more likely to be estrogen receptor negative, whereas \( BRCA2 \)-associated tumors are more likely to be estrogen receptor positive. Currently, studies are ongoing to determine whether \( BRCA1 \) and \( BRCA2 \) status can be used to guide systemic therapy choices for breast cancer. Some targeted therapies are showing activity in \( BRCA \) mutation carriers with breast cancer such as PARP - poly(ADP-ribose) polymerase-inhibitors.

**APC Gene and Familial Adenomatous Polyposis**

Patients affected with familial adenomatous polyposis (FAP) characteristically develop hundreds to thousands of polyps in the colon and rectum. The polyps usually appear in adolescence and, if left untreated, progress to colorectal cancer. FAP is associated with benign extracolonic manifestations that may be useful in identifying new cases, including congenital hypertrophy of the retinal pigment epithelium, epidermoid cysts, and osteomas. In addition to colorectal cancer, patients with FAP are at risk for upper intestinal neoplasms (gastric and duodenal polyps, duodenal and periampullary cancer), hepatobiliary tumors (hepatoblastoma, pancreatic cancer, and cholangiocarcinoma), thyroid carcinomas, desmoid tumors, and medulloblastomas.

The product of the adenomatous polyposis coli tumor-suppressor gene (\( APC \)) plays an important role in cell-cell interactions, cell adhesion, regulation of \( \beta \)-catenin, and maintenance of cytoskeletal microtubules. Alterations in \( APC \) lead to dysregulation of several physiologic processes that govern colonic epithelial cell homeostasis, including cell-cycle progression, migration, differentiation, and apoptosis. Mutations in the \( APC \) have been identified in FAP and in 80% of sporadic colorectal cancers.67 Furthermore, \( APC \) mutations are the earliest known genetic alterations in colorectal cancer progression, which emphasizes its importance in cancer initiation. The germline mutations in \( APC \) may arise from point mutations, insertions, or deletions that lead to a premature stop codon and a truncated, functionally inactive protein. The risk of developing specific manifestations of FAP is correlated with the position of the \( APC \) mutations, a phenomenon referred to as genotype-phenotype correlation. For example, desmoids usually are associated with mutations between codons 1403 and 1578.68,69 Mutations in the extreme 5’ or 3’ ends of \( APC \), or in the alternatively spliced region of exon 9, are associated with an attenuated version of FAP. Better understanding of the genotype-phenotype correlations may assist in patient counseling and therapeutic planning.

**Mismatch Repair Genes and Hereditary Nonpolyposis Colorectal Cancer.** Hereditary nonpolyposis colorectal cancer (HNPCC), also referred to as Lynch syndrome, is an autosomal dominant hereditary cancer syndrome that predisposes to a wide spectrum of cancers, including colorectal cancer without polyposis. Some have proposed that HNPCC consists of at least two syndromes: Lynch syndrome 1, which entails hereditary nonpolyposis colorectal cancer, and Lynch syndrome 2, featuring a similar colonic phenotype accompanied by a high risk for carcinoma of the endometrium, transitional cell carcinoma of the ureter and renal pelvis, and carcinomas of the stomach, small bowel, ovary, and pancreas.70 The diagnostic criteria for HNPCC are referred to as the Amsterdam criteria, or the 3-2-1-0 rule. The classic Amsterdam criteria were revised to include other HNPCC-related cancers (Table 10-4).71 These criteria are met when three or more family members have histologically verified, HNPCC-associated cancers (one of whom is a first-degree relative of the other two), two or more generations are involved, at least one individual was diagnosed before age 50 years, and no individuals have FAP.71

During DNA replication, DNA polymerases may introduce single nucleotide mismatches or small insertion or deletion loops. These errors are corrected through a process referred to as mismatch repair. When mismatch repair genes are inactivated, DNA mutations in other genes that are critical to cell growth and proliferation accumulate rapidly. In HNPCC, germline mutations have been identified in several genes that play a key role in DNA nucleotide mismatch repair: \( hMLH1 \) (human mutL homologue 1), \( hMSH2 \) (human mutS homologue 2), \( hMSH6 \), \( hMSH4 \), and \( hMSH5 \).
and hPMS1 and hPMS2 (human postmeiotic segregation 1 and 2), of which hMLH1 and hMSH2 are the most common. The hallmark of HNPCC is microsatellite instability, which occurs on the basis of unpaired mismatches and small insertion or deletion loops. Microsatellite instability can be tested by comparing the DNA of a patient’s tumor with DNA from adjacent normal epithelium, amplifying the DNA with polymerase chain reaction (PCR) using a standard set of markers, comparing the amplified genomic DNA sequences, and classifying the degree of microsatellite instability as high, low, or stable. Such microsatellite instability testing may help select patients who are more likely to have germline mutations. An analysis of patients with early-onset colorectal cancer (age less than 50) showed that 8% had an unsuspected germline mutation in a mismatch repair gene and could be considered as having Lynch syndrome. Thus, genetic counseling and testing with a multigene panel should be considered for such patients.78

PTEN and Cowden Disease

Somatic deletions or mutations in the tumor-suppressor gene PTEN (phosphatase and tensin homologue deleted on chromosome 10) have been observed in a number of glioma breast, prostate, and renal carcinoma cell lines and several primary tumor specimens. PTEN encodes a 403-amino-acid protein, tyrosine phosphatase. PTEN negatively controls the PI3K signaling pathway for the regulation of cell growth and survival by dephosphorylating phosphoinositol 3,4,5-triphosphate; thus, mutation of PTEN leads to constitutive activation of the PI3K/Akt signaling pathway. The “hot spot” for PTEN mutations has been identified in exon 5. Forty-three percent of CD mutations have been identified in this exon, which contains the tyrosine phosphatase core domain. This suggests that the PTEN catalytic activity is vital for its biologic function. PTEN was identified as the susceptibility gene for the autosomal dominant syndrome Cowden disease (CD) or multiple hamartoma syndrome.80 Trichilemmomas, benign tumors of the hair follicle infundibulum, and mucocutaneous papillomatosis are pathognomonic of CD. Other common features include thyroid adenomas and multinodular goiters, breast fibroadenomas, and hamartomatous GI polyps. The diagnosis of CD is made when an individual or family has a combination of pathognomonic major and/or minor criteria proposed by the International Cowden Consortium.81 CD is associated with an increased risk of breast and thyroid cancers. Breast cancer develops in 25% to 50% of affected women.81

p16 and Hereditary Malignant Melanoma. The gene p16, also known as INK4A, CDKN1, CDKN2A, and MTS1, is a tumor suppressor that acts by binding CDK4 and CDK6 and inhibiting the catalytic activity of the CDK4-CDK6/cyclin D complex that is required for phosphorylation of Rb and subsequent cell-cycle progression. Studies suggest that germline mutations in p16 can be found in 20% of melanoma-prone families.82 Mutations in p16 that alter its ability to inhibit the catalytic activity of the CDK4-CDK6/cyclin D complex not only increase the risk of melanoma by 75-fold but also increase the risk of pancreatic cancer by 22-fold.83 Interestingly, p16 mutations that do not appear to alter its function increase the risk of melanoma by 38-fold and do not increase the risk of pancreatic cancer.83 Genomic characterization of primary tumors has revealed that p16 is inactivated through point mutation, promoter methylation, or deletion in a significant portion of sporadic tumors, including cancers of the pancreas, esophagus, head and neck, stomach, breast, and colon, as well as melanomas.

E-cadherin and Hereditary Diffuse Gastric Cancer. E-cadherin is a cell adhesion molecule that plays an important role in normal architecture and function of epithelial cells. The adhesive function of E-cadherin is dependent on interaction of its cytoplasmic domain with β- and γ-catenins and may be regulated by phosphorylation of β-catenin.

Hereditary diffuse gastric carcinoma is an autosomal dominant cancer syndrome that results from germline mutations in the E-cadherin gene, CDH1. Carriers of CDH1 mutations have a 70% to 80% chance of developing gastric cancer.84 Furthermore, mutations of CDH1 have been described in sporadic cancers of the ovary, endometrium, breast, and thyroid. However, frequent mutations have been identified in only two particular tumors: diffuse gastric carcinomas and lobular breast carcinomas. Invasive lobular breast carcinomas often show inactivating mutations in combination with a loss of heterozygosity of the wild-type CDH1 allele.85 Interestingly, in gastric carcinomas the predominant mutations are exon skipping causing in-frame deletions, whereas most mutations identified in lobular breast cancers are premature stop codons; this suggests a genotype-phenotype correlation.

RET Proto-Oncogene and Multiple Endocrine Neoplasia Type 2

The RET (rearranged during transfection) gene encodes for a transmembrane receptor tyrosine kinase that plays a role in proliferation, migration, and differentiation of cells derived from the neural crest. Gain-of-function mutations in the RET gene are associated with medullary thyroid carcinoma in isolation or multiple endocrine neoplasia type 2 (MEN2) syndromes. MEN2A is associated with medullary thyroid carcinoma in isolation or multiple endocrine neoplasia type 2 (MEN2) syndromes. MEN2A is associated with medullary thyroid carcinoma and pheochromocytoma (in 50%) or parathyroid adenoma (in 20%), whereas MEN2B is associated with medullary thyroid carcinoma, marfanoid habitus, mucosal neuromas, and ganglioneuromatosis of the gastrointestinal tract.86 RET mutations lead to uncontrolled growth of the thyroid C cells, and in familial medullary cancer, C-cell hyperplasia progresses to bilateral, multicentric medullary thyroid cancer. Mutations in the RET gene have also been identified in half of sporadic medullary thyroid cancers.

Genetic Modifiers of Risk. Individuals carrying identical germline mutations vary in regard to cancer penetrance (whether cancer will develop or not) and cancer phenotype (the tissues involved). It is thought that this variability may be due to environmental influences or, if genetic, to genetic modifiers...
of risk. Similarly, genetic modifiers of risk also can play a role in determining whether an individual will develop cancer after exposure to carcinogens.

**Chemical Carcinogens**

The first report indicating that cancer could be caused by environmental factors was by John Hill, who in 1761 noted the association between nasal cancer and excessive use of tobacco snuff. Currently, approximately 60% to 90% of cancers are thought to be due to environmental factors. Any agent that can contribute to tumor formation is referred to as a carcinogen and can be a chemical, physical, or viral agent. Chemicals are classified into three groups based on how they contribute to tumor formation. The first group of chemical agents, the genotoxins, can initiate carcinogenesis by causing a mutation. The second group, the cocarcinogens, by themselves cannot cause cancer but potentiate carcinogenesis by enhancing the potency of genotoxins. The third group, tumor promoters, enhances tumor formation when given after exposure to genotoxins.

The International Agency for Research on Cancer (IARC) maintains a registry of human carcinogens that is available through the World Wide Web (http://www.iarc.fr). The compounds are categorized into five groups based on an analysis of epidemiologic studies, animal models, and short-term mutagenesis tests. Group 1 contains what are considered to be proven human carcinogens, based on formal epidemiologic studies among workers who were exposed for long periods (several years) to the chemicals. Group 2A contains what are considered to be probable human carcinogens. Suggestive epidemiologic evidence exists for compounds in this group, but the data are insufficient to establish causality. There is evidence of carcinogenicity, however, from animal studies carried out under conditions relevant to human exposure. Group 2B contains what are considered to be possible carcinogens because these substances are associated with a clear statistically and biologically significant increase in the incidence of malignant tumors in more than one animal species or strain. Group 3 agents are not classifiable, and Group 4 agents are probably not carcinogenic to humans. Selected substances that have been classified as proven carcinogens (group 1) by the IARC in an expert panel review are listed in Table 10-5.

**Physical Carcinogens**

Physical carcinogenesis can occur through induction of inflammation and cell proliferation over a period of time or through exposure to physical agents that induce DNA damage. Foreign bodies can cause chronic irritation that can expose cells to carcinogenesis due to other environmental agents. In animal models, for example, subcutaneous implantation of a foreign body can lead to the development of tumors that have been attributed to chronic irritation from the foreign objects. In humans, clinical scenarios associated with chronic irritation and inflammation such as chronic nonhealing wounds, burns, and inflammatory bowel syndrome have all been associated with an increased risk of cancer. *H. pylori* infection is associated with gastritis and gastric cancer, and thus, its carcinogenicity may be considered physical carcinogenesis. Infection with the liver fluke *Opisthorchis viverrini* similarly leads to local inflammation and cholangiocarcinoma.

The induction of lung and mesothelial cancers by asbestos fibers and nonfibrous particles such as silica are other examples of foreign body-induced physical carcinogenesis. Animal experiments have demonstrated that the dimensions and durability of the asbestos and other fibrous minerals are the key determinants of their carcinogenicity. Short fibers can be inactivated by phagocytosis, whereas long fibers (>10 μm) are cleared less effectively and are encompassed by proliferating epithelial cells. The long fibers support cell proliferation and have been shown to preferentially induce tumors. Asbestos-associated biologic effects also may be mediated through reactive oxygen and nitrogen species. Furthermore, an interaction occurs between asbestos and silica and components of cigarette smoke. Polycyclic aromatic hydrocarbons (PAHs) in cigarette smoke are metabolized by epithelial cells and form DNA adducts. If PAH is coated on asbestos, PAH uptake is increased. Both PAH and asbestos impair lung clearance, potentially increasing uptake further. Therefore, physical carcinogens may be synergistic with chemical carcinogens.

Radiation is the best-known agent of physical carcinogenesis and is classified as ionizing radiation (X-rays, gamma rays, and alpha and beta particles) or nonionizing radiation (UV). The carcinogenic potential of ionizing radiation was recognized soon after Wilhelm Conrad Roentgen’s discovery of X-rays in 1895. Within the next 20 years, a large number of radiation-related skin cancers were reported. Long-term follow-up of survivors of the atomic bombing of Hiroshima and Nagasaki revealed that virtually all tissues exposed to radiation are at risk for cancer. Radiation can induce a spectrum of DNA lesions that includes damage to the nucleotide bases and cross-linking, and DNA single- and double-strand breaks (DSBs). Misrepaired DSBs are the principal lesions of importance in the induction of chromosomal abnormalities and gene mutations. DSBs in irradiated cells are repaired primarily by a nonhomologous end-joining process, which is error prone; thus, DSBs facilitate the production of chromosomal rearrangements and other large-scale changes such as chromosomal deletions. It is thought that radiation may initiate cancer by inactivating tumor-suppressor genes. Activation of oncogenes appears to play a lesser role in radiation carcinogenesis.

Although it has been assumed that the initial genetic events induced by radiation constitute direct mutagenesis from radiation, other indirect effects may contribute to carcinogenesis. For example, radiation induces genomic instability in cells that persists for at least 30 generations after irradiation. Therefore, even if cells do not acquire mutations at initial irradiation, they remain at risk for developing new mutations for several generations. Moreover, even cells that have not been directly irradiated appear to be at risk, a phenomenon referred to as the bystander effect.

Nonionizing UV radiation is a potent DNA-damaging agent and is known to induce skin cancer in experimental animals. Most nonmelanoma human skin cancers are thought to be induced by repeated exposure to sunlight, which leads to a series of mutations that allow the cells to escape normal growth control. Patients with inherited xeroderma pigmentosum lack one or more DNA repair pathways, which confers susceptibility to UV-induced cancers, especially on sun-exposed body parts. Patients with ataxia telangiectasia mutated syndrome also have a radiation-sensitive phenotype.

**Viral Carcinogens**

One of the first observations that cancer may be caused by transmissible agents was by Peyton Rous in 1910 when he demonstrated that cell-free extracts from sarcomas in chickens could transmit sarcomas to other animals injected with these extracts. This was subsequently discovered to represent viral transmission of cancer by the Rous sarcoma virus. At present, several human viruses are known to have oncogenic properties, and several
### Group 1 chemical carcinogens and evidence for carcinogenicity in humans and for genotoxicity as the main mechanism

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Tumor sites or types with sufficient evidence in humans</th>
<th>Evidence of genotoxicity as the main mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-Aminobiphenyl</td>
<td>Urinary bladder</td>
<td>Strong</td>
</tr>
<tr>
<td>Benzidine</td>
<td>Urinary bladder</td>
<td>Strong</td>
</tr>
<tr>
<td>Dyes metabolized to benzidine</td>
<td></td>
<td>Strong^a</td>
</tr>
<tr>
<td>4,4’-Methylenebis(2-chloroaniline)</td>
<td></td>
<td>Strong^a</td>
</tr>
<tr>
<td>2-Naphthylamine</td>
<td>Urinary bladder</td>
<td>Strong</td>
</tr>
<tr>
<td>Ortho-toluidine</td>
<td>Urinary bladder</td>
<td>Moderate</td>
</tr>
<tr>
<td>Auramine production</td>
<td>Urinary bladder</td>
<td>Weak/lack of data^b</td>
</tr>
<tr>
<td>Magenta production</td>
<td>Urinary bladder</td>
<td>Weak/lack of data^b</td>
</tr>
<tr>
<td>Benzo[a]pyrene</td>
<td>...</td>
<td>Strong^a</td>
</tr>
<tr>
<td>Soot (chimney sweeping)</td>
<td>Skin, lung</td>
<td>Moderate</td>
</tr>
<tr>
<td>Coal gasification</td>
<td>Lung</td>
<td>Strong</td>
</tr>
<tr>
<td>Coal-tar distillation</td>
<td>Skin</td>
<td>Strong</td>
</tr>
<tr>
<td>Coke production</td>
<td>Lung</td>
<td>Strong</td>
</tr>
<tr>
<td>Coal-tar pitches (paving, roofing)</td>
<td>Lung</td>
<td>Strong</td>
</tr>
<tr>
<td>Aluminum production</td>
<td>Lung, urinary bladder</td>
<td>Weak/moderate^b,c</td>
</tr>
<tr>
<td>Aflatoxins</td>
<td>Hepatocellular carcinoma</td>
<td>Strong</td>
</tr>
<tr>
<td>Benzene</td>
<td>ANLL</td>
<td>Strong</td>
</tr>
<tr>
<td>Bis (chloromethyl)ether/chloromethyl methyl ether</td>
<td>Lung</td>
<td>Moderate/strong</td>
</tr>
<tr>
<td>1,3-Butadiene</td>
<td>Haematolymphatic organs</td>
<td>Strong</td>
</tr>
<tr>
<td>Dioxin (2,3,7,8-TCDD)</td>
<td>All cancers combined^a</td>
<td>See text^d</td>
</tr>
<tr>
<td>2,3,4,7,8-Pentachlorodibenzofuran</td>
<td>...</td>
<td>See text^a,d</td>
</tr>
<tr>
<td>3,3’,4,4’,5-Pentachlorobiphenyl (PCB-126)</td>
<td>...</td>
<td>See text^a,d</td>
</tr>
<tr>
<td>Ethylene oxide</td>
<td>...</td>
<td>Strong^a</td>
</tr>
<tr>
<td>Formaldehyde</td>
<td>Nasopharynx, Leukemia^e</td>
<td>Strong, Moderate</td>
</tr>
<tr>
<td>Sulfur mustard</td>
<td>Lung</td>
<td>Strong</td>
</tr>
<tr>
<td>Vinyl chloride</td>
<td>Hepatic angiosarcoma, hepatocellular carcinoma</td>
<td>Strong</td>
</tr>
<tr>
<td>Iron and steel founding</td>
<td>Lung</td>
<td>Weak/moderate</td>
</tr>
<tr>
<td>Isopropyl alcohol manufacture using strong acids</td>
<td>Nasal cavity</td>
<td>Weak/lack of data</td>
</tr>
<tr>
<td>Mineral oils</td>
<td>Skin</td>
<td>Weak/lack of data</td>
</tr>
<tr>
<td>Occupational exposure as a painter</td>
<td>Lung, urinary bladder, pleural mesothelioma</td>
<td>Strong^c</td>
</tr>
<tr>
<td>Rubber-manufacturing industry</td>
<td>Leukaemia, lymphoma,^e urinary bladder,^e lung,^e stomach^e</td>
<td>Strong^e</td>
</tr>
<tr>
<td>Shale oils</td>
<td>Skin</td>
<td>Weak/lack of data</td>
</tr>
<tr>
<td>Strong inorganic acid mists</td>
<td>Larynx</td>
<td>Weak/lack of data</td>
</tr>
</tbody>
</table>

ANLL = acute nonlymphocytic leukemia; ALL = acute lymphocytic leukemia; CLL = chronic lymphocytic leukemia; MM = multiple myeloma; NH = non-Hodgkin lymphoma; STS = soft-tissue sarcoma.

^aAgents classified in Group 1 on the basis of mechanistic information.

^bWeak evidence in workers, but strong evidence for some chemicals in this industry.

^cDue to the diversity and complexity of these exposures, other mechanisms may also be relevant.

^dStrong evidence for an aryl hydrocarbon receptor (AhR)-mediated mechanism.

^eParticularly myeloid leukemia.

^fAfter maternal exposure (before or during pregnancy, or both).

^gNew epidemiological findings.

**Table 10-6**

Selected viral carcinogens

<table>
<thead>
<tr>
<th>VIRUS</th>
<th>PREDOMINANT TUMOR TYPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epstein-Barr virus</td>
<td>Burkitt’s lymphoma</td>
</tr>
<tr>
<td></td>
<td>Hodgkin’s disease</td>
</tr>
<tr>
<td></td>
<td>Immunosuppression-related</td>
</tr>
<tr>
<td></td>
<td>lymphoma</td>
</tr>
<tr>
<td></td>
<td>Sinonasal angiocentric T-cell</td>
</tr>
<tr>
<td></td>
<td>lymphoma</td>
</tr>
<tr>
<td></td>
<td>Nasopharyngeal carcinoma</td>
</tr>
<tr>
<td>Hepatitis B virus</td>
<td>Hepatocellular carcinoma</td>
</tr>
<tr>
<td>Hepatitis C virus</td>
<td>Hepatocellular carcinoma</td>
</tr>
<tr>
<td>HIV type 1</td>
<td>Kaposi’s sarcoma</td>
</tr>
<tr>
<td></td>
<td>Cervical cancer</td>
</tr>
<tr>
<td></td>
<td>Non-Hodgkin’s lymphoma</td>
</tr>
<tr>
<td>Human herpes virus 8</td>
<td>Kaposi’s sarcoma</td>
</tr>
<tr>
<td>Human papillomavirus</td>
<td>Cervical cancer</td>
</tr>
<tr>
<td>16 and 18</td>
<td>Vulvar and vaginal cancer</td>
</tr>
<tr>
<td></td>
<td>Penile cancer</td>
</tr>
<tr>
<td></td>
<td>Oropharyngeal cancer (commonly</td>
</tr>
<tr>
<td></td>
<td>base of tongue and tonsil)</td>
</tr>
<tr>
<td></td>
<td>Anal cancer</td>
</tr>
<tr>
<td>Human T-cell</td>
<td>Adult T-cell leukemia/lymphoma</td>
</tr>
<tr>
<td>lymphotropic viruses</td>
<td></td>
</tr>
<tr>
<td>Merkel cell polyoma</td>
<td>Merkel cell carcinoma</td>
</tr>
<tr>
<td>virus</td>
<td></td>
</tr>
</tbody>
</table>

*Data based on information in the International Agency for Research on Cancer monographs.

*Only tumor types for which causal relationships are established are listed. Other cancer types may be linked to the agents with a lower frequency or with insufficient data to prove causality.*

have been causally linked to human cancers (Table 10-6). It is estimated that 15% of all human tumors worldwide are caused by viruses.

Viruses may cause or increase the risk of malignancy through several mechanisms, including direct transformation, expression of oncogenes that interfere with cell-cycle checkpoints or DNA repair, expression of cytokines or other growth factors, and alteration of the immune system. Oncogenic viruses may be RNA or DNA viruses. Oncogenic RNA viruses are retroviruses and contain a reverse transcriptase. After the viral infection, the single-stranded RNA viral genome is transcribed into a double-stranded DNA copy, which is then integrated into the chromosomal DNA of the cell. Retroviral infection of the cell is permanent; thus, integrated DNA sequences remain in the host chromosome. Oncogenic transforming retroviruses carry oncogenes derived from cellular genes. These cellular genes, referred to as *proto-oncogenes*, usually are involved in mitogenic signaling and growth control, and include protein kinases, G proteins, growth factors, and transcription factors (Table 10-7).

Integration of the provirus upstream of a proto-oncogene may produce chimeric virus-cell transcripts and recombination during the next round of replication that could lead to incorporation of the cellular gene into the viral genome. Then again, many retroviruses do not possess oncogenes but can cause tumors in animals regardless. This occurs by integration of the provirus near a normal cellular proto-oncogene and activation of the expression of these genes by the strong promoter and enhancer sequences in the integrated viral sequence.

Unlike the oncogenes of the RNA viruses, those of the DNA tumor viruses are viral, not cellular, in origin. These genes are required for viral replication using the host cell machinery. In permissive hosts, infection with an oncogenic DNA virus may result in a productive lytic infection, which leads to cell death and the release of newly formed viruses. In nonpermissive cells, the viral DNA can be integrated into the cellular chromosomal DNA, and some of the early viral genes can be synthesized persistently, which leads to transformation of cells to a neoplastic state. The binding of viral oncoproteins to cellular tumor-suppressor proteins p53 and Rb is fundamental to the carcinogenesis induced by most DNA viruses, although some target different cellular proteins.

Like other types of carcinogenesis, viral carcinogenesis is a multistep process. Some retroviruses contain two cellular oncogenes, rather than one, in their genome and are more rapidly tumorigenic than single-gene transforming retroviruses, which emphasizes the cooperation between transforming genes. Furthermore, some viruses encode genes that suppress or delay apoptosis.

Although immunocompromised individuals are at elevated risk, most patients infected with oncogenic viruses do not develop cancer. When cancer does develop, it usually occurs several years after the viral infection. It is estimated, for example, that the risk of hepatocellular carcinoma (HCC) among individuals infected with hepatitis C virus is 1% to 3% after 30 years. There may be synergy between various environmental factors and viruses in carcinogenesis.

Recognition of a viral origin for some tumors has led to the pursuit of vaccination as a preventive strategy. The use of childhood hepatitis B vaccination has already translated into a decrease in liver cancer incidence in the East Asia. Similarly, it is recognized that cervical cancer and its obligate precursors, cervical intraepithelial neoplasia grades 2 and 3, and adenocarcinoma in situ, are caused by oncogenic human papillomavirus (HPV); administration of HPV vaccine to HPV-naive women, substantially reduces the incidence of HPV16/18-related cervical precancers and cervical cancer. Studies suggest that HPV vaccination may also reduce oral HPV infections that are a risk factor for the development of oropharyngeal cancer. The American Cancer Society recommends routine HPV vaccination for girls and boys starting at age 11 or 12. The vaccination series can be started as early as age 9. HPV vaccination is also recommended for females 13 to 26 years old and for males 13 to 21 years old who have not started the vaccines, or who have started but have not completed the series. Males 22 to 26 years old may also be vaccinated. HPV vaccination is also recommended up until age 26 for men who have sex with men and for people with weakened immune systems (including people with HIV infection), if they have not previously been vaccinated. It is important for patients to know that vaccination at older ages is less effective in lowering cancer risk.

**CANCER RISK ASSESSMENT**

Cancer risk assessment is an important part of the initial evaluation of any patient. A patient’s cancer risk not only is an important determinant of cancer screening recommendations but also
**Table 10-7**

Selected cellular oncogenes in retroviruses

<table>
<thead>
<tr>
<th>ONCOGENE</th>
<th>VIRUS NAME</th>
<th>ORIGIN</th>
<th>PROTEIN PRODUCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>abl</td>
<td>Abelson murine leukemia virus</td>
<td>Mouse</td>
<td>Tyrosine kinase</td>
</tr>
<tr>
<td>fes</td>
<td>ST feline sarcoma virus</td>
<td>Cat</td>
<td>Tyrosine kinase</td>
</tr>
<tr>
<td>fps</td>
<td>Fujinami sarcoma virus</td>
<td>Chicken</td>
<td>Tyrosine kinase</td>
</tr>
<tr>
<td>src</td>
<td>Rous sarcoma virus</td>
<td>Chicken</td>
<td>Tyrosine kinase</td>
</tr>
<tr>
<td>erbB</td>
<td>Avian erythroblastosis virus</td>
<td>Chicken</td>
<td>Epidermal growth factor receptor</td>
</tr>
<tr>
<td>fms</td>
<td>McDonough feline sarcoma virus</td>
<td>Cat</td>
<td>Colony-stimulating factor receptor</td>
</tr>
<tr>
<td>kit</td>
<td>Hardy-Zuckerman 4 feline sarcoma virus</td>
<td>Cat</td>
<td>Stem cell factor receptor</td>
</tr>
<tr>
<td>mil</td>
<td>Avian myelocytoma virus</td>
<td>Chicken</td>
<td>Serine/threonine kinase</td>
</tr>
<tr>
<td>mos</td>
<td>Moloney murine sarcoma virus</td>
<td>Mouse</td>
<td>Serine/threonine kinase</td>
</tr>
<tr>
<td>raf</td>
<td>Murine sarcoma virus 3611</td>
<td>Mouse</td>
<td>Serine/threonine kinase</td>
</tr>
<tr>
<td>sis</td>
<td>Simian sarcoma virus</td>
<td>Monkey</td>
<td>Platelet-derived growth factor</td>
</tr>
<tr>
<td>H-ras</td>
<td>Harvey murine sarcoma virus</td>
<td>Rat</td>
<td>GDP/GTP binding</td>
</tr>
<tr>
<td>K-ras</td>
<td>Kirsten murine sarcoma virus</td>
<td>Rat</td>
<td>GDP/GTP binding</td>
</tr>
<tr>
<td>erbA</td>
<td>Avian erythroblastosis virus</td>
<td>Chicken</td>
<td>Transcription factor (thyroid hormone receptor)</td>
</tr>
<tr>
<td>ets</td>
<td>Avian myeloblastosis virus E26</td>
<td>Chicken</td>
<td>Transcription factor</td>
</tr>
<tr>
<td>fes</td>
<td>FBJ osteosarcoma virus</td>
<td>Mouse</td>
<td>Transcription factor (AP1 component)</td>
</tr>
<tr>
<td>jun</td>
<td>Avian sarcoma virus 17</td>
<td>Chicken</td>
<td>Transcription factor (AP1 component)</td>
</tr>
<tr>
<td>myb</td>
<td>Avian myeloblastosis virus</td>
<td>Chicken</td>
<td>Transcription factor</td>
</tr>
<tr>
<td>myc</td>
<td>MC29 myelocytoma virus</td>
<td>Chicken</td>
<td>Transcription factor (NF-kB family)</td>
</tr>
</tbody>
</table>

AP1 = activator protein 1; FBJ = Finkel-Biskis-Jinkins; GDP = guanosine diphosphate; GTP = guanosine triphosphate; NF-κB = nuclear factor κB.


may alter how aggressively an indeterminant finding will be pursued for diagnosis. A “probably benign” mammographic lesion, for example, defined as one with <2% probability of malignancy (American College of Radiology category III) is usually managed with a 6-month follow-up mammogram in a patient at baseline cancer risk, but obtaining a tissue diagnosis may be preferable in a patient at high risk for breast cancer.98

Cancer risk assessment starts with taking a complete history that includes history of environmental exposures to potential carcinogens and a detailed family history. Risk assessment for breast cancer, for example, includes obtaining a family history to determine whether another member of the family is known to carry a breast cancer susceptibility gene; whether there is familial clustering of breast cancer, ovarian cancer, thyroid cancer, sarcoma, adenocortical carcinoma, endometrial cancer, brain tumors, dermatologic manifestations, leukemia, or lymphoma; and whether the patient is from a population at increased risk, such as individuals of Ashkenazi Jewish descent. Patients who have a family history suggestive of a cancer susceptibility syndrome such as hereditary breast-ovarian syndrome, Li-Fraumeni Syndrome, or Cowden’s Disease would benefit from genetic counseling and possibly genetic testing.

There are several models that can estimate risk based on complex family histories and assist clinicians in estimating breast cancer risk or the likelihood that a BRCA mutation is present, including the Claus model, Tyrer-Cuzick model, BRCAPRO model, and the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA) model.99-102 Patients who do have a strong hereditary component of risk can be evaluated on the basis of their age, race, personal history, and exposures. One of the most commonly used models for risk assessment in breast cancer is the Gail model.103 Gail and colleagues analyzed the data from 2852 breast cancer cases and 3146 controls from the Breast Cancer Detection and Demonstration Project, a mammography screening project conducted in the 1970s, and developed a model for projecting breast cancer incidence. The model uses risk factors such as an individual’s age, age at menarche, age at first live birth, number of first-degree relatives with breast cancer, number of previous breast biopsy specimens, and whether the biopsy specimen results revealed atypical ductal hyperplasia (Table 10-8).103 This model has led to the development of a breast cancer risk assessment tool, which is available on the World Wide Web.104 This tool incorporates the risk factors used in the Gail model, as well as race and ethnicity, and allows a health professional to project a woman’s individualized estimated risk for invasive breast cancer over a 5-year period and over her lifetime (to age 90 years). Notably, these risk projections assume that the woman is undergoing regular clinical breast examinations and screening mammograms. Also of note is that this program underestimates the risk for women who have already had a diagnosis of invasive or noninvasive breast cancer and does not take into account specific genetic predispositions such as mutations in BRCA1 or BRCA2. However, risk assessment tools such as this have been validated and are now in widespread clinical use. Similar models are in development or are being validated for...
Table 10-8
Assessment of risk for invasive breast cancer

<table>
<thead>
<tr>
<th>RISK FACTOR</th>
<th>RELATIVE RISK (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at menarche (years)</td>
<td></td>
</tr>
<tr>
<td>&gt;14</td>
<td>1.00</td>
</tr>
<tr>
<td>12–13</td>
<td>1.10</td>
</tr>
<tr>
<td>&lt;12</td>
<td>1.21</td>
</tr>
<tr>
<td>Age at first live birth (years)</td>
<td></td>
</tr>
<tr>
<td>Patients with no first-degree relatives with cancer</td>
<td></td>
</tr>
<tr>
<td>&lt;20</td>
<td>1.00</td>
</tr>
<tr>
<td>20–24</td>
<td>1.24</td>
</tr>
<tr>
<td>25–29 or nulliparous</td>
<td>1.55</td>
</tr>
<tr>
<td>≥30</td>
<td>1.93</td>
</tr>
<tr>
<td>Patients with one first degree-relative with cancer</td>
<td></td>
</tr>
<tr>
<td>&lt;20</td>
<td>1.00</td>
</tr>
<tr>
<td>20–24</td>
<td>2.64</td>
</tr>
<tr>
<td>25–29 or nulliparous</td>
<td>2.76</td>
</tr>
<tr>
<td>≥30</td>
<td>2.83</td>
</tr>
<tr>
<td>Patients with ≥2 first-degree relatives with cancer</td>
<td></td>
</tr>
<tr>
<td>&lt;20</td>
<td>6.80</td>
</tr>
<tr>
<td>20–24</td>
<td>5.78</td>
</tr>
<tr>
<td>25–29 or nulliparous</td>
<td>4.91</td>
</tr>
<tr>
<td>≥30</td>
<td>4.17</td>
</tr>
<tr>
<td>Breast biopsies (number)</td>
<td></td>
</tr>
<tr>
<td>Patients aged &lt;50 y at counseling</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.00</td>
</tr>
<tr>
<td>1</td>
<td>1.70</td>
</tr>
<tr>
<td>≥2</td>
<td>2.88</td>
</tr>
<tr>
<td>Patients aged ≥50 y at counseling</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.00</td>
</tr>
<tr>
<td>1</td>
<td>1.27</td>
</tr>
<tr>
<td>≥2</td>
<td>1.62</td>
</tr>
<tr>
<td>Atypical hyperplasia</td>
<td></td>
</tr>
<tr>
<td>No biopsies</td>
<td>1.00</td>
</tr>
<tr>
<td>At least 1 biopsy, no atypical hyperplasia</td>
<td>0.93</td>
</tr>
<tr>
<td>No atypical hyperplasia, hyperplasia status unknown for at least 1 biopsy</td>
<td>1.00</td>
</tr>
<tr>
<td>Atypical hyperplasia in at least 1 biopsy</td>
<td>1.82</td>
</tr>
</tbody>
</table>


Early detection is the key to success in cancer therapy. Screening for common cancers using relatively noninvasive tests is expected to lead to early diagnosis, allow more conservative surgical therapies with decreased morbidity, and potentially improve surgical cure rates and overall survival rates. Key factors that influence screening guidelines are how prevalent the cancer is in the population, what risk is associated with the screening measure, and whether early diagnosis actually affects outcome. The value of a widespread screening measure is likely to go up with the prevalence of the cancer in a population, which often determines the age cutoffs for screening and explains why screening is done only for common cancers. The risks associated with the screening measure are a significant consideration, especially with more invasive screening measures such as colonoscopy. The consequences of a false-positive screening test result also need to be considered. For example, when 1000 screening mammograms are taken, only 2 to 4 new cases of cancer will be identified; this number is slightly higher (6 to 10 prevalent cancers per 1000 mammograms) for initial screening mammograms. However, as many as 10% of screening mammograms may be potentially suggestive of an abnormality, which requires further imaging (i.e., a 10% recall rate). Of those women with abnormal mammogram findings, only 5% to 10% will be determined to have a breast cancer. Among women for whom biopsy specimen is recommended, 25% to 40% will have a breast cancer. A false-positive screening result is likely to induce significant emotional distress in patients, leads to unnecessary biopsy specimens, and has cost implications for the health care system.

American Cancer Society guidelines for the early detection of cancer are listed in Table 10-9. These guidelines are updated periodically to incorporate emerging technologies and new data on the efficacy of screening measures. Besides the American Cancer Society, several other professional bodies make recommendations for screening. Although the screening guidelines differ somewhat, most organizations do not emphasize one screening strategy as superior to another, but all emphasize the importance of age-appropriate screening.

Screening guidelines are developed for the general baseline-risk population. These guidelines need to be modified for patients who are at high risk. For example, more intensive colorectal cancer screening is recommended for individuals at increased risk because of a history of adenomatous polyps, a personal history of colorectal cancer, a family history of either colorectal cancer or colorectal adenomas diagnosed in a first-degree relative before age 60 years, a personal history of inflammatory bowel disease of significant duration, or a family history or genetic test result indicating FAP or HNPCC. For some diseases, in higher risk populations, both the screening modality and the screening intensity may be altered. For example, breast magnetic resonance imaging is recommended as an adjunct to

other cancers. For example, a lung cancer risk prediction model, which includes age, sex, asbestos exposure history, and smoking history, has been found to predict risk of lung cancer. There is now growing interest in using each individuals genotype, such as presence or absence of single nucleotide polymorphisms which each may confer low or intermediate cancer risk. Risk models that include biologic as well as environmental factors may accurately predict cancer risk, providing better guidance as to which patients should undergo more intensive screening (e.g., screening with magnetic resonance imaging of the breast, computerized tomography screening of the lung), and should be considered for preventive strategies.

CANCER SCREENING

Early detection is the key to success in cancer therapy. Screening for common cancers using relatively noninvasive tests is expected to lead to early diagnosis, allow more conservative surgical therapies with decreased morbidity, and potentially improve surgical cure rates and overall survival rates. Key factors that influence screening guidelines are how prevalent the cancer is in the population, what risk is associated with the screening measure, and whether early diagnosis actually affects outcome. The value of a widespread screening measure is likely to go up with the prevalence of the cancer in a population, which often determines the age cutoffs for screening and explains why screening is done only for common cancers. The risks associated with the screening measure are a significant consideration, especially with more invasive screening measures such as colonoscopy. The consequences of a false-positive screening test result also need to be considered. For example, when 1000 screening mammograms are taken, only 2 to 4 new cases of cancer will be identified; this number is slightly higher (6 to 10 prevalent cancers per 1000 mammograms) for initial screening mammograms. However, as many as 10% of screening mammograms may be potentially suggestive of an abnormality, which requires further imaging (i.e., a 10% recall rate). Of those women with abnormal mammogram findings, only 5% to 10% will be determined to have a breast cancer. Among women for whom biopsy specimen is recommended, 25% to 40% will have a breast cancer. A false-positive screening result is likely to induce significant emotional distress in patients, leads to unnecessary biopsy specimens, and has cost implications for the health care system.

American Cancer Society guidelines for the early detection of cancer are listed in Table 10-9. These guidelines are updated periodically to incorporate emerging technologies and new data on the efficacy of screening measures. Besides the American Cancer Society, several other professional bodies make recommendations for screening. Although the screening guidelines differ somewhat, most organizations do not emphasize one screening strategy as superior to another, but all emphasize the importance of age-appropriate screening.

Screening guidelines are developed for the general baseline-risk population. These guidelines need to be modified for patients who are at high risk. For example, more intensive colorectal cancer screening is recommended for individuals at increased risk because of a history of adenomatous polyps, a personal history of colorectal cancer, a family history of either colorectal cancer or colorectal adenomas diagnosed in a first-degree relative before age 60 years, a personal history of inflammatory bowel disease of significant duration, or a family history or genetic test result indicating FAP or HNPCC. For some diseases, in higher risk populations, both the screening modality and the screening intensity may be altered. For example, breast magnetic resonance imaging is recommended as an adjunct to

other cancers. For example, a lung cancer risk prediction model, which includes age, sex, asbestos exposure history, and smoking history, has been found to predict risk of lung cancer. There is now growing interest in using each individuals genotype, such as presence or absence of single nucleotide polymorphisms which each may confer low or intermediate cancer risk. Risk models that include biologic as well as environmental factors may accurately predict cancer risk, providing better guidance as to which patients should undergo more intensive screening (e.g., screening with magnetic resonance imaging of the breast, computerized tomography screening of the lung), and should be considered for preventive strategies.
<table>
<thead>
<tr>
<th>CANCER SITE</th>
<th>POPULATION</th>
<th>TEST OR PROCEDURE</th>
<th>FREQUENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast</td>
<td>Women age 40 and above</td>
<td>Mammography</td>
<td>Women with an average risk of breast cancer should undergo regular screening mammography starting at age 45 years. Women age 45 to 54 years should be screened annually. Women 55 years and older should transition to biennial screening or have the opportunity to continue screening annually. Women should continue screening mammography as long as their overall health is good and they have a life expectancy of 10 years or longer. The ACS does not recommend clinical breast examination for breast cancer screening among average-risk women at any age.</td>
</tr>
<tr>
<td>Cervix</td>
<td>Women, age 21–65 y</td>
<td>Pap test and HPV DNA test</td>
<td>Cervical cancer screening should begin at age 21 y. For women age 21–29 y, screening should be done every 3 y with conventional or liquid-based Pap tests. For women age 30–65 y, screening should be done every 5 y with both the HPV test and the Pap test (preferred), or every 3 y with the Pap test alone (acceptable). Women age &gt;65 y who have had ≥3 consecutive negative Pap tests or ≥2 consecutive negative HPV and Pap tests within the last 10 y, with the most recent test occurring within the last 5 y, and women who have had a total hysterectomy should stop cervical cancer screening.</td>
</tr>
<tr>
<td>Colorectal</td>
<td>Men and women age ≥50 y</td>
<td>gFOBT, or FIT, or sDNA with a high sensitivity for cancer</td>
<td>Annual, starting at age 50 y.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FSIG, or</td>
<td>Every 5 y, starting at age 50 y. FSIG can be performed alone, or consideration can be given to combining FSIG performed every 5 y with a highly sensitive guaiac-based FOBT or FIT performed annually.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DCBE, or</td>
<td>Every 5 y, starting at age 50 y.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Colonoscopy</td>
<td>Every 10 y, starting at age 50 y.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CT colonography</td>
<td>Every 5 yr, starting at age 50 y.</td>
</tr>
<tr>
<td>Endometrial</td>
<td>Women, at menopause</td>
<td></td>
<td>At the time of menopause, women at average risk should be informed about the risks and symptoms of endometrial cancer and strongly encouraged to report any unexpected bleeding or spotting to their physicians.</td>
</tr>
<tr>
<td>Lung</td>
<td>Current or former smokers age 50–74 in good health with at least a 30 pack/year history</td>
<td>LDCT</td>
<td>Clinicians with access to high-volume, high-quality lung cancer screening and treatment centers should initiate a discussion about lung cancer screening with apparently healthy patients age 55–74 y who have at least a 30 pack-y smoking history, and who currently smoke or have quit within the past 15 y. A process of informed and shared decision-making with a clinician related to the potential benefits, limitations, and harms associated with screening for lung cancer with LDCT should occur before any decision is made to initiate lung cancer screening. Smoking cessation counseling remains a high priority for clinical attention in discussions with current smokers, who should be informed of their continuing risk of lung cancer. Screening should not be viewed as an alternative to smoking cessation.</td>
</tr>
</tbody>
</table>

(Continued)
**Table 10-9**

American Cancer Society recommendations for early detection of cancer in average-risk, asymptomatic individuals (Continued)

<table>
<thead>
<tr>
<th>CANCER SITE</th>
<th>POPULATION</th>
<th>TEST OR PROCEDURE</th>
<th>FREQUENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prostate</td>
<td>Men, age ≥50 y</td>
<td>DRE and PSA</td>
<td>Men who have at least a 10-y life expectancy should have an opportunity to make an informed decision with their health care provider about whether to be screened for prostate cancer, after receiving information about the potential benefits, risks, and uncertainties associated with prostate cancer screening. Prostate cancer screening should not occur without an informed decision-making process. Men at average risk should receive this information beginning at age 50 years. Men in higher risk groups should receive this information before age 50 years. Men should either receive this information directly from their healthcare providers or be referred to reliable and culturally appropriate sources. Patient decision aids are helpful in preparing men to make a decision whether to be tested.</td>
</tr>
</tbody>
</table>

Cancer-related checkup Men and women age ≥20 y On the occasion of a periodic health examination, the cancer-related checkup should include examination for cancers of the thyroid, testicles, ovaries, lymph nodes, oral cavity, and skin, as well as health counseling about tobacco, sun exposure, diet and nutrition, risk factors, sexual practices, and environmental and occupational exposures.

ACS = American Cancer Society; BSE = breast self-examination; CBE = clinical breast examination; Pap = Papanicolaou; HPV = human papillomavirus; gFOBT = guaiac-based fecal occult blood test; FIT = fecal immunochemical test; sDNA, stool DNA; DRE = digital rectal examination; FSIG = flexible sigmoidoscopy; DCBE = double-contrast barium enema; CT = computed tomography; LDCT = low-dose helical CT; PSA = prostate-specific antigen.


mammography for breast cancer screening in BRCA mutation carriers, first-degree relatives of carriers, and women with a lifetime breast cancer risk of 20% to 25% or higher.107

The National Lung Screening Trial demonstrated a 20% reduction in lung cancer deaths in adults age 55 to 74 years who were at high risk of lung cancer and randomized to low-dose helical computed tomography (LDCT) screening compared with screening with annual CXR.108 In 2013, the American Cancer Society updated their lung cancer screening recommendations to emphasize that clinicians with access to high-volume, high-quality lung cancer screening and treatment centers should ascertain the smoking history of their patients 55 to 74 years of age, and should discuss lung cancer screening with those who have at least a 30 pack per year smoking history, currently smoke, or have quit within the past 15 years, and who are in relatively good health.109 It is recommended that this discussion include the benefits, uncertainties, and harms associated with screening for lung cancer with LDCT.

**CANCER DIAGNOSIS**

The definitive diagnosis of solid tumors is obtained by performing a biopsy specimen of the lesion. Biopsy findings determine the tumor histology and grade and thus, assist in definitive therapeutic planning. Biopsy specimens of mucosal lesions usually are obtained endoscopically (e.g., via colonoscope, bronchoscope, or cystoscope). Lesions that are easily palpable, such as those of the skin, can either be excised or sampled by punch biopsy specimen. Deep-seated lesions can be localized with computed tomographic (CT) scan or ultrasound guidance for acquisition of a biopsy specimen.

A sample of a lesion can be obtained with a needle or with an open incisional or excisional biopsy specimen. Fine-needle aspiration is easy and relatively safe, but has the disadvantage of not giving information on tissue architecture. For example, fine-needle aspiration biopsy specimen of a breast mass can make the diagnosis of malignancy but cannot differentiate between an invasive and noninvasive tumor. Therefore core-needle biopsy specimen is more advantageous when the histologic findings will affect the recommended therapy. Core biopsy specimen, like fine-needle aspiration, is relatively safe and can be performed either by direct palpation (e.g., a breast mass or a soft tissue mass) or can be guided by an imaging study (e.g., stereotactic core biopsy specimen of the breast). Core biopsy specimens, like fine-needle aspirations, have the disadvantage of introducing sampling error. For example, 19% to 44% of patients with a diagnosis of atypical ductal hyperplasia based on core biopsy specimen findings of a mammographic abnormality are found to have carcinoma upon excision of the entire lesion.110 It is crucial to ensure that the histologic findings are consistent with the clinical scenario and to know the appropriate interpretation of each histologic finding. A needle biopsy specimen for which the report is inconsistent with the clinical scenario should be either repeated or followed by an open biopsy procedure.

Open biopsy specimens have the advantage of providing more tissue for histologic evaluation and the disadvantage of being an operative procedure. Incisional biopsy specimens are
Cancer staging is a system used to describe the anatomic extent of a malignant process in an individual patient. Staging systems may incorporate relevant clinical prognostic factors such as tumor size, location, extent, grade, and dissemination to regional lymph nodes or distant sites. Accurate staging is essential in designing an appropriate treatment regimen for an individual patient. Staging of the lymph node basin is considered a standard part of primary surgical therapy for most surgical procedures and is discussed later in this chapter. Cancer patients who are considered to be at high risk for distant metastasis usually undergo a preoperative staging work-up. This involves a set of imaging studies of sites of preferential metastasis for a given cancer type. For a patient with breast cancer, for example, a staging work-up would include a chest radiograph, bone scan, and liver ultrasound, or CT scans to evaluate for lung, bone, and liver metastases, respectively. A distant staging work-up usually is performed only for patients likely to have metastasis based on the characteristics of the primary tumor; for example, a staging work-up for a patient with ductal carcinoma in situ of the breast or a small invasive breast tumor is likely to be low yield and not cost effective.

Recently there also is increased usage of molecular imaging with positron emission tomography (PET) scanning, or PET/CT, for cancer staging. Most commonly PET scanning is performed with fluorine 18 incorporated into fluorodeoxyglucose (FDG). FDG PET assesses the rate of glycolysis. FDG uptake is increased in most malignant tissues but also in benign pathologic conditions such as inflammatory disorders, trauma, infection, and granulomatous disease. PET/CT combines a PET scanner and an X-ray CT scanner in a single gantry, in order to acquire sequential images from both devices in the same session. These separate images are combined into a single coregistered image that gives information on the size and shape of abnormal masses in conjunction with their metabolic activity. It has been especially useful in the staging and management of lymphoma, lung cancer, and colorectal cancer. The role of PET scanning in evaluating many other cancers is evolving, and additional molecular tracers, such as 3′-deoxy-3′-(18F)-fluorothymidine, used to assess proliferation, are being actively pursued.

A PET scan can be useful in staging a cancer that potentially can be treated radically, such as small cell lung cancer. In the case of some cancers such as GIST, a PET scan can be used to establish baseline staging before commencing targeted therapy and assessing the overall response to therapy. Another use for PET scanning is the evaluation of an indeterminate lesion as in the case of a solitary pulmonary nodule that is suspected to be malignant in nature. In testicular cancer (seminoma) and lymphoma, this imaging modality has been shown to be effective in assessing suspected disease recurrence, relapse, and/or residual disease. Finally, PET scans have been effective in guiding biopsies in the setting of mesothelioma. Standardization of staging systems is essential to allow comparison of results from different studies from different institutions and worldwide. The staging systems proposed by the American Joint Committee on Cancer (AJCC) and the Union Internationale Contre le Cancer (International Union Against Cancer, or UICC) are among the most widely accepted staging systems. Both the AJCC and the UICC have adopted a shared tumor, node, and metastasis (TNM) staging system that defines the cancer in terms of the anatomic extent of disease and is based on assessment of three components: the size of the primary tumor (T), the presence (or absence) and extent of nodal metastases (N), and the presence (or absence) and extent of distant metastases (M).

The TNM staging applies only to tumors that have been microscopically confirmed to be malignant. Standard TNM staging (clinical and pathologic) is completed at initial diagnosis. Clinical staging (cTNM or TNM) is based on information gained up until the initial definitive treatment. Pathologic staging (pTNM) includes clinical information and information obtained from pathologic examination of the resected primary tumor and regional lymph nodes. Tumor size following neoadjuvant chemotherapy is designated as ypT, and should be based on the largest single focus of residual invasive cancer. Other classifications, such as staging at the time of retreatment for recurrence (rTNM) or autopsy staging (aTNM), should be clearly identified as such.

The clinical measurement of tumor size (T) is the one judged to be the most accurate for each individual case based on physical examination and imaging studies. For example, in breast cancer the size of the tumor could be obtained from a physical examination, mammogram, or ultrasound, and the tumor size is based only on the invasive component.

If even one lymph node is involved by tumor, the N component is at least N1. For many solid tumor types, simply the absence or presence of lymph node involvement is recorded, and the tumor is categorized either as N0 or N1. For other tumor types, the number of lymph nodes involved, the size of the lymph nodes or the lymph node metastasis, or the regional lymph node basin involved also has been shown to have prognostic value. In these cancers, the designations N1, N2, and N3 suggest an increasing abnormality of lymph nodes based on size, characteristics, and location. NX indicates that the lymph nodes cannot be fully assessed.

Cases in which there is no distant metastasis are designated M0, cases in which one or more distant metastases are detected
are designated M1, and cases in which the presence of distant metastasis cannot be assessed are designated MX. In clinical practice, negative findings on clinical history and examination are sufficient to designate a case as M0. However, in clinical trials, routine follow-up often is performed to standardize the detection of distant metastases.

The practice of dividing cancer cases into groups according to stage is based on the observation that the survival rates are higher for localized (lower-stage) tumors than for tumors that have extended beyond the organ of origin. Therefore, staging assists in selection of therapy, estimation of prognosis, evaluation of treatments, and exchange of information among treatment centers. Notably, the AJCC regularly updates its staging system to incorporate advances in prognostic technology to improve the predictive accuracy of the TNM system. Therefore, it is important to know which revision of a staging system is being used when evaluating studies.

**TUMOR MARKERS**

**Prognostic and Predictive Tissue Markers**

Tumor markers are substances that can be detected in higher than normal amounts in the serum, urine, or tissues of patients with certain types of cancer. Tumor markers are produced either by the cancer cells themselves or by the body in a response to the cancer.

Over the past decade, there has been an especially high interest in identifying tissue tumor markers that can be used as prognostic or predictive markers. Although the terms **prognostic marker** and **predictive marker** are sometimes used interchangeably, the term **prognostic marker** generally is used to describe molecular markers that predict disease-free survival, disease-specific survival, and overall survival, whereas the term **predictive marker** often is used in the context of predicting response to certain therapies.

The goal is to identify prognostic markers that can give information on prognosis independent of other clinical characteristics and therefore can provide information to supplement the projections based on clinical presentation. This would allow practitioners to further classify patients as being at higher or lower risk within clinical subgroups and to identify patients who may benefit most from adjuvant therapy. For example, ideal prognostic tumor markers would be able to help determine which patients with node-negative breast cancer are at higher risk of relapse so that adjuvant systemic therapy could be given only to that group. However, although a large number of studies have identified potential novel prognostic markers, most have not been tested with enough vigor to be shown to be of clinical utility. In the 2017, American Society of Clinical Oncology (ASCO) guidelines, it was decided that level of uPA/PAI-1 measured by enzyme-linked immunosorbent assay could be used to determine prognosis in cases of newly diagnosed node-negative, hormone receptor positive breast cancer. In contrast, the data for many other markers, including Ki-67, p27, HER1/EGFR, and p53 were felt to be insufficient to support their use in the management of these breast cancer patients. Similarly, guidelines are available for the management of patients with colorectal cancer and emphasize the processes that need to be in place for accurate measurement of abnormalities in DNA mismatch repair genes as well as EGFR and BRAF mutational status.

Predictive markers are markers that can prospectively identify patients who will benefit from a certain therapy. For example in breast cancer, estrogen receptor (ER), and HER2 assessment can identify patients who can benefit from anti-estrogen therapies (e.g., tamoxifen) and anti-HER2 targeted therapies (e.g., trastuzumab), respectively, and the ASCO guidelines recommend that these markers be routinely assessed. High-throughput techniques such as transcriptional profiling allow for assessment of the relative mRNA levels of thousands of genes simultaneously in a given tumor using microarray technology. With the advent of such molecular profiling technologies, researchers have focused on identifying expression profiles that are prognostic for different cancer types. For breast cancer, although many such multiparameter tests are under development, few have reached the large-scale validation stage. In 2007, ASCO guidelines suggested that one of these, the Oncotype DX assay, can be used to predict recurrence in women with node-negative, ER-positive breast cancer who are treated with tamoxifen. Oncotype DX is a quantitative reverse-transcriptase polymerase chain reaction (RT-PCR) test that used paraffin-fixed tissue. A 21-gene recurrence score (RS) is generated based on the expression of 16 cancer genes and 5 reference genes. The levels of expression are used to derive an RS that ranges from 0 to 100, using a prospectively defined mathematical algorithm. This novel quantitative approach to the evaluation of the best-known molecular pathways in breast cancer has produced impressive results. Use of this multigene assay to predict recurrence was validated in the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-14 trial, in which ER-positive, node-negative patients had received tamoxifen. By multivariate Cox proportional analysis, RS was found to be independently associated with recurrence risk, with a hazard ratio of 3.21 (95% confidence interval of 2.23 to 4.65, \( P < .001 \)). The RS was indeed able to stratify patients by freedom from distant recurrence (Fig. 10-13). The Trial Assessing Individualized Options for Treatment for breast cancer (TAILORx) is evaluating the utility of Oncotype DX for predicting prognosis in patients with ER-positive, node-negative tumors and will focus on women with intermediate RS scores in whom the role of chemotherapy is unclear. Several other multigene predictors for breast cancer are available including MammaPrint, a gene expression profiling platform assessing a
70-gene transcriptional signature. This assay was approved by the Food and Drug Administration (FDA) in February 2007. The usefulness of this assay in making therapy-related decisions was tested prospectively in a large-scale study, the Microarray in Node-Negative Disease May Avoid Chemotherapy (MIND- ACT) trial. Among women with early-stage breast cancer who were determined to be at high clinical risk and low genomic risk for cancer recurrence, the receipt of no adjuvant chemotherapy on the basis of the 70-gene signature result led to a 5-year metastasis survival rate that was 1.5% lower than the rate in patients who had received chemotherapy. In light of these findings, 45% of women with breast cancer who are considered to be at high clinical risk might not require adjuvant chemotherapy.

Multigene profiles to predict prognosis are in development or in validation phases for many other solid tumor types, including lung cancer, ovarian cancer, pancreatic cancer, colorectal cancer, and melanoma. Gene signatures and genomic alterations also are being studied for their ability to predict response to specific chemotherapy regimens or targeted therapies. Many of these multigene marker sets will likely be incorporated into clinical practice in the years to come.

Serum Markers

Serum markers are under active investigation because they may allow early diagnosis of a new cancer or may be used to follow cancer response to therapy or monitor for recurrence. Unfortunately, identification of serum markers of clinical value has been challenging. Many of the tumor markers proposed so far have had low sensitivities and specificities. Tumor marker levels may not be elevated in all patients with cancer, especially in the early stages, when a serum marker would be most useful for diagnosis. Therefore, when a tumor marker is used to monitor recurrence, it is important to be certain that the level of the tumor marker was elevated before primary therapy. Moreover, tumor marker levels can be elevated in benign conditions. Many tumor markers are not specific for a certain type of cancer and can be elevated with more than one type of tumor. Since there may be significant laboratory variability, it is important to obtain serial results from the same laboratory. In spite of these many clinical limitations, several serum markers are in clinical use. A few of the commonly measured serum tumor markers are discussed in the following sections.

Prostate-Specific Antigen. Prostate-specific antigen (PSA) is an androgen-regulated serine protease produced by the prostate epithelium. PSA is normally present in low concentrations in the blood of all adult males. PSA levels may be elevated in the blood of men with benign prostate conditions such as prostatitis and benign prostatic hyperplasia, as well as in men with prostate cancer. PSA levels have been shown to be useful in evaluating the effectiveness of prostate cancer treatment and monitoring for recurrence after therapy. In monitoring for recurrence, a trend of increasing levels is considered more significant than a single absolute elevated value.

Although PSA has been widely used for prostate cancer screening, the utility of PSA screening remains controversial. There is concern that the number of men who avoid dying from prostate cancer due to screening is small, while the harms related to the treatment of screen-detected cancers, including incontinence and erectile dysfunction are at least moderate. In 2012, the U.S. Preventive Services Task Force concluded with moderate certainty that the harms of PSA testing outweigh the benefits and on that basis recommended against PSA-based screening for all men. In 2010, the American Cancer Society updated its guidelines for the early detection of prostate cancer to state that men who have at least a 10-year life expectancy should have an opportunity to make an informed decision with their health care provider about whether to be screened for prostate cancer with digital rectal exam and serum PSA, after receiving information about the benefits, risks, and uncertainties associated with prostate cancer screening; this recommendation was reinforced in their most recent guidelines.

Carcinoembryonic Antigen. Carcinoembryonic antigen (CEA) is a glycoprotein found in the embryonic endodermal epithelium. Elevated CEA levels have been detected in patients with primary colorectal cancer as well as in patients with breast, lung, ovarian, prostate, liver, and pancreatic cancer. Levels of CEA also may be elevated in benign conditions, including diverticulitis, peptic ulcer disease, bronchitis, liver abscess, and alcoholic cirrhosis, especially in smokers and in elderly persons.

CEA measurement is most commonly used in the management of colorectal cancer. However, the appropriate use of CEA testing in patients with colorectal cancer has been debated. Use of CEA level as a screening test for colorectal cancer is not recommended. CEA levels may be useful if obtained preoperatively and postoperatively in patients with a diagnosis of colorectal cancer. Preoperative elevation of CEA level is an indicator of poor prognosis. However, the ASCO clinical practice guidelines state that the data are insufficient to support the use of CEA to determine whether to give a patient adjuvant therapy; the data are stronger for the use of CEA for monitoring for postoperative recurrence. CEA measurement is the most cost-effective approach for detecting metastasis, with over 60% of recurrences being detected first by an elevation in CEA level. Therefore, in cases in which the patient would be a candidate for resection of recurrent colorectal cancer or systemic therapy, the ASCO guidelines recommend that postoperative CEA testing be performed every 3 months in patients with stage II or III disease for at least 3 years. CEA is the marker of choice for monitoring metastatic colorectal cancer during systemic therapy.

Alpha-Fetoprotein. Alpha-fetoprotein (AFP) is a glycoprotein normally produced by a developing fetus. AFP levels decrease soon after birth in healthy adults. An elevated level of AFP suggests the presence of either primary liver cancer or a germ cell tumor of the ovary or testicle. Rarely, other types of cancer such as gastric are associated with an elevated AFP level. Benign conditions that can cause elevations of AFP include cirrhosis, hepatic necrosis, acute hepatitis, chronic active hepatitis, ataxia-telangiectasia, Wiskott-Aldrich syndrome, and pregnancy. The sensitivity of an elevated AFP level for detecting HCC is approximately 60%. AFP is considered to be sensitive and specific enough to be used for screening for HCC in high-risk populations. Current consensus recommendations are to screen healthy hepatitis B virus carriers with annual or semiannual measurement of AFP level and to screen carriers with cirrhosis or chronic hepatitis and patients with cirrhosis of any etiology with twice-yearly measurement of AFP level and liver ultrasonography. Although AFP testing has been used widely for a long time, its efficacy in early diagnosis of HCC is limited. With improvements in imaging technology, a larger proportion of patients diagnosed with HCC are now AFP seronegative.

Cancer Antigen 19-9. Cancer antigen 19-9 (CA 19-9) is a tumor-related antigen that was originally defined by a monoclonal antibody produced by a hybridoma prepared from murine
spleen cells immunized with a human colorectal cancer cell line. The data are insufficient to recommend use of CA 19-9 for screening, diagnosis, surveillance, or monitoring of therapy for colon cancer. Based on the 2006 ASCO guidelines, there are also insufficient data to recommend use of CA 19-9 for screening, diagnosis, or determination of the operability of pancreatic cancer. However, for patients with locally advanced or metastatic cancer receiving active therapy, CA 19-9 can be measured at the start of therapy and every 1 to 3 months while therapy is given; elevations in serial CA 19-9 levels may indicate progressive disease and should be confirmed by additional studies.

**Cancer Antigen 15-3.** Cancer antigen 15-3 (CA 15-3) is an epitope of a large membrane glycoprotein encoded by the MUC1 gene that tumor cells shed into the bloodstream. The CA 15-3 epitope is recognized by two monoclonal antibodies in a sandwich radioimmunoassay. CA 15-3 levels are most useful in following the course of treatment in women diagnosed with advanced breast cancer. CA 15-3 levels are infrequently elevated in early-stage breast cancer. CA 15-3 levels can be increased in benign conditions such as chronic hepatitis, tuberculosis, sarcoidosis, pelvic inflammatory disease, endometriosis, systemic lupus erythematosus, pregnancy, and lactation, and in other types of cancer such as lung, ovarian, endometrial, and GI cancers.

The sensitivity of CA 15-3 is higher for metastatic disease, and in these cases studies have shown sensitivity to be between 54% and 87%, with specificity as high as 96%. This has led to interest in using CA 15-3 for monitoring patients with advanced breast cancer for recurrence. Elevated CA 15-3 levels have been reported before relapse in 54% of patients, with a lead time of 4.2 months. Therefore, detection of elevated CA 15-3 levels during follow-up should prompt evaluation for recurrent disease. However, 6% to 8% of patients without recurrence will have elevated CA 15-3 levels that require evaluation. Furthermore, monitoring with the use of CA 15-3 levels has shown no demonstrated impact on survival. Therefore, the 2007 ASCO guidelines state that the routine use of CA 15-3 for screening, diagnosis, staging, or surveillance of breast cancer is not recommended because available data are insufficient. For monitoring patients during active therapy, CA 15-3 can be used in conjunction with diagnostic imaging and history and physical examination. In the absence of measurable disease, an increase may be interpreted to indicate treatment failure. However, caution is advised when interpreting rising levels in the first 4 to 6 weeks of therapy.

**Cancer Antigen 27-29.** The MUC-1 gene product in the serum may be quantitated by using radioimmunoassay with a monoclonal antibody against the cancer antigen 27-29 (CA 27-29). CA 27-29 levels can be elevated in breast cancer as well as in cancers of the colon, stomach, kidney, lung, ovary, pancreas, uterus, and liver. First-trimester pregnancy, endometriosis, benign breast disease, kidney disease, and liver disease also may be associated with elevated CA 27-29 levels.

CA 27-29 has been reported to have a sensitivity of 57%, a specificity of 98%, a positive predictive value of 83%, and a negative predictive value of 93% in detecting breast cancer recurrences. Although CA 27-29 has been found to predict recurrence an average of 5.3 months before other symptoms or tests, testing of CA 27-29 levels has not been demonstrated to affect disease-free and overall survival rates. Therefore, the ASCO guidelines state that, as with CA 15-3, the routine use of CA 27-29 for screening, diagnosis, staging, or surveillance of breast cancer is not recommended because available data are insufficient. CA 27-29 levels can be used together with diagnostic imaging and history and physical examination to monitor patients during active therapy. When no measurable disease is present, an increase in level may be considered to indicate treatment failure. However, rising levels in the first 4 to 6 weeks of therapy should be interpreted with caution.

**Circulating Tumor Cells**

Circulating tumor cells (CTCs) are cells present in the blood that possess antigenic or genetic characteristics of a specific tumor type. One CTC detection methodology is capture and quantitation of CTCs with immunomagnetic beads coated with antibody specific for cell-surface, epithelial, or cancer antigens. Another methodology used to detect cancer cells in the peripheral blood is RT-PCR. It has been suggested that measurement of CTCs can be an effective tool for selecting patients who have a high risk of relapse and for monitoring efficacy of cancer therapy.

CTCs have probably been most extensively studied in breast cancer. The most promising data come from the use of CTC measures in metastatic breast cancer. In a prospective multicenter trial, the number of CTCs (≥5 CTCs vs. <5 CTCs per 7.5 mL of whole blood) before treatment of metastatic breast cancer was an independent predictor of progression-free and overall survival rates. The presence of >5 CTCs after the first course of therapy predicted lack of response to treatment. This technology, known as CellSearch, has been approved by the FDA for clinical use. In a well-designed single institutional study, detection of one or more CTCs in stage I to III breast cancer patients was associated with both decreased progression-free survival and overall survival. A large clinical trial was launched by the German SUCCESS study group to evaluate the prognostic significance of CTCs in early breast cancer. Blood samples were obtained from over 2000 average-to-high-risk nonmetastatic breast cancer patients before chemotherapy and nearly 1500 patients post chemotherapy and examined for CTCs. Women with detectable CTCs before chemotherapy had significantly worse disease-free and overall survival, and those with five or more CTCs had the highest relapse risk. However, there is limited data to prove that the use of CTC testing leads to improved survival or improved quality of life; thus, the ASCO 2017 guidelines update did not recommend the use of CTC measurement in any clinical setting. The clinical utility of measuring CTC response to initial therapy is now being tested prospectively in multiple clinical trials. The ability to conduct comprehensive analyses of cancer genomes within individual cells is becoming a real possibility, although the clinical utility of such information is still evolving. In addition to CTC, it has become possible in recent years to conduct a so-called “liquid tumor biopsy” by measuring levels of circulating tumor DNA and also circulating microRNAs. These new techniques are investigational but have the potential to provide prognostic and diagnostic information via their presumed correlation with tumor progression and the development of genomic alterations. The prognostic implications of detection of CTCs by RT-PCR have been intensively studied for melanoma. In the recent multicenter Sunbelt Melanoma Trial, serial RT-PCR was performed on peripheral blood samples using four markers—tyrosinase, melanoma antigen reacting to T cell (MART-1), melanoma antigen 3 (MAGE3), and gp 100—to detect occult
melanoma cells in the bloodstream.\textsuperscript{127} Although there were no differences in survival rates among patients in whom at least one marker was detected and those in whom no markers were detected, the disease-free survival and distant disease-free survival rates were worse for patients in whom more than one marker was detected at any time during follow-up.\textsuperscript{127}

**Bone Marrow Micrometastases**

Micrometastatic disease in the bone marrow, also referred to as minimal residual disease, continues to be investigated as a potential prognostic marker. Bone marrow micrometastatic disease usually is detected by staining bone marrow aspirates with monoclonal antibodies to cytokeratin, but other methodologies such as flow cytometry and RT-PCR are in use. Breast cancer patients with bone marrow micrometastasis have larger tumors, tumors with a higher histologic grade, more lymph node metastases, and more hormone receptor-negative tumors than patients without bone marrow micrometastasis. In 4700 patients with stage I, II, or III breast cancer, micrometastasis was a significant prognostic factor associated with poor overall survival, breast cancer-specific survival, disease-free survival, and distant disease-free survival during a 10-year observation period.\textsuperscript{128} The American College of Surgeons Oncology Group Z0010 trial enrolled women with clinical T1 to T2N0M0 invasive breast carcinoma in a prospective observational study to determine the association between survival and metastases detected by immunohistochemical staining of bone marrow specimens from patients with early-stage breast cancer.\textsuperscript{129} Of 3413 bone marrow specimens examined by immunocytochemistry, only 104 (3.0\%) were positive for tumor. Bone marrow involvement was associated with a decreased overall survival, but this association was not significant on multivariable analysis. The prognostic implication of bone marrow involvement continues to be studied by the National Surgical Adjuvant Breast and Bowel Project.

At this time, the routine use of bone marrow testing is not recommended.\textsuperscript{131} Ongoing clinical trials are evaluating the role of routine assessment of bone marrow status in the care of patients with early and advanced breast cancer. The utility of assessment of bone marrow micrometastasis has also been evaluated in other tumor types, including gastric, esophageal, colorectal, lung, cervical, and ovarian cancer.\textsuperscript{130}

**SURGICAL APPROACHES TO CANCER THERAPY**

**Multidisciplinary Approach to Cancer**

Although surgery is an effective therapy for most solid tumors, patients who die from cancer usually die of metastatic disease. Therefore, to improve patient survival rates, a multimodality approach, including systemic therapy and radiation therapy, is key for most tumors. It is important that surgeons involved in cancer care not only know the techniques for performing a cancer operation but also know the alternatives to surgery and be well versed in reconstructive options. It is also crucial that the surgeon be familiar with the indications for and complications of preoperative and postoperative chemotherapy and radiation therapy. Although the surgeon may not be delivering these other therapies, as the first physician to see a patient with a cancer diagnosis, he or she is ultimately responsible for initiating the appropriate consultations. For this reason, the surgeon often is responsible for determining the most appropriate adjuvant therapy for a given patient as well as the best sequence for therapy. In most instances, a multidisciplinary approach beginning at the patient’s initial presentation is likely to yield the best result.

**Surgical Management of Primary Tumors**

The goal of surgical therapy for cancer is to achieve oncologic cure. A curative operation presupposes that the tumor is confined to the organ of origin or to the organ and the regional lymph node basin. Patients in whom the primary tumor is not resectable with negative surgical margins are considered to have inoperable disease. The operability of primary tumors is best determined before surgery with appropriate imaging studies that can define the extent of local-regional disease. For example, a preoperative thin-section CT scan is obtained to determine resectability of pancreatic cancer, which is based on the absence of extrapancreatic disease, the absence of tumor extension to the superior mesenteric artery and celiac axis, and a patent superior mesenteric vein-portal vein confluence.\textsuperscript{130} Disease involving multiple distant metastases is deemed inoperable because it is usually not curable with surgery of the primary tumor. Therefore, patients who are at high risk of having distant metastasis should undergo a staging work-up before surgery for the primary tumor. On occasion, primary tumors are resected in these patients for palliative reasons, such as improving the quality of life by alleviating pain, infection, or bleeding. An example of this is toilet mastectomies for large ulcerated breast tumors. Patients with limited metastases from a primary tumor on occasion are considered surgical candidates if the natural history of isolated distant metastases for that cancer type is favorable or the potential complications associated with leaving the primary tumor intact are significant.

In the past, it was presumed that the more radical the surgery, the better the oncologic outcome would be. Over the past three decades, this has been recognized as not necessarily being true, which has led to more conservative operations, with wide local excisions replacing compartmental resections of sarcomas, and partial mastectomies, skin-sparing mastectomies, and breast-conserving therapies replacing radical mastectomies for breast cancer. The uniform goal for all successful oncologic operations seems to be achieving widely negative margins with no evidence of macroscopic or microscopic tumor at the surgical margins. The importance of negative surgical margins for local tumor control and/or survival has been documented for many tumor types, including sarcoma, breast cancer, pancreatic cancer, and rectal cancer. Thus, it is clear that every effort should be made to achieve microscopically negative surgical margins. Inking of the margins, orientation of the specimen by the surgeon, and immediate gross evaluation of the margins by a pathologist using frozen-section analysis when necessary may assist in achieving negative margins at the first operation. In the end, although radiation therapy and systemic therapy can assist in decreasing local recurrence rates in the setting of positive margins, adjuvant therapy cannot substitute for adequate surgery.

Although it is clear that the surgical gold standard is negative surgical margins, the appropriate surgical margins for optimal local control are controversial for many cancer types. In contrast, in melanoma the optimal margin width for any tumor depth has been better defined, owing to the systematic study of this question in randomized clinical trials.\textsuperscript{132,133} Although such randomized studies may not be possible for all tumor types, it is important to determine optimal surgical margins for each cancer type so that adjuvant radiation and systemic therapy can be offered to patients deemed to be at increased risk for local treatment failure.
Most neoplasms have the ability to metastasize via the lymphatics. Therefore, most oncologic operations have been designed to remove the primary tumor and draining lymphatics en bloc. This type of operative approach usually is undertaken when the lymph nodes draining the primary tumor site lie adjacent to the tumor bed, as is the case for colorectal cancers and gastric cancers. For tumors in which the regional lymph node basin is not immediately adjacent to the tumor (e.g., melanomas), lymph node surgery can be performed through a separate incision. Unlike most carcinomas, soft tissue sarcomas rarely metastasize to the lymph nodes (<5%); therefore, lymph node surgery usually is not necessary.

It is generally accepted that a formal lymphadenectomy is likely to minimize the risk of regional recurrence of most cancers. For example, the introduction of total mesorectal excision of rectal cancer has been associated with a large decline in local-regional recurrence, and this procedure has become the new standard of operative management. On the other hand, there have been two opposing views regarding the role of lymphadenectomy in survival of cancer patients. The traditional Halsted view states that lymphadenectomy is important for staging and survival. The opposing view counters that cancer is systemic at inception and that lymphadenectomy, although useful for staging, does not affect survival. For most cancers, involvement of the lymph nodes is one of the most significant prognostic factors. Interestingly, in some studies removal of a larger number of lymph nodes has been found to be associated with an improved overall survival rate for many tumors, including breast cancer, colon cancer, and lung cancer. Although this seems to support the Halsted theory, the practice of sentinel lymph node biopsy specimens yields the risk of regional recurrence, there may be alternative explanations for the same finding. For example, the surgeon who performs a more extensive lymphadenectomy may obtain wider margins around the tumor or even provide better overall care, such as ensuring that patients receive the appropriate adjuvant therapy or undergo a more thorough staging workup. Alternatively, the pathologist may perform a more thorough examination, identifying more nodes and more accurately staging the nodes. The effect of appropriate staging on survival is twofold. Patients with nodal metastases may be offered adjuvant therapy, which improves their survival chances. Further, the enhanced staging can improve perceived survival rates through a “Will Rogers effect.” Such a phenomenon is observed when moving an element from one set to another set raises the average value of both sets. When commenting on the 1930s migration of poor farmers from a dustbowl state to a more prosperous western state, humorist Will Rogers quipped that this event raised the average intelligence of both states. Thus, identification of small metastases that had formerly been silent and unidentified leads to stage migration for these patients and thus to a perceived improvement in chances of survival for the higher stage. In addition, there is improved survival for the lower stage, which is now minus the patients with low volume nodal disease. Clearly the impact of lymphadenectomy on survival will continue to be a topic of clinical research.

Surgical management of the clinically negative regional lymph node basin has evolved with the introduction of lymphatic mapping technology (Fig. 10-14). Lymphatic mapping and sentinel lymph node biopsy specimen were first reported in 1977 by Cabanas for penile cancer. Now, sentinel node biopsy specimen is the standard of care for the management of melanoma and breast cancer. The utility of sentinel node biopsy is being explored in other cancer types such as head and neck squamous cell cancer and vulvar cancer.

The first node to receive drainage from the tumor site is termed the sentinel node. This node is the node most likely to contain metastases, if metastases to that regional lymph node basin are present. The goal of lymphatic mapping and sentinel lymph node biopsy specimen is to identify and remove the lymph node most likely to contain metastases in the least invasive fashion. The practice of sentinel lymph node biopsy specimen followed by regional lymph node dissection for selected patients with a positive sentinel lymph node avoids the morbidity of lymph node dissections in patients with negative nodes. An additional advantage of the sentinel lymph node technique is that it directs attention to a single node, which allows more careful analysis of the lymph node most likely to have a positive yield and increases the accuracy of nodal staging. Two criteria are used to assess the efficacy of a sentinel lymph node biopsy specimen: the sentinel lymph node identification rate and the false-negative rate. The sentinel lymph node identification rate

**Figure 10-14.** Lymphatic mapping and sentinel lymph node biopsy specimen for breast cancer. A. Peritumoral injection of blue dye. B. Blue dye draining into the sentinel lymph node.
is the proportion of patients in whom a sentinel lymph node was identified and removed among all patients undergoing an attempted sentinel lymph node biopsy specimen. The false-negative rate is the proportion of patients with regional lymph node metastases in whom the sentinel lymph node was found to be negative. False-negative biopsy specimen results may be due to identifying the wrong node or to missing the sentinel node (i.e., surgical error) or they may be due to the cancer cells’ establishing metastases not in the first node encountered but in a second-echelon node (i.e., biologic variation). Alternatively, false-negative biopsy specimen results may be due to inadequate histologic evaluation of the lymph node. The false-negative rates for sentinel lymph node biopsy specimen in study series range between 0% and 11%. Both increases in the identification rate and decreases in the false-negative rate have been observed as surgeons gain experience with the technique.

Lymphatic mapping is performed by using isosulfan blue dye, technetium-labeled sulfur colloid or albumin, or a combination of both techniques to detect sentinel nodes. The combination of blue dye and technetium has been reported to improve the capability of detecting sentinel lymph nodes. The nodal drainage pattern usually is determined with a preoperative lymphoscintigram, and the “hot” and/or blue nodes are identified with the assistance of a gamma probe and careful nodal basin exploration. Careful manual palpation is a crucial part of the procedure to minimize the false-negative rate.

The nodes are evaluated with serial sectioning, hematoxylin and eosin staining, and immunohistochemical analysis with S-100 protein and homatropine methylbromide staining for melanoma and cytokeratin staining for breast cancer. The utility of molecular techniques such as RT-PCR to assess the sentinel nodes is still being explored.

Another area of active investigation is the prognostic value of minimal nodal involvement. For example, in breast cancer, nodes with isolated tumor cell deposits of <0.2 mm are considered to be N0 by the sixth edition of the AJCC staging manual. However, some retrospective studies have suggested that even this amount of nodal disease burden has negative prognostic implications. Molecular ultrastaging with RT-PCR for patients with node-negative disease was assessed in a prospective multicenter trial and was found to not be prognostic in malignant melanoma. However, a recent meta-analysis of 22 studies enrolling 4019 patients found that PCR positivity was associated with worse overall and disease-free survival. Further study of the utility of ultrastaging of nodes in breast cancer, melanoma, and several other tumor types is ongoing.

Until recently, in breast cancer management, when sentinel node mapping revealed a positive sentinel node, this was followed by a completion axillary lymph node dissection. Results of the American College of Surgeons Oncology Group Z0011 trial, challenged this practice. ACOSOG Z11 was a phase 3 multicenter noninferiority trial conducted to determine the effects of completion lymph node dissections in melanoma is under investigation. In the MSLT-II clinical trial, melanoma patients with sentinel-node metastases were randomized to immediate completion lymph-node dissection or nodal observation with ultrasonography. The primary end point of this study was melanoma-specific survival. Immediate completion lymph-node dissection led to increased regional disease control and provided prognostic information but did not increase melanoma-specific survival among patients with intermediate-thickness melanoma and sentinel-node metastases.

Surgical Management of Distant Metastases

The treatment of a patient with distant metastases depends on the number and sites of metastases, the cancer type, the rate of tumor growth, the previous treatments delivered and the responses to these treatments, and the patient’s age, physical condition, and desires. Although once a tumor has metastasized it usually is not curable with surgical therapy, such therapy has resulted in cure in selected cases with isolated metastases to the liver, lung, or brain.

Patient selection is the key to the success of surgical therapy for distant metastases. The cancer type is a major determinant in surgical decision making. A liver metastasis from a colon cancer is more likely to be an isolated and thus resectable lesion than a liver metastasis from a pancreatic carcinoma. The growth rate of the tumor also plays an important role and can be determined in part by the disease-free interval and the time between treatment of the primary tumor and detection of the distant recurrence. Patients with longer disease-free intervals have a higher survival rate after surgical metastasectomy than those with a short disease-free interval. Similarly, patients who have synchronous metastases (metastases diagnosed at the initial cancer diagnosis) do worse after metastasectomy than patients who develop metachronous metastases (metastasis diagnosed after a disease-free interval). The natural history of metastatic disease is so poor for some tumors (e.g., pancreatic cancer) that there is no role at this time for surgical metastasectomy. In cancers with a more favorable outlook, observation for several weeks or months, potentially with initial treatment with systemic therapy, can allow the surgeon to monitor for metastases at other sites.

In curative surgery for distant metastases, as with surgery for primary tumors, the goal is to resect the metastases with negative margins. In patients with hepatic metastases that are unresectable because their location near intrahepatic blood vessels precludes a margin-negative resection, or because they are multifocal or hepatic function is inadequate, tumor ablation with cryotherapy or radiofrequency ablation is an alternative. Curative resections or ablative procedures should be attempted...
only if the lesions are accessible and the procedure can be performed safely.

**CHEMOTHERAPY**

**Clinical Use of Chemotherapy**

In patients with documented distant metastatic disease, chemotherapy is usually the primary modality of therapy. The goal of therapy in this setting is to decrease the tumor burden, thus prolonging survival. It is rare to achieve cure with chemotherapy for metastatic disease for most solid tumors. Chemotherapy administered to a patient who is at high risk for distant recurrence but has no evidence of distant disease is referred to as *adjuvant chemotherapy*. The goal of adjuvant chemotherapy is eradication of micrometastatic disease, with the intent of decreasing relapse rates and improving survival rates.

Adjuvant therapy can be administered after surgery (postoperative chemotherapy) or before surgery (neoadjuvant chemotherapy, or induction therapy). A portion or all of the planned adjuvant chemotherapy can be administered before the surgical removal of the primary tumor. Preoperative chemotherapy has three potential advantages. The first is that preoperative regression of tumor can facilitate resection of tumors that were initially inoperable, or allow more conservative surgery for patients whose cancer was operable to begin with. In the NSABP B-18 project, for example, women were randomly assigned to receive adjuvant doxorubicin and cyclophosphamide preoperatively or postoperatively. More patients treated before surgery than after surgery underwent breast-conserving surgery (68% vs. 60%). The second advantage of preoperative chemotherapy is the treatment of micrometastases without the delay of postoperative recovery. The third advantage is the ability to assess a cancer’s response to treatment clinically, after a number of courses of chemotherapy, and pathologically, after surgical resection. This is especially important if alternative treatment regimens are available to be offered to patients whose disease responded inadequately. Molecular characterization of the residual disease may also give insight into mechanisms of chemoresistance and possible therapeutic targets.

There are some potential disadvantages to preoperative chemotherapy, however. Although disease progression while the patient is receiving preoperative chemotherapy is rare in chemotherapy-sensitive tumors such as breast cancer, it is more frequent in relatively chemotherapy-resistant tumors such as sarcomas. Thus, patient selection is critical to ensure that the opportunity to treat disease surgically is not lost by giving preoperative chemotherapy. Often, rates of postoperative wound infection, flap necrosis, and delays in postoperative adjuvant therapy do not differ between patients who are treated with preoperative chemotherapy and patients who are treated with surgery first. However, preoperative chemotherapy can introduce special challenges to tumor localization, margin analysis, lymphatic mapping, and pathologic staging.

Response to chemotherapy is monitored clinically with imaging studies as well as physical examinations. Response usually is defined as complete response, partial response, stable disease, or progression. Response generally is assessed using the Response Evaluation Criteria in Solid Tumors (RECIST) criteria. Objective tumor response assessment is critical because tumor response is used as a prospective endpoint in clinical trials and tumor response is a guide to clinicians regarding continuation of current therapy.

**Principles of Chemotherapy**

Chemotherapy destroys cells by first-order kinetics, which means that with the administration of a drug a constant percentage of cells is killed, not a constant number of cells. If a patient with $10^{12}$ tumor cells is treated with a dose that results in 99.9% cell kill (3-log cell kill), the tumor burden will be reduced from $10^{12}$ to $10^9$ cells (or 1 kg to 1 g). If the patient is retreated with the same drug, which theoretically could result in another 3-log cell kill, the cells would decrease in number from $10^9$ to $10^6$ (1 g to 1 mg) rather than being eliminated totally.

Chemotherapeutic agents can be classified according to the phase of the cell cycle during which they are effective. Cell-cycle phase nonspecific agents (e.g., alkylating agents) have a linear dose-response curve, such that the fraction of cells killed increases with the dose of the drug. In contrast, the cell-cycle phase-specific drugs have a plateau with respect to cell killing ability, and cell kill will not increase with further increases in drug dose.

**Anticancer Agents**

**Alkylating Agents.** Alkylating agents are cell-cycle–nonspecific agents, that is, they are able to kill cells in any phase of the cell cycle. They act by cross-linking the two strands of the DNA helix or by causing other direct damage to the DNA. The damage to the DNA prevents cell division and, if severe enough, leads to apoptosis. The alkylating agents are composed of three main subgroups: classic alkylators, nitrosoureas, and miscellaneous DNA-binding agents (Table 10-10).

**Antitumor Antibiotics.** Antitumor antibiotics are the products of fermentation of microbial organisms. Like the alkylating agents, these agents are cell-cycle nonspecific. Antitumor antibiotics damage the cell by interfering with DNA or RNA synthesis, although the exact mechanism of action may differ by agent.

**Antimetabolites.** Antimetabolites are generally cell-cycle–specific agents that have their major activity during the S phase of the cell cycle and have little effect on cells in G0. These drugs are most effective, therefore, in tumors that have a high growth fraction. Antimetabolites are structural analogues of naturally occurring metabolites involved in DNA and RNA synthesis. Therefore, they interfere with normal synthesis of nucleic acids by substituting for purines or pyrimidines in the metabolic pathway to inhibit critical enzymes in nucleic acid synthesis. The antimetabolites include folate antagonists, purine antagonists, and pyrimidine antagonists.

**Plant Alkaloids.** Plant alkaloids are derived from plants such as the periwinkle plant, *Vinca rosea* (e.g., vincristine, a vinca alkaloid), or the root of American mandrake, *Podophyllum peltatum* (e.g., etoposide, a podophyllotoxin). Vinca alkaloids affect the cell by binding to tubulin in the S phase. This blocks microtubule polymerization, which results in impaired mitotic spindle formation in the M phase. Taxanes such as paclitaxel, on the other hand, cause excess polymerization and stability of microtubules, which blocks the cell cycle in mitosis. The epipodophyllotoxins (e.g., etoposide) act to inhibit a DNA enzyme called topoisomerase II by stabilizing the DNA-topoisomerase II complex. This results in an inability to synthesize DNA, and thus the cell cycle is stopped in the G1 phase.
Combination Chemotherapy

Combination chemotherapy may provide greater efficacy than single-agent therapy by three mechanisms: (a) it provides maximum cell kill within the range of toxicity for each drug that can be tolerated by the host, (b) it offers a broader range of coverage of resistant cell lines in a heterogeneous population, and (c) it prevents or delays the emergence of drug-resistant cell lines. When combination regimens are devised, drugs known to be active as single agents usually are selected. Drugs with different mechanisms of action are combined to allow for additive or synergistic effects. Combining cell-cycle–specific and cell-cycle–nonspecific agents may be especially advantageous. Drugs with differing dose-limiting toxic effects are combined to allow for each drug to be given at therapeutic doses. Drugs with different patterns of resistance are combined whenever possible to minimize cross-resistance. The treatment-free interval between cycles is kept to the shortest possible time that will allow for recovery of the most sensitive normal tissue.

Drug Toxicity

Tumors are more susceptible than normal tissue to chemotherapeutic agents, in part because they have a higher proportion of dividing cells. Normal tissues with a high growth fraction, such as the bone marrow, oral and intestinal mucosa and hair follicles are sensitive to chemotherapeutic effects. Therefore, treatment with chemotherapeutic agents can produce toxic effects such as bone marrow suppression, stomatitis, ulceration of the GI tract, and alopecia. Toxic effects usually are graded from 0 to 4 on the basis of World Health Organization standard criteria. Significant drug toxicity may necessitate a dosage reduction. A toxic effect requiring a dose modification or change in dose intensity is referred to as a dose-limiting toxic effect. Because maintaining dose intensity is important to preserve as high a tumor cell kill as possible, several supportive strategies have been developed, such as administration of colony-stimulating factors and erythropoietin to treat poor bone marrow reserve and administration of cytoprotectants such as mesna and amifostine to prevent renal dysfunction. Some toxicities, such as neuropathy, are not as easily reversible, and their potential effects on lifestyle must be considered when evaluating a patient prior to the initiation of chemotherapy.

Administration of Chemotherapy

Chemotherapy usually is administered systemically (IV, IM, SC, or PO). Systemic administration treats micrometastases at widespread sites and prevents systemic recurrence. However, it increases the drug’s toxicity to a wide range of organs throughout the body. One method to minimize systemic toxicity while enhancing target organ delivery of chemotherapy is regional administration of chemotherapy. Many of these approaches require surgical access, such as intrahepatic delivery of chemotherapy for hepatic carcinomas or metastatic colorectal cancer using a hepatic artery infusion pump, limb perfusion for extremity melanoma and sarcoma, and intraperitoneal hyperthermic

Table 10-10

<table>
<thead>
<tr>
<th>Classification of chemotherapeutic agents</th>
<th>Alkylating agents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classic alkylating agents</td>
<td>Busulfan</td>
</tr>
<tr>
<td></td>
<td>Chlorambucil</td>
</tr>
<tr>
<td></td>
<td>Cyclophosphamide</td>
</tr>
<tr>
<td></td>
<td>Ifosfamide</td>
</tr>
<tr>
<td></td>
<td>Mechlorethamine (nitrogen mustard)</td>
</tr>
<tr>
<td></td>
<td>Melphalan</td>
</tr>
<tr>
<td></td>
<td>Mitomycin C</td>
</tr>
<tr>
<td></td>
<td>Triethylene thiophosphoramid (thiotepa)</td>
</tr>
<tr>
<td>Nitrosoureas</td>
<td>Carmustine (BCNU)</td>
</tr>
<tr>
<td></td>
<td>Lomustine (CCNU)</td>
</tr>
<tr>
<td></td>
<td>Semustine (MeCCNU)</td>
</tr>
<tr>
<td></td>
<td>Streptozocin</td>
</tr>
<tr>
<td>Miscellaneous DNA-binding agents</td>
<td>Carboplatin</td>
</tr>
<tr>
<td></td>
<td>Cisplatin</td>
</tr>
<tr>
<td></td>
<td>Dacarbazine (DTIC)</td>
</tr>
<tr>
<td></td>
<td>Hexamethylmelamine</td>
</tr>
<tr>
<td></td>
<td>Procarbazine</td>
</tr>
<tr>
<td>Antitumor antibiotics</td>
<td>Bleomycin</td>
</tr>
<tr>
<td></td>
<td>Dactinomycin (actinomycin D)</td>
</tr>
<tr>
<td></td>
<td>Daunorubicin</td>
</tr>
<tr>
<td></td>
<td>Doxorubicin</td>
</tr>
<tr>
<td></td>
<td>Idarubicin</td>
</tr>
<tr>
<td></td>
<td>Plicamycin (mithramycin)</td>
</tr>
<tr>
<td>Antimetabolites</td>
<td>Folate analogues</td>
</tr>
<tr>
<td></td>
<td>Methotrexate</td>
</tr>
<tr>
<td>Purine analogues</td>
<td>Azathioprine</td>
</tr>
<tr>
<td></td>
<td>Mercaptopurine</td>
</tr>
<tr>
<td></td>
<td>Thioguanine</td>
</tr>
<tr>
<td></td>
<td>Cladribine (2-chlorodeoxyadenosine)</td>
</tr>
<tr>
<td></td>
<td>Fludarabine</td>
</tr>
<tr>
<td></td>
<td>Pentostatin</td>
</tr>
<tr>
<td>Pyrimidine analogues</td>
<td>Capecitabine</td>
</tr>
<tr>
<td></td>
<td>Cytarabine</td>
</tr>
<tr>
<td></td>
<td>Flouxuridine</td>
</tr>
<tr>
<td></td>
<td>Gemcitabine</td>
</tr>
<tr>
<td>Ribonucleotide reductase inhibitors</td>
<td>Hydroxyurea</td>
</tr>
<tr>
<td>Plant alkaloids</td>
<td>Vinca alkaloids</td>
</tr>
<tr>
<td></td>
<td>Vinblastine</td>
</tr>
<tr>
<td></td>
<td>Vincristine</td>
</tr>
<tr>
<td></td>
<td>Vindesine</td>
</tr>
<tr>
<td></td>
<td>Vinorelbine</td>
</tr>
<tr>
<td>Epopodophyllotoxins</td>
<td>Etoposide</td>
</tr>
<tr>
<td></td>
<td>Teniposide</td>
</tr>
<tr>
<td>Taxanes</td>
<td>Paclitaxel</td>
</tr>
<tr>
<td></td>
<td>Docetaxel</td>
</tr>
<tr>
<td>Miscellaneous agents</td>
<td>Asparaginase</td>
</tr>
<tr>
<td></td>
<td>Estramustine</td>
</tr>
<tr>
<td></td>
<td>Mitotane</td>
</tr>
</tbody>
</table>

Brunicardi_Ch10_p0305-p0354.indd   341 22/02/19   2:14 PM
perfusion for pseudomyxoma peritonei. Alternately, percutaneous access may be utilized, such as limb infusion with percutaneously placed catheters.

HORMONAL THERAPY

Some tumors, most notably breast and prostate cancers, originate from tissues whose growth is under hormonal control. The first attempts at hormonal therapy were through surgical ablation of the organ producing the hormones involved, such as oophorectomy for breast cancer. Currently, hormonal anticancer agents include androgens, antiandrogens, antiestrogens, estrogens, glucocorticoids, gonadotropin inhibitors, progestins, aromatase inhibitors, and somatostatin analogues. Hormones or hormone-like agents can be administered to inhibit tumor growth by blocking or antagonizing the naturally occurring substance, such as with the estrogen antagonist tamoxifen. Other substances that block the synthesis of the natural hormone can be administered as alternatives. Aromatase inhibitors, for example, block the peripheral conversion of endogenous androgens to estrogens in postmenopausal women. Hormonal therapy provides a highly tumor-specific form of therapy in sensitive tissues. In breast cancer, estrogen and progesterone receptor status is used to predict the success of hormonal therapy. Androgen receptor is also being pursued as a therapeutic target for breast cancer treatment.

TARGETED THERAPY

Over the past decade, increased understanding of cancer biology has fostered the emerging field of molecular therapeutics. The basic principle of molecular therapeutics is to exploit the molecular differences between normal cells and cancer cells to develop targeted therapies. Thus, targeted therapies usually are directed at the processes involved in tumor growth rather than directly targeting the tumor cells. The ideal molecular target would be exclusively expressed in the cancer cells, be the driving force of the proliferation of the cancer cells, and be critical to their survival. A large number of molecular targets are currently being explored, both preclinically and in clinical trials. The major groups of targeted therapy agents are inhibitors of growth factor receptors, inhibitors of intracellular signal transduction, cell-cycle inhibitors, apoptosis-based therapies, and anti-angiogenic compounds.

Protein kinases have come to the forefront as attractive therapeutic targets with the success of imatinib mesylate (Gleevec) in treating chronic myelogenous leukemia and GI stromal tumors, and trastuzumab (Herceptin) in treating breast cancer, and vemurafenib in treating melanoma. These drugs work by targeting bcr-abl and c-kit (imatinib) and HER2 and BRAF, respectively. For example, a phase 3 randomized trial demonstrated that, compared with dacarbazine, standard of care chemotherapy option for patients with metastatic melanoma with a V600E BRAF mutation, the BRAF inhibitor vemurafenib led to significantly higher response rates (48% vs. 5%). At 6 months, overall survival was 84% (95% CI, 78–89) in the vemurafenib group and 64% (95% CI, 56–73) in the dacarbazine group. The hazard ratio for tumor progression in the vemurafenib group was 0.26 (95% CI, 0.20–0.33; P<0.001). The estimated median progression-free survival was 5.3 months in the vemurafenib group and 1.6 months in the dacarbazine group. This trial highlights the fact that in at least some tumor types targeted therapies that inhibit a genomic alteration that is a driver is likely to be more effective than an unselected therapeutic option.

Sequencing of the human genome has revealed approximately 500 protein kinases. Several tyrosine kinases have been shown to have oncogenic properties and many other protein kinases have been shown to be aberrantly activated in cancer cells. Therefore, protein kinases involved in these aberrantly activated pathways are being aggressively pursued in molecular therapeutics. Potential targets like HER2 can be targeted via different strategies, such as transcriptional downregulation, targeting of mRNA, RNA inhibition, antisense strategies, direct inhibition of protein activity, and induction of immunity against the protein. Most of the compounds in development are monoclonal antibodies like trastuzumab or small-molecule kinase inhibitors like imatinib or vemurafenib. Some other agents, such as sunitinib, are multitargeted kinase inhibitors. Selected FDA-approved targeted therapies are listed in Table 10-11. Many of the promising pathways, such as the PI3K/Akt/mTOR pathway, are being pursued as therapeutic targets with several drugs in development, targeting different aspects of the pathway (Fig. 10-15).

Development of molecularly targeted agents for clinical use presents several unique challenges. Once an appropriate compound is identified and confirmed to have activity in preclinical testing, predictive markers for activity in the preclinical setting must be defined. Expression of a target may not be sufficient to predict response because the pathway of interest may not be activated or critical to the cancer’s survival. Although in traditional phase 1 trials the goal is to identify the maximum tolerated dosage, the maximum dosage of biologic agents may not be necessary to achieve the desired biologic effect. Thus, assays to verify modulation of the target need to be developed to determine at what dosage the desired effect is achieved. When phases 2 and 3 clinical trials are initiated, biomarker modulation studies should be integrated into the trial to determine whether clinical response correlates with target modulation and thus to identify additional parameters that impact response. Rational dose selection and limitation of study populations to patients most likely to respond to the molecular therapy as determined by predictive markers are most likely to lead to successful clinical translation of a product. Finally, most biologic agents are cytostatic, not cytotoxic. Thus, rational combination therapy mixing new biologic agents with either established chemotherapeutic agents that have synergy or with other biologic agents is more likely to lead to cancer cures.

IMMUNOTHERAPY

The aim of immunotherapy is to induce or potentiate inherent antitumor immunity that can destroy cancer cells. Central to the process of antitumor immunity is the ability of the immune system to recognize tumor-associated antigens present on human cancers and to direct cytotoxic responses through humoral or T-cell–mediated immunity. Overall, T-cell–mediated immunity appears to have the greater potential of the two for eradicating tumor cells. T cells recognize antigens on the surfaces of target cells as small peptides presented by class I and class II MHC molecules.

Several antitumor strategies are under investigation. One approach to antitumor immunity is nonspecific immunotherapy, which stimulates the immune system as a whole through administration of bacterial agents or their products, such as bacille
### Table 10-11
Selected FDA-approved targeted therapies

<table>
<thead>
<tr>
<th>GENERIC NAME</th>
<th>TRADE NAME</th>
<th>TARGET</th>
<th>FDA-APPROVED INDICATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ado-trastuzumab emtansine</td>
<td>Kadcyla</td>
<td>HER2</td>
<td>Breast cancer</td>
</tr>
<tr>
<td>Axitinib</td>
<td>Inlyta</td>
<td>KIT, FDGFR(\beta), VEGFR1/2/3</td>
<td>RCC</td>
</tr>
<tr>
<td>Bevacizumab</td>
<td>Avastin</td>
<td>VEGF</td>
<td>Colorectal cancer, lung cancer, glioblastoma, NSCLC, RCC</td>
</tr>
<tr>
<td>Bortezomib</td>
<td>Velcade</td>
<td>Proteasome</td>
<td>Myeloma</td>
</tr>
<tr>
<td>Bosutinib</td>
<td>Bosulif</td>
<td>ABL</td>
<td>CML (Philadelphia chromosome+)</td>
</tr>
<tr>
<td>Cabozantinib</td>
<td>Cometriq</td>
<td>FLT3, KIT, MET, RET, VEGFR2</td>
<td>Medullary thyroid cancer</td>
</tr>
<tr>
<td>Cetuximab</td>
<td>Erbitux</td>
<td>EGFR</td>
<td>Colorectal cancer (KRAS wild-type), Squamous cell cancer of the head and neck</td>
</tr>
<tr>
<td>Crizotinib</td>
<td>Xalkori</td>
<td>ALK (anaplastic lymphoma kinase) and ROS1 (c-ros oncogene 1) inhibitor</td>
<td>Non-small cell lung carcinoma</td>
</tr>
<tr>
<td>Dabrafenib</td>
<td>Tafinolar</td>
<td>BRAF V600E mutation</td>
<td>Melanoma</td>
</tr>
<tr>
<td>Dasatinib</td>
<td>Sprycel</td>
<td>ABL, src family, KIT, EPHA2, PDGFR-(\beta)</td>
<td>CML</td>
</tr>
<tr>
<td>Erlotinib</td>
<td>Tarceva</td>
<td>EGFR</td>
<td>NSCLC, Pancreatic cancer</td>
</tr>
<tr>
<td>Everolimus</td>
<td>Afinitor</td>
<td>mTOR</td>
<td>PNET, RCC, Breast cancer, Nonresectable subependymal giant cell astrocytoma associated with tuberous sclerosis</td>
</tr>
<tr>
<td>Gefitinib</td>
<td>Iressa</td>
<td>EGFR</td>
<td>NSCLC with known/previous benefit from gefitinib (limited approval)</td>
</tr>
<tr>
<td>Ibrutinib</td>
<td>Imbruvica</td>
<td>Bruton’s Tyrosine Kinase</td>
<td>Chronic lymphocytic leukemia</td>
</tr>
<tr>
<td>Imatinib</td>
<td>Gleevec</td>
<td>KIT, ABL, PDGFR</td>
<td>CML, GIST (KIT+), Dermatofibrosarcoma protuberans</td>
</tr>
<tr>
<td>Lapatinib</td>
<td>Tykerb</td>
<td>EGFR and HER2</td>
<td>Breast cancer (HER2+)</td>
</tr>
<tr>
<td>Nilotinib</td>
<td>Tasigna</td>
<td>ABL</td>
<td>CML (Philadelphia chromosome+)</td>
</tr>
<tr>
<td>Panitumumab</td>
<td>Vectibix</td>
<td>EGFR</td>
<td>Colorectal cancer (KRAS wild type)</td>
</tr>
<tr>
<td>Pazopanib</td>
<td>Votrient</td>
<td>VEGFR, PDGFR, KIT</td>
<td>RCC</td>
</tr>
<tr>
<td>Pertuzumab</td>
<td>Perjeta</td>
<td>HER2</td>
<td>Breast cancer (HER2+)</td>
</tr>
<tr>
<td>Ponatinib</td>
<td>Iclusig</td>
<td>ABL, FGFR1-3, FLT3, VEGFR2</td>
<td>CML, ALL (Philadelphia chromosome+)</td>
</tr>
<tr>
<td>Regorafenib</td>
<td>Stivarga</td>
<td>KIT, PDGFR(\beta), RAF, RET, VEGFR1/2/3</td>
<td>Colorectal cancer, GIST</td>
</tr>
<tr>
<td>Sorafenib</td>
<td>Nexavar</td>
<td>VEGFR, PDGFR, KIT, RAF</td>
<td>HCC, RCC</td>
</tr>
<tr>
<td>Sunitinib</td>
<td>Sutent</td>
<td>VEGFR PDGFR KIT, FIt-3, RET</td>
<td>GIST, RCC, PNET</td>
</tr>
<tr>
<td>Temsirolimus</td>
<td>Torisel</td>
<td>mTOR</td>
<td>RCC</td>
</tr>
<tr>
<td>Trastuzumab</td>
<td>Herceptin</td>
<td>HER2</td>
<td>Breast cancer (HER2+), Gastric cancer (HER2+)</td>
</tr>
<tr>
<td>Vandetanib</td>
<td>Caprelsa</td>
<td>EGFR, RET, VEGFR2</td>
<td>Medullary thyroid cancer</td>
</tr>
<tr>
<td>Vemurafenib</td>
<td>Zelboraf</td>
<td>BRAF</td>
<td>Melanoma (BRAF V600E mutant)</td>
</tr>
<tr>
<td>Vorinostat</td>
<td>Zolinza</td>
<td>Histone deacetylases</td>
<td>Cutaneous T-cell lymphoma</td>
</tr>
</tbody>
</table>

CML = chronic myelogenous leukemia; EGFR = epidermal growth factor receptor; EPHA2 = ephrin A2; FDA = Food and Drug Administration; Flt-3 = fms-related tyrosine kinase 3; GIST = GI stromal tumor; HCC = hepatocellular cancer, HER2 = human epidermal growth factor receptor 2; mTOR = mammalian target of rapamycin; NSCLC = non-small cell lung cancer, PDGF = platelet-derived growth factor; PDGFR = platelet-derived growth factor receptor; PNET = pancreatic neuroendocrine tumor; RCC = renal cell carcinoma; RET = rearranged during transfection; VEGF = vascular endothelial growth factor; VEGFR = vascular endothelial growth factor receptor.
Calmette-Guérin. This approach is thought to activate the effectors of antitumor response such as natural killer cells and macrophages, as well as polyclonal lymphocytes. Another approach to nonspecific immunotherapy is systemic administration of cytokines such as interleukin-2, interferon-α, and interferon-γ. Interleukin-2 stimulates proliferation of cytotoxic T lymphocytes and maturation of effectors such as natural killer cells into lymphokine-activated killer cells. Interferons, on the other hand, exert antitumor effects directly by inhibiting tumor cell proliferation and indirectly by activating host immune cells, including macrophages, dendritic cells, and natural killer cells, and by enhancing human leukocyte antigen (HLA) class I expression on tumor cells.

Antigen-specific immunotherapy can be active, as is achieved through antitumor vaccines, or passive. In passive immunotherapy, antibodies to specific tumor-associated antigens can be produced by hybridoma technique and then administered to patients whose cancers express these antigens, inducing antibody-dependent cellular cytotoxicity.

The early attempts at vaccination against cancers used allogeneic cultured cancer cells, including irradiated cells, cell lysates, and shed antigens isolated from tissue culture supernatants. An alternate strategy is the use of autologous tumor vaccines. These have the potential advantage of being more likely to contain antigens relevant for the individual patient but have the disadvantage of requiring a large amount of tumor tissue for preparation, which restricts eligibility of patients for this modality. Strategies to enhance immunogenicity of tumor cells include the introduction of genes encoding cytokines or chemokines, and fusion of the tumor cells to allogeneic MHC class II-bearing cells. Alternatively, heat shock proteins derived from a patient’s tumor can be used because heat shock protein peptide complexes are readily taken up by dendritic cells for presentation to T cells.

Identification of tumor antigens has made it possible to perform antigen-specific vaccination. For example, in the case of melanoma, several antigens have been identified that can be recognized by both CD8+ cytotoxic T cells and CD4+ helper T cells, including MART-1, gp 100, MAGE1, tyrosinase, TRP-1, TRP-2, and NY-ESO-1. Antigens tested usually are overexpressed or mutated in cancer cells. Tissue specificity and immunogenicity are important determinants in choosing an appropriate target. Vaccines directed at defined tumor antigens aim to combine selected tumor antigens and appropriate routes for delivering these antigens to the immune system to optimize antitumor immunity. Several different vaccination approaches have been studied, including tumor cell-based vaccines, peptide-based vaccines, recombinant virus-based vaccines, DNA-based vaccines, and dendritic cell vaccines.

In adoptive transfer, antigen-specific effector cells (i.e., cytotoxic T lymphocytes) or antigen-nonspecific effector cells
(i.e., natural killer cells) can be transferred to a patient. These effector cells can be obtained from the tumor (tumor-infiltrating lymphocytes) or the peripheral blood.

Clinical experience in patients with metastatic disease has shown objective tumor responses to a variety of immunotherapeutic modalities. It is thought, however, that the immune system is overwhelmed with the tumor burden in this setting, and thus adjuvant therapy may be preferable, with immunotherapy reserved for decreasing tumor recurrences. Trials to date suggest that immunotherapy is a potentially useful approach in the adjuvant setting. How to best select patients for this approach and how to integrate immunotherapy with other therapies are not well understood for most cancer types.

Tolerance to self-antigens expressed in tumors is a limitation in generating antitumor responses. Recently, several pathways that modulate tolerance and approaches to manipulating these pathways have been identified: pathways that activate professional antigen-presenting cells such as Toll-like receptors, growth factors, and the CD40 pathway; cytokines to enhance immunostimulation; and pathways that inhibit T-cell inhibitory signals or block the activity of immune-suppressive regulatory T cells (Tregs).

A new and highly effective strategy to activate the T-cell arm of anticancer immunity is the use of monoclonal antibodies to block inhibitory signaling pathways employed by the immune system to prevent T cell overactivation and the development of autoimmunity. CTLA-4 and PD-1 are two important inhibitory T-cell checkpoints that can be blocked with neutralizing antibodies and result in an effective antigen-specific anti-tumor response.

CTLA-4 is an inhibitory receptor expressed by activated T cells that belongs to the immunoglobulin superfamily. CTLA4 is related to the T-cell costimulatory receptor, CD28, and both are bound by CD80 and CD86 (also known as B7-1 and B7-2) which are expressed on antigen-presenting cells. CTLA4 conveys an inhibitory signal to the T cell, whereas engagement of CD28 with ligand sends a stimulatory signal. CTLA-4 is able to outcompete CD28 for CD80 and CD86 ligands and therefore is able to dominate immune signaling in the setting of antigen recognition. CTLA-4 is also expressed by regulatory T cells, which contributes to their ability to inhibit T-cell function.

Programmed death ligand 1 (PD-L1) is a 40 kDa type 1 transmembrane protein that is thought to play an important role in suppressing the immune system. PD-L1 binds to its receptor, PD-1, which is found on activated T cells. The PD1/PDL1 pathway is increasingly recognized as a key contributor to tumor-mediated immune suppression. The interaction between PD-1 leads to reduced proliferation, altered production of stimulatory cytokines, and reduced T-cell lytic activity. Thus, both anti-PD1 and anti-PD-L1 strategies are actively being pursued for cancer therapy.

The FDA-approved CTLA-4 blocking antibody ipilimumab has shown efficacy in patients with metastatic melanoma. Nivolumab and pembrolizumab are antibodies that target PD-1, whereas blockade of PD-L1 is accomplished with agents such as atezolizumab. Cancers for which checkpoint inhibitors have found utility include melanoma, renal cell carcinoma, bladder carcinoma, squamous cell carcinoma of the head and neck, and carcinoma of the lung. These agents produce durable shrinkage of advanced disease in 20% to 40% of patients, and combination strategies that employ checkpoint inhibitors with cytokines, vaccines, cellular therapies, and other targeted agents are under active investigation.

**GENE THERAPY**

Gene therapy is being pursued as a possible approach to modifying the genetic program of cancer cells as well as treating metabolic diseases. The field of cancer gene therapy uses a variety of strategies, ranging from replacement of mutated or deleted tumor-suppressor genes to enhancement of immune responses to cancer cells. Indeed, in preclinical models, approaches such as replacement of tumor-suppressor genes leads to growth arrest or apoptosis. However, the translation of these findings into clinically useful tools presents special challenges.

One of the main difficulties in getting gene therapy technology from the laboratory to the clinic is the lack of a perfect delivery system. An ideal vector would be administered through a noninvasive route and would transduce all of the cancer cells and none of the normal cells. Furthermore, the ideal vector would have a high degree of activity, that is, it would produce an adequate amount of the desired gene product to achieve target cell kill. Unlike genetic diseases in which delivery of the gene of interest into only a portion of the cells may be sufficient to achieve clinical effect, cancer requires either that the therapeutic gene be delivered to all of the cancer cells or that a therapeutic effect be achieved on nontransfected cells as well as transfected cells through a bystander effect. However, treatment of a metabolic disease requires prolonged gene expression, whereas transient expression may be sufficient for cancer therapy.

Several vector systems are under study for gene therapy; however, none is considered ideal. One of the promising approaches to increase the number of tumor cells transduced is the use of a replication-competent virus like a parovirus, human reovirus, or vesicular stomatitis virus that selectively replicates within malignant cells and lysed them more efficiently than it does normal cells. Another strategy for killing tumor cells with suicide genes exploits tumor-specific expression elements, such as the MUC-1, PSA, CEA, or VEGF promoters, that can be used to achieve tissue-specific or tumor-specific expression of the desired gene.

Because the goal in cancer therapy is to eradicate systemic disease, optimization of delivery systems is the key to success for gene therapy strategies. Gene therapy is likely to be most successful when combined with standard therapies, but it will provide the advantage of customization of therapy based on the molecular status of an individual’s tumor.

**MECHANISMS OF INTRINSIC AND ACQUIRED DRUG RESISTANCE**

Several tumor factors influence tumor cell kill. Tumors are heterogeneous, and, according to the Goldie-Coldman hypothesis, tumor cells are genetically unstable and tend to mutate to form different cell clones. This has been used as an argument for giving chemotherapy as soon as possible in treatment to reduce the likelihood that resistant clones will emerge. Tumor size is another important variable. Larger tumors may have greater heterogeneity, although heterogeneity may also differ based on biologic subtype. Tumor growth may be described by a Gompertz curve, named after Benjamin Gompertz, which has the form of a sigmoid function. Gompertzian models have thus been used to describe changes in tumor cell numbers over time where growth is slowest at the start and end of a time period, but are quite rapid in the middle. Theoretically, for any tumor, there is a period of time where cancer cells grow rapidly (exponential growth...
Cancer cells demonstrate adaptive responses to targeted therapy, like activating alternate pathways of survival; thus, these alterations may blunt therapeutic efficacy. Cancer cells also acquire resistance upon prolonged treatment with targeted therapy through a variety of mechanisms. One mechanism is through the loss of the target. For example, this was observed in a study of patients with HER2-positive breast cancer patients who were treated with neoadjuvant trastuzumab-based chemotherapy. Post neoadjuvant treatment, a third of the samples from patients who did not have a complete pathologic response displayed loss of the HER2 amplification that had been present in their pretreatment-biopsy specimens. Another means by which cancers develop resistance is the acquisition of additional genetic aberrations. In lung cancer, a second mutation in EGFR (T790M) and MET amplification have been described as two main mechanisms of drug resistance to EGFR inhibitors erlotinib and gefitinib. Other mechanisms like novel genetic changes, including HER2 and EGFR amplification, PIK3CA mutations, and markers of epithelial-to-mesenchymal transition have also been reported in EGFR inhibitor resistant lung. Analysis of metastases from patients with colorectal cancer who developed resistance to cetuximab or panitumumab showed the emergence of KRAS amplification in one sample and acquisition of secondary KRAS mutations in 60% of the cases. These studies emphasize the utility of repeat tumor biopsy specimens at the time of relapse or progression to identify mechanisms of resistance and best combinatorial therapies.

### RADIATION THERAPY

#### Physical Basis of Radiation Therapy

**Ionizing radiation** is energy strong enough to remove an orbital electron from an atom. This radiation can be electromagnetic, like a high-energy photon, or particulate, such as an electron, proton, neutron, or alpha particle. Radiation therapy is delivered primarily as high-energy photons (gamma rays and X-rays) and charged particles (electrons). Gamma rays are photons that are released from the nucleus of a radioactive atom. X-rays are photons that are created electronically, such as with a clinical linear accelerator. Currently, high-energy radiation is delivered to tumors primarily with linear accelerators. X-rays traverse the tissue, depositing the maximum dose beneath the surface, and thus spare the skin. Electrons are used to treat superficial skin lesions, superficial tumors, or surgical beds to a depth of 5 cm. Gamma rays typically are produced by radioactive sources used in brachytherapy.

The dose of radiation absorbed correlates with the energy of the beam. The basic unit is the amount of energy absorbed per unit of mass (joules per kilogram) and is known as a gray (Gy). One gray is equivalent to 100 rads, the unit of radiation measurement used in the past.

#### Biologic Basis of Radiation Therapy

Radiation deposition results in DNA damage manifested by single- and double-strand breaks in the sugar phosphate backbone of the DNA molecule. Cross-linking between the DNA strands and chromosomal proteins also occurs. The mechanism of DNA damage differs by the type of radiation delivered. Electromagnetic radiation is indirectly ionizing through the actions of short-lived hydroxyl radicals, which are produced primarily by the ionization of cellular hydrogen peroxide ($H_2O_2$). Protons and other heavy particles are directly ionizing and directly damage DNA.
Radiation damage is manifested primarily by the loss of cellular reproductive integrity. Most cell types do not show signs of radiation damage until they attempt to divide, so slowly proliferating tumors may persist for months and appear viable. Some cell types, however, undergo apoptosis.

The extent of DNA damage after radiation exposure is dependent on several factors. The most important of these is cellular oxygen. Hypoxic cells are significantly less radiosensitive than aerated cells. The presence of oxygen is thought to prolong the half-life of free radicals produced by the interaction of X-rays and cellular H$_2$O$_2$, and thus indirectly ionizing radiation is less efficacious in tumors with areas of hypoxia. In contrast, radiation damage from directly ionizing radiation is independent of cellular oxygen levels.

The extent of DNA damage from indirectly ionizing radiation is dependent on the phase of the cell cycle. The most radiation-sensitive phases are G$_1$ and M, whereas G$_2$ and late S phases are less sensitive. Thus, irradiation of a population of tumor cells results in killing of a greater proportion of cells in G$_2$ and M phases. However, delivery of radiation in divided doses, a concept referred to as fractionation, allows the surviving G$_1$ and S phase cells to progress to more sensitive phases, a process referred to as reassortment. In contrast to DNA damage after indirectly ionizing radiation, that after exposure to directly ionizing radiation is less dependent on the cell-cycle phase.

Several chemicals can modify the effects of ionizing radiation. These include hypoxic cell sensitizers such as metronidazole and misonidazole, which mimic oxygen and increase cell kill of hypoxic cells. A second category of radiation sensitizers are the thymidine analogues iododeoxyuridine and bromodeoxyuridine. These molecules are incorporated into the DNA in place of thymidine and render the cells more susceptible to radiation damage; however, they are associated with considerable acute toxicity. Several other chemotherapeutic agents sensitize cells to radiation through various mechanisms, including 5-fluorouracil, actinomycin D, gemcitabine, paclitaxel, topotecan, doxorubicin, and vinorelbine. The development of novel radiosensitizers is an active area of research and multiple small molecules as well as novel nanomaterials are under investigation.

Radiation Therapy Planning

Radiation therapy is delivered in a homogeneous dose to a well-defined region that includes tumor and/or surrounding tissue at risk for subclinical disease. The first step in planning is to define the target to be irradiated as well as the dose-limiting organs in the vicinity. Treatment planning includes evaluation of alternative treatment techniques, which is done through a process referred to as simulation. Once the beam distribution that will best achieve homogeneous delivery to the target volume and minimize the dose to the normal tissue is determined, immobilization devices and markings or tattoos on the patient’s skin are used to ensure that each daily treatment is given in the same way. Conventional fractionation is 1.8 to 2 Gy/d, administered 5 days each week for 3 to 7 weeks.

Radiation therapy may be used as the primary modality for palliation in certain patients with metastatic disease, primarily patients with bony metastases. In these cases, radiation is recommended for symptomatic metastases only. However, lytic metastases in weight-bearing bones such as the femur, tibia, or humerus also are considered for irradiation. Another circumstance in which radiation therapy might be appropriate is spinal cord compression due to metastases to the vertebral body that extend posteriorly to the spinal canal.

The goal of adjuvant radiation therapy is to decrease local-regional recurrence rates. Adjuvant radiation therapy can be given before surgery, after surgery, or in selected cases, during surgery. Preoperative radiation therapy has several advantages. It may minimize seeding of the tumor during surgery and it allows for smaller treatment fields because the operative bed has not been contaminated with tumor cells. Also, radiation therapy for inoperable tumors may achieve adequate reduction to make them operable. The disadvantages of preoperative therapy are an increased risk of postoperative wound healing problems and the difficulty in planning subsequent radiation therapy in patients who have positive surgical margins. If radiation therapy is given postoperatively, it is usually given 3 to 4 weeks after surgery to allow for wound healing. The advantage of postoperative radiation therapy is that the surgical specimen can be evaluated histologically and radiation therapy can be reserved for patients who are most likely to benefit from it. Further, the radiation therapy can be modified on the basis of margin status. The disadvantages of postoperative radiation therapy are that the volume of normal tissue requiring irradiation may be larger owing to surgical contamination of the tissue planes and that the tumor may be less sensitive to radiation owing to poor oxygenation. Postlaparotomy adhesions may decrease the mobility of the small bowel loops, increasing the risk for radiation injury in abdominal or pelvic irradiation. Given the potential advantages and disadvantages of both approaches, the roles of preoperative and postoperative radiation therapy are being actively evaluated and compared for many cancer types.

Another mode of postoperative radiation therapy is brachytherapy. In brachytherapy, unlike in external beam therapy, the radiation source is in contact with the tissue being irradiated. The radiation source may be cesium, gold, iridium, or radium. Brachytherapy is administered via temporary or permanent delivery implants such as needles, seeds, or catheters. Temporary brachytherapy catheters are placed either during open surgery or percutaneously soon after surgery. The implants are loaded interstitially, and treatment usually is given postoperatively for a short duration, such as 1 to 3 days. Although brachytherapy has the disadvantages of leaving scars at the catheter insertion site and requiring special facilities for inpatient brachytherapy, the advantage of patient convenience owing to the shorter treatment duration has made intracavitary treatment approaches popular for the treatment of breast cancer.

Another short delivery approach is intraoperative radiotherapy (IORT), often used in combination with external beam therapy. The oncologic consequences of the limited treatment volume and duration associated with brachytherapy and IORT are not well understood. Accelerated partial breast irradiation with interstitial brachytherapy, intracavitary brachytherapy (MammoSite), IORT, and three-dimensional conformal external beam radiotherapy is being compared with whole breast irradiation in an intergroup phase 3 trial (NSABP B-39/Radiation Therapy Oncology Group 0413). Several additional studies of adjuvant IORT also are ongoing internationally. There has also been increased interest in utilizing intensity-modulated radiation therapy (IMRT). IMRT is a complex technique for the delivery of radiation therapy preferentially to target structures while minimizing doses to adjacent normal critical structures. It is widely utilized for the treatment of a variety of tumor types, including
the central nervous system, head and neck, breast, prostate, gastrointestinal tract, and gynecologic organs, as well as in patients where previous radiation therapy has been delivered. Stereotactic radiosurgery uses extremely accurate image-guidance and patient positioning to deliver a high dose of radiation to a small tumor with well-defined margins. In this manner, the dose of radiation being applied to normal tissues can be minimized. It is most commonly used for the treatment of brain and spinal tumors. Protons are a charged particle that can be also used in external beam radiation therapy. Proton therapy employs a beam of protons as a means of delivering radiation to a tumor. In contrast to photons, which deposit energy continuously during their passage through tissue, protons deposit a large amount of their energy near the end of their path (known as the Bragg peak) and release less energy along the way. Thus, proton therapy could theoretically reduce the exposure of normal tissue to radiation, allowing the delivery of higher doses of radiation to a tumor. It is thought that chemoradiation given concurrently with radiation improves survival rates. Chemotherapy before radiation has the advantage of reducing the tumor burden, which facilitates radiation therapy. On the other hand, some chemotherapy regimens, when given concurrently with radiation, may sensitize the cells to radiation therapy. Chemoradiation is being investigated in many tumor types, including rectal cancer, pancreatic cancer, and esophageal cancer. In a Cochrane review of six randomized controlled trials, it was demonstrated that in patients with T3/4 rectal cancer, chemoradiation was associated with a significantly lower local recurrence rate compared with radiation therapy alone (OR 0.56, 95% CI 0.42–0.75, \( P < 0.0001 \)) but was not associated with improved survival.

**Side Effects**

Both tumor and normal tissue have radiation dose-response relationships that can be plotted as a sigmoidal curve (Fig. 10-16). A minimum dose of radiation must be given before any response is seen. The response to radiation then increases slowly with an increase in dose. At a certain dose level the curves become exponential, with increases in tumor response and normal tissue toxicity with each incremental dose increase. The side effects of radiation therapy can be acute, occurring during or 2 to 3 weeks after therapy, or chronic, occurring weeks to years after therapy. The side effects depend on the tissue included in the target volume. Some of the major acute and chronic sequelae of radiation are summarized in Table 10-13. In addition to these effects, a small increase in the risk for secondary malignancies is attributable to radiation therapy.

**CANCER PREVENTION**

The truth of the old axiom “An ounce of prevention is worth a pound of cure” is being increasingly recognized in oncology. Cancer prevention can be divided into three categories: (a) primary prevention (i.e., prevention of initial cancers in healthy individuals), (b) secondary prevention (i.e., prevention of cancer in individuals with premalignant conditions), and (c) tertiary prevention (i.e., prevention of second primary cancers in patients cured of their initial disease).

The systemic or local administration of therapeutic agents to prevent the development of cancer, called chemoprevention, is being actively explored for several cancer types. In breast cancer, the NSABP Breast Cancer Prevention Trial demonstrated that tamoxifen administration reduces the risk of breast cancer by one half and reduces the risk of estrogen receptor-positive tumors by 69% in high-risk patients. Therefore, tamoxifen has been approved by the FDA for breast cancer chemoprevention. The subsequent NSABP P-2 trial demonstrated that raloxifene is as effective as tamoxifen in reducing the risk of invasive breast cancer and is associated with a lower risk of thromboembolic events and cataracts but a nonstatistically significant higher risk of noninvasive breast cancer; these findings led the FDA to approve raloxifene for prevention as well. Several other agents are also under investigation. Celecoxib (a cyclooxygenase 2 [COX-2] inhibitor) has been shown to reduce polyp number and polyp burden in patients with FAP, which led to its approval by the FDA for these patients. However, celecoxib is no longer widely used as a primary preventative treatment in this setting due to the association between COX-2 inhibitors and coronary artery disease. In head and neck cancer, 13-cis-retinoic acid has been shown both to reverse oral leukoplakia and to reduce second primary tumor development. However, a large phase 3 clinical trial that utilized low-dose 13-cis-retinoic acid in patients with early stage squamous cell carcinoma of the head and neck showed no significant difference in the incidence of tumor recurrence or the second primary tumors between the placebo and chemoprevention arms. Thus, the chemoprevention trials completed so far have had mixed results. Much remains to be done over the next few years to improve outcomes and decrease therapy-related toxic effects. It is important for surgeons to be aware of these preventive options because they are likely to be involved in the diagnosis of premalignant and malignant conditions and will be the ones to counsel patients about their chemopreventive options.

In selected circumstances, the risk of cancer is high enough to justify surgical prevention. These high-risk settings include hereditary cancer syndromes such as hereditary breast-ovarian cancer syndrome, hereditary diffuse gastric cancer, multiple endocrine neoplasia type 2, FAP, and hereditary nonpolyposis colorectal cancer, as well as some nonhereditary...
conditions such as chronic ulcerative colitis. Most prophylactic surgeries are large ablative surgeries (e.g., bilateral risk-reducing mastectomy or total proctocolectomy). Therefore, it is important that the patient be completely informed about potential surgical complications as well as long-term lifestyle consequences. Further, the conservative options of close surveillance and chemoprevention need to be discussed. The patient’s cancer risk needs to be assessed accurately and implications for survival discussed. Ultimately, the decision to proceed with surgical prevention should be individualized and made with caution.

### TRENDS IN ONCOLOGY

#### Cancer Screening and Diagnosis

It is clear that the practice of oncology will change dramatically over the next few decades because our understanding of the molecular basis of cancer and available technologies are evolving rapidly. One of the critical changes expected is earlier detection of cancers. With improvements in available imaging modalities and development of newer functional imaging techniques, it is likely that many tumors will be detected at earlier, more curable stages in the near future.

Another area of rapid development is the identification of serum markers. High-throughput technologies such as matrix-assisted laser desorption ionization time-of-flight mass spectrometry and liquid chromatography ion-spray tandem mass spectrometry have revolutionized the field of proteomics and are now being used to compare the serum protein profiles of patients with cancer with those of individuals without cancer. Identification of unique proteins as well as unique proteomic profiles for most cancer types is being pursued actively by many researchers and, if successful, could dramatically enhance our ability to detect cancers early.

DNA fragments that are derived from tumors and are circulating in the blood stream are referred to as circulating tumor DNA (ctDNA). Analysis of ctDNA can potentially provide information on the entire tumor genome and has potential clinical utility as a so-called “liquid biopsy” when blood samples are obtained during important junctures of a clinical scenario. ctDNA may originate directly from the tumor or from circulating tumor cells, which refers to intact tumor cells that are shed from primary tumors and enter the bloodstream. The precise mechanism of ctDNA release has not been determined; however, there is evidence to show that the length of the DNA fragments are similar to those seen during the process of apoptosis. ctDNA can be reliably procured from peripheral blood and analyzed via a number of advanced techniques, including next generation sequencing. The main advantages of using ctDNA in genomic studies is the ability to obtain information on the entire tumor genome, thus avoiding the difficulties of tumor heterogeneity that are encountered with needle biopsies, and the ability to obtain multiple samples with much less risk to the patient.

#### Surgical Therapy

The current trend in surgery is toward more conservative resections. With earlier identification of tumors, more conservative operations may be possible. The goal, however, is always to remove the tumor en bloc with wide negative margins. Another interesting area being explored is the destruction of tumors by techniques such as radiofrequency ablation, cryoablation, and heat-producing technologies like lasers, microwaves, or focused ultrasound.

The debate over how to manage the regional lymph node basins for certain cancer types continues. With an increasing understanding of the metastatic process, surgeons may be able to stratify patients on the basis of the likelihood that their disease will spread metastatically, based on the gene expression profile of their primary tumors, and offer regional therapy accordingly. There is also a growing interest in minimally invasive surgical treatments for a variety of cancer types.

#### Systemic Therapy

The current trend in systemic therapy is toward individualized therapy. Therefore, the intent is to determine the underlying biology of each tumor to tailor therapy accordingly. Genomic, transcriptional, and proteomic profiling approaches are being used to identify molecular signatures that correlate

<table>
<thead>
<tr>
<th>ORGAN</th>
<th>ACUTE CHANGES</th>
<th>CHRONIC CHANGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin</td>
<td>Erythema, wet or dry desquamation, epilation</td>
<td>Telangiectasia, subcutaneous fibrosis, ulceration</td>
</tr>
<tr>
<td>GI tract</td>
<td>Nausea, diarrhea, edema, ulceration, hepatitis</td>
<td>Stricture, ulceration, perforation, hematochezia</td>
</tr>
<tr>
<td>Kidney</td>
<td>—</td>
<td>Nephropathy, renal insufficiency</td>
</tr>
<tr>
<td>Bladder</td>
<td>Dysuria</td>
<td>Hematuration, ulceration, perforation</td>
</tr>
<tr>
<td>Gonads</td>
<td>Sterility</td>
<td>Atrophy, ovarian failure</td>
</tr>
<tr>
<td>Hematopoietic tissue</td>
<td>Lymphopenia, neutropenia, thrombocytopenia</td>
<td>Pancytopenia</td>
</tr>
<tr>
<td>Bone</td>
<td>Epiphysyal growth arrest</td>
<td>Necrosis</td>
</tr>
<tr>
<td>Lung</td>
<td>Pneumonitis</td>
<td>Pulmonary fibrosis</td>
</tr>
<tr>
<td>Heart</td>
<td>—</td>
<td>Pericarditis, vascular damage</td>
</tr>
<tr>
<td>Upper aerodigestive tract</td>
<td>Mucositis, xerostomia, anosmia</td>
<td>Xerostomia, dental caries</td>
</tr>
<tr>
<td>Eye</td>
<td>Conjunctivitis</td>
<td>Cataract, keratitis, optic nerve atrophy</td>
</tr>
<tr>
<td>Nervous system</td>
<td>Cerebral edema</td>
<td>Necrosis, myelitis</td>
</tr>
</tbody>
</table>

Table 10-13

Local effects of radiation

<table>
<thead>
<tr>
<th>ORGAN</th>
<th>ACUTE CHANGES</th>
<th>CHRONIC CHANGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin</td>
<td>Erythema, wet or dry desquamation, epilation</td>
<td>Telangiectasia, subcutaneous fibrosis, ulceration</td>
</tr>
<tr>
<td>GI tract</td>
<td>Nausea, diarrhea, edema, ulceration, hepatitis</td>
<td>Stricture, ulceration, perforation, hematochezia</td>
</tr>
<tr>
<td>Kidney</td>
<td>—</td>
<td>Nephropathy, renal insufficiency</td>
</tr>
<tr>
<td>Bladder</td>
<td>Dysuria</td>
<td>Hematuration, ulceration, perforation</td>
</tr>
<tr>
<td>Gonads</td>
<td>Sterility</td>
<td>Atrophy, ovarian failure</td>
</tr>
<tr>
<td>Hematopoietic tissue</td>
<td>Lymphopenia, neutropenia, thrombocytopenia</td>
<td>Pancytopenia</td>
</tr>
<tr>
<td>Bone</td>
<td>Epiphysyal growth arrest</td>
<td>Necrosis</td>
</tr>
<tr>
<td>Lung</td>
<td>Pneumonitis</td>
<td>Pulmonary fibrosis</td>
</tr>
<tr>
<td>Heart</td>
<td>—</td>
<td>Pericarditis, vascular damage</td>
</tr>
<tr>
<td>Upper aerodigestive tract</td>
<td>Mucositis, xerostomia, anosmia</td>
<td>Xerostomia, dental caries</td>
</tr>
<tr>
<td>Eye</td>
<td>Conjunctivitis</td>
<td>Cataract, keratitis, optic nerve atrophy</td>
</tr>
<tr>
<td>Nervous system</td>
<td>Cerebral edema</td>
<td>Necrosis, myelitis</td>
</tr>
</tbody>
</table>
with response to certain agents. It is likely that in the near future all tumors can be tested and treatments individualized. Patients who will respond to conventional therapies can be treated with these regimens, whereas patients who will not respond will not, which spares them the toxicity. Instead, the latter patients can be offered novel therapies. Furthermore, with the advent of effective immune-based therapies, it is likely that patients may be given treatments that can specifically target the alterations driving tumor growth in combination with drugs that can enhance the anticancer immune response. Patients can be genotyped for critical alleles that may affect drug metabolism and thus, may influence the efficacy as well as the side effects of the drugs given. Finally, stratification of patients by gene expression profile for prognosis may assist in determining which patients are at higher risk of relapse so that patients whose tumors have less aggressive biologic characteristics can be spared further therapy.

REFERENCES

Entries highlighted in bright blue are key references.


BACKGROUND

Organ transplantation is a relatively novel field of medicine that has made significant progress since the second half of the 20th century. Advances in surgical technique and a better understanding of immunology are the two main reasons that transplants have evolved from experimental procedures, just several decades ago, to a widely accepted treatment today for patients with end-stage organ failure. Throughout the world, for a variety of indications, kidney, liver, pancreas, intestine, heart, and lung transplants are now the current standard of care.

But the success of transplantation has created new challenges. A better understanding of the pathophysiology of end-stage organ failure as well as advances in critical care medicine and in the treatment of various diseases led to expanding the criteria for, and decreasing the contraindications to, transplants. As a result, the discrepancy between the ever-growing number of patients awaiting a transplant and the limited number of organs available is one of the field’s biggest challenges (Fig. 11-1). In 2017 alone, according to the United Network for Organ Sharing (UNOS), about 115,000 patients in the United States were awaiting a transplant, yet the number of transplants performed approached only about 35,000 (Fig. 11-2).

DEFINITIONS

In addition to being the overall name of this relatively new field of medicine, transplantation is the process of transferring an organ, tissue, or cell from one place to another. An organ transplant is a surgical procedure in which a failing organ is replaced by a functioning one. The organ is transplanted either orthotopically (implanted in the same anatomic location in the recipient as it was in the donor) or heterotopically (implanted in...
Key Points

1. The field of transplantation has made tremendous advances in the last 50 years, mainly due to refinements in surgical technique and development of effective immunosuppressive medications.

2. Although immunosuppressive medications are essential for transplantation, they are associated with significant short- and long-term morbidity.

3. Opportunistic infections can be significantly lowered by the use of appropriate antimicrobial agents.

4. Kidney transplantation represents the treatment of choice for almost all patients with end-stage renal disease. The gap between demand (patients on the waiting list) and supply (available kidneys) continues to widen.

5. Pancreas transplantation represents the most reliable way to achieve euglycemia in patients with poorly controlled diabetes.

6. The results of islet transplantation continue to improve but still trail those of pancreas transplantation.

7. Liver transplantation has become the standard of care for many patients with end-stage liver failure and/or liver cancer.

Another anatomic location. Orthotopic transplants require the removal of the diseased organ (heart, lungs, liver, or intestine); in heterotopic transplants, the diseased organ is kept in place (kidney, pancreas).

According to the degree of immunologic similarity between the donor and recipient, transplants are divided into three main categories: (a) an autotransplant is the transfer of cells, tissue, or an organ from one part of the body to another part in the same person, so no immunosuppression is required; this type of transplant includes skin, artery or vein, bone, cartilage, nerve, and islet cell transplants; (b) an allotransplant is the transfer of cells, tissue, or an organ from one person to another of the same species; with the exception of identical twins, the immune system of the recipient recognizes the donated organ as a foreign body, so immunosuppression is required in order to avoid rejection; and (c) a xenotransplant is the transfer of cells, tissue, or an organ from one organism to another of a different species. To date, animal-to-human transplants are still experimental procedures, given the very complex immunologic and infectious issues that have yet to be solved.

HISTORY

Over the centuries, many different references to transplantation can be found in the world’s literature, yet transplantation as a recognized scientific and medical field began to emerge only in the middle of the 20th century. Two major events led to the rise of transplantation.

First, the surgical technique of the vascular anastomosis was developed by the French surgeon Alexis Carrel.1 This led to increased transplant activity, especially in animal models. Russian surgeon Yu Yu Voronoy was the first to report a series of human-to-human kidney transplants in the 1940s.2 But the outcomes were dismal, mainly because of the lack of understanding of the underlying immunologic processes.

Second, the findings of British scientist Sir Peter B. Medawar in the 1940s were also key.3 In his work with skin grafts in animal models and in human burn patients, he learned that the immune system plays a crucial role in the failure of skin grafts. His research led to a better understanding of the immune system and is considered to be the birth of transplant immunobiology.

The first human transplant with long-term success was performed by Joseph Murray in Boston, Massachusetts, in 1954.4 Because it was a living related kidney transplant between identical twins, no immunosuppression was required; the recipient lived for another 8 years before he died of issues unrelated to the transplanted kidney. Other centers performed similar transplants and could reproduce similar good results.

Ultimately, attempts were made to perform kidney transplants between nonidentical individuals. For immunosuppression, total-body radiation and an anticancer agent called # Waiting # Transplanted Figure 11-1. Patients on the waiting list and the number of organ transplants performed, 2000 to 2009. (U.S. data from the Scientific Registry of Transplant Recipients Annual Report, http://srtr.org)
6-mercaptopurine were used, but given the profound toxicity of both those methods of immunosuppression, results were discouraging. A breakthrough was achieved in the early 1960s with the introduction of maintenance immunosuppression through a combination of corticosteroids and a less toxic derivative of 6-mercaptopurine, azathioprine.\(^5,6\)

Increasing experience with kidney transplants and the better results achieved with maintenance immunosuppression paved the way for the era of nonrenal transplants (Table 11-1). In 1963, the first liver transplant was performed by Thomas Starzl in Denver, Colorado, and the first lung transplant was performed by James Hardy in Jackson, Mississippi. In 1966, the first pancreas transplant was performed by William Kelly and Richard Lillehei in Minneapolis, Minnesota. In 1967, the first successful heart transplant was performed by Christiaan Barnard in Cape Town, South Africa. The early years of transplantation were marked by high mortality, mainly because of irreversible rejection. However, dramatic advances occurred with the further development of new forms of immunosuppression. The groundbreaking event was the introduction of the first anti-T lymphocyte (T cell) drug, cyclosporine, in the early 1980s.\(^7\) Since then, with an even better understanding of immunologic processes, many other drugs have been introduced that target specific pathways that lead to rejection. As a result, rejection rates have decreased substantially, allowing a 1-year graft survival rate in excess of 80% in all types of transplants.

The gradual increase in the organ shortage led to innovative surgical techniques. For example, deceased donor split-liver transplants and living donor liver transplants have helped expand the liver donor pool. Similarly, living donor intestine and pancreas techniques have been developed. The evolution of donor nephrectomy from an open to a minimally invasive procedure (laparoscopic or robotic) has helped increase the pool of living kidney donors.

## TRANSPLANT IMMUNOBIOLOGY

The outcomes of early transplants were less than satisfactory. The limiting factor was the lack of understanding of immunologic processes, and irreversible rejection was the reason for graft loss in the vast majority of recipients. A better understanding of transplant immunobiology led to significant improvements in patient and graft survival rates.\(^8,9\) The immune system is designed as a defense system to protect the body from foreign pathogens, such as viruses, bacteria, and fungi, but it also acts to reject transplanted cells, tissues, and organs, recognizing them as foreign. It mediates other complex processes as well, such as the body’s response to trauma or to tumor growth. No matter what type of agent, the immune system recognizes it as foreign and triggers a strong response that is designed to either eradicate pathogenic organisms or reject foreign cells or tissue.

<table>
<thead>
<tr>
<th>Table 11-1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Transplant history</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ORGAN</th>
<th>YEAR</th>
<th>SURGEON</th>
<th>LOCATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kidney</td>
<td>1954</td>
<td>Joseph E. Murray</td>
<td>Boston, MA</td>
</tr>
<tr>
<td>Liver</td>
<td>1963</td>
<td>Thomas E. Starzl</td>
<td>Denver, CO</td>
</tr>
<tr>
<td>Lung</td>
<td>1963</td>
<td>James D. Hardy</td>
<td>Jackson, MS</td>
</tr>
<tr>
<td>Pancreas</td>
<td>1966</td>
<td>Richard C. Lillehei</td>
<td>Minneapolis, MN</td>
</tr>
<tr>
<td>Heart</td>
<td>1967</td>
<td>Christiaan N. Barnard</td>
<td>Cape Town, South Africa</td>
</tr>
<tr>
<td>Small intestine</td>
<td>1967</td>
<td>Richard C. Lillehei</td>
<td>Minneapolis, MN</td>
</tr>
<tr>
<td>Heart/lung</td>
<td>1981</td>
<td>Bruce Reitz</td>
<td>Stanford, CA</td>
</tr>
<tr>
<td>Multivisceral</td>
<td>1989</td>
<td>Thomas E. Starzl</td>
<td>Pittsburgh, PA</td>
</tr>
</tbody>
</table>
**TRANSPORT ANTIGENS**

Transplants between genetically nonidentical persons lead to recognition and rejection of the organ by the recipient’s immune system, if no intervention is undertaken. The main antigens responsible for this process are part of the major histocompatibility complex (MHC). In humans, these antigens make up the human leukocyte antigen (HLA) system. The antigen-encoding genes are located on chromosome 6. Two major classes of HLA antigens are recognized. They differ in their structure, function, and tissue distribution. Class I antigens (HLA-A, HLA-B, and HLA-C) are expressed by all nucleated cells. Class II antigens (HLA-DR, HLA-DP, and HLA-DQ) are expressed by antigen-presenting cells (APCs) such as B lymphocytes, dendritic cells, macrophages, and other phagocytic cells.

The principal function of HLA antigens is to present the fragments of foreign proteins to T lymphocytes. This leads to recognition and elimination of the foreign antigen with great specificity. HLA molecules play a crucial role in transplant recipients as well. They can trigger rejection of a graft via two different mechanisms. The most common mechanism is cellular rejection, in which the damage is caused by activated T lymphocytes. The process of activation and proliferation is triggered by exposure of T lymphocytes to the donor’s HLA molecules. The other mechanism is humoral rejection, in which the damage is mediated by circulating antibodies against the donor’s HLA molecules. The donor-specific antibodies can be present either pretransplant, due to previous exposure (because of a previous transplant, pregnancy, blood transfusion, or immunization), or posttransplant. After antibody binding to the donor’s HLA molecules, the complement cascade is activated, leading to cellular lysis.

**ALLORECOGNITION AND LYMPHOCYTE ACTIVATION**

The immune system of each person is designed to discriminate between self and nonself cells and tissues. This process is called allore cognition, with T cells playing the crucial role. The recognition of foreign HLA antigens by the recipient’s T cells may occur by either a direct or an indirect pathway. Direct recognition occurs when the recipient’s T cells are activated by direct interaction with the donor’s HLA molecules. Indirect recognition occurs when the recipient’s T cells are activated by interaction with APCs that have processed and presented the foreign antigen. The foreign antigen can be shed from the graft into the circulation, or it can be identified by the APCs within the graft itself.

Independent of the pathway of foreign HLA antigen presentation, the ensuing activation of T cells is similar and involves complex cell surface receptors and markers, i.e., the T-cell receptor (TCR) and an array of cluster differentiation markers (CDs). A two-signal model, T-cell activation begins with the engagement of the TCR/CD3 complex with the foreign molecule. This interaction causes transmission of the signal into the cell, namely signal 1. However, this signal alone is not sufficient to activate the T cell. An additional costimulatory signal is required, i.e., signal 2. Two well-characterized costimulatory interactions are the CD40/CD154 and B7/CD28 pathways. The “master switch” is turned on by the interaction of CD40 protein with APCs, along with the interaction of CD154 protein with T cells; this ligation induces the upregulation of other costimulatory molecules. Transmission of signal 1 and signal 2 into the cell nucleus leads to upregulation of the transcription of genes for several cytokines, including the T-cell growth factor interleukin-2 (IL-2). In turn, IL-2 activates a number of pathways, leading to proliferation and differentiation of T cells. Rejection is the result of an attack of activated T cells on the transplanted organ.

Although T-cell activation is the main culprit in rejection, B-cell activation and subsequent antibody production also play a role. After the foreign HLA antigen is processed by B cells, it interacts with activated helper T cells, leading to differentiation of B cells into plasma cells and subsequently to their proliferation and antibody production.

**CLINICAL REJECTION**

Graft rejection is due to a complex interaction of different components of the immune system, including B and T lymphocytes, APCs, and cytokines. The end result is graft damage caused by inflammatory injury. According to its onset and pathogenesis, rejection is divided into three main types: hyperacute, acute, and chronic, and each is described in the following sections.

**Hyperacute**

Hyperacute rejection, a very rapid type of rejection, results in irreversible damage and graft loss within minutes to hours after organ reperfusion. It is triggered by preformed antibodies against the donor’s HLA or ABO blood group antigens. These antibodies activate a series of events that result in diffuse intravascular coagulation, causing ischemic necrosis of the graft. Fortunately, pretransplant blood group typing and cross-matching (in which the donor’s cells are mixed with the recipient’s serum, and then the cells are observed for any destruction) have virtually eliminated the incidence of hyperacute rejection.

**Acute**

Acute rejection, the most common type of rejection, usually occurs within a few days or weeks posttransplant. According to the mechanism involved, it is further divided into cellular (T-cell–mediated) rejection, humoral (antibody-mediated) rejection, or a combination of both. The diagnosis is based on the results of biopsies of the transplanted organ, special immunologic stains, and laboratory tests (such as elevated creatinine levels in kidney transplant recipients, elevated liver test values in liver transplant recipients, and elevated levels of glucose, amylase, and lipase in pancreas transplant recipients).

**Chronic**

Chronic rejection occurs slowly and usually is progressive. It can manifest within the first year posttransplant but most often takes place gradually over several years. The mechanisms are not well understood, but the pathologic changes eventually lead to tissue fibrosis and loss of graft function. As advances in immunosuppression have diminished the incidence of acute rejection, this form of rejection is becoming more common.

**CLINICAL IMMUNOSUPPRESSION**

A successful transplant hinges upon a balance between the extent of the recipient’s immune response, the health and viability of the donor allograft, and pharmacologic immunosuppression. Immunosuppressive regimens are critical to graft and
patient survival posttransplant. Immunosuppression has evolved from the use of azathioprine and steroids in the 1960s and 1970s to the development in the 1980s of cyclosporine, the latter which markedly increased allograft survival. \(^2\) The introduction of tacrolimus and mycophenolate mofetil (MMF) in the 1990s further advanced the field of transplantation, enabling a variety of combinations to be used for immunosuppression often “tailored” for each recipient (Table 11-2).

Presently, immunosuppressants are used in multidrug regimens aimed at increasing efficacy by targeting multiple pathways to lower the immune response and to decrease the toxicity of individual agents. Certain regimens may involve withdrawal, avoidance, or minimization of certain classes of drugs. Transplant centers generally institute their immunosuppressive protocols based on experience, risk profiles, cost considerations, and outcomes. Immunosuppression is delivered in two phases: induction (starting immediately posttransplant, when the risk of rejection is highest) and maintenance (usually starting within days posttransplant and usually continuing for the life of the graft and the recipient). Thus, the degree of immunosuppression is highest in the first 3 to 6 months posttransplant; during this time, prophylaxis against a number of different bacterial, viral, or even antifungal opportunistic pathogens is administered. \(^12\) \(^13\)

A conventional immunosuppressive protocol might include (a) induction with anti-T-lymphocyte–depleting or nondepleting antibodies and (b) maintenance with calcineurin inhibitors, antiproliferative agents, and corticosteroids. Characteristics of the most common immunosuppressive agents are listed in Table 11-3.

**Table 11-2**

<table>
<thead>
<tr>
<th>Immunosuppressive drugs by grouping</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Immunophilin binders</strong></td>
</tr>
<tr>
<td>Calcineurin inhibitors</td>
</tr>
<tr>
<td>Cyclosporine</td>
</tr>
<tr>
<td>Tacrolimus</td>
</tr>
<tr>
<td>Noninhibitors of calcineurin</td>
</tr>
<tr>
<td>Sirolimus</td>
</tr>
<tr>
<td><strong>Antimetabolites</strong></td>
</tr>
<tr>
<td>Inhibitors of de novo purine synthesis</td>
</tr>
<tr>
<td>Azathioprine</td>
</tr>
<tr>
<td>Mycophenolate mofetil</td>
</tr>
<tr>
<td><strong>Biologic immunosuppression</strong></td>
</tr>
<tr>
<td>Polyclonal antibodies</td>
</tr>
<tr>
<td>Atgam</td>
</tr>
<tr>
<td>Antithymocyte immunoglobulin</td>
</tr>
<tr>
<td>Monoclonal antibodies</td>
</tr>
<tr>
<td>Muromonab-CD3</td>
</tr>
<tr>
<td>Basiliximab</td>
</tr>
<tr>
<td>Belatacept</td>
</tr>
<tr>
<td>Alemtuzumab</td>
</tr>
<tr>
<td>Rituximab</td>
</tr>
<tr>
<td>Bortezomib</td>
</tr>
<tr>
<td>Eculizumab</td>
</tr>
<tr>
<td><strong>Other</strong></td>
</tr>
<tr>
<td>Corticosteroids</td>
</tr>
</tbody>
</table>

**INDUCTION**

Induction includes the use of depleting (polyclonal) antibodies or nondepleting antibodies within the first month posttransplant. Studies have shown that induction with antibody regimens may prevent acute rejection, potentially leading to improved graft survival and the use of less maintenance immunosuppression.

**Depleting Antibodies**

Rabbit antithymocyte globulin (Thymoglobulin) is a purified gamma globulin obtained by immunizing rabbits with human thymocytes. Atgam, which has largely been replaced by Thymoglobulin, is a purified gamma globulin obtained by immunizing horses with human thymocytes. These agents contain antibodies to T cells and B lymphocytes (B cells), integrins, and other adhesion molecules, thereby resulting in rapid depletion of peripheral lymphocytes. Typically, the total dose of Thymoglobulin is roughly 6 mg/kg, a dose that has been shown to confer adequate lymphocyte depletion and better allograft survival. Doses of 3 mg/kg may not effectively prevent acute rejection, but more doses and prolonged administration increase the risk of infection and the potential occurrence of lymphoma. Thymoglobulin administration causes a cytokine release syndrome, so premedications (acetaminophen and diphenhydramine) are usually given. The principal side effects of Thymoglobulin include fever, chills, arthralgias, thrombocytopenia, leukopenia, and an increased incidence of a variety of infections. \(^14\) \(^15\)

**Nondepleting Antibodies**

Basiliximab (Simulect) is an anti-CD25 monoclonal antibody. The alpha subunit of the IL-2 receptor, also known as Tac or CD25, is found exclusively on activated T cells. Blockade of this component by this monoclonal antibody selectively prevents IL-2–induced T-cell activation. No lymphocyte depletion occurs with basiliximab; thus, it is not designed to be used to treat acute rejection. Its selectivity in blocking IL-2–mediated responses makes it a powerful induction agent without the added risks of infections, malignancies, or other major side effects. Currently, basiliximab is the only available anti-CD25 monoclonal antibody approved for clinical use. Usually, it is followed by the use of calcineurin inhibitors, corticosteroids, and MMF as maintenance immunosuppression. \(^16\)

Alemtuzumab (Campath, Lemtrada), another anti-CD52 monoclonal antibody, was initially used to treat chronic lymphocytic leukemia. The use of alemtuzumab has grown in the field of transplantation, given its profound lymphocyte-depleting effects. It causes cell death by complement-mediated cytolsis, antibody-mediated cytotoxicity, and apoptosis. One dose alone (30 mg) depletes 99% of lymphocytes. Monocyte recovery can be seen at 3 months posttransplant; B-cell recovery at 12 months; and T-cell recovery, albeit only to 50% of baseline, at 36 months. Alemtuzumab causes a significant cytokine release reaction and often requires premedications (steroids and antihistamines). Because of the long-lasting T-cell depletion, the risks of infection and posttransplant lymphoproliferative disorder remain. Initially, alemtuzumab was available only through a limited distribution program, but more recently has been studied in a number of clinical trials. \(^17\) \(^18\)

**MAINTENANCE**

**Corticosteroids**

Corticosteroids have had a role in immunosuppression since the beginning of the field of transplantation. Despite numerous
Table 11-3

Summary of the main immunosuppressive drugs

<table>
<thead>
<tr>
<th>DRUG</th>
<th>MECHANISM OF ACTION</th>
<th>ADVERSE EFFECTS</th>
<th>CLINICAL USES</th>
<th>DOSAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyclosporine (CSA)</td>
<td>Binds to cyclophilin Inhibits calcineurin and IL-2 synthesis</td>
<td>Nephrotoxicity Tremor Hypertension Hirsutism</td>
<td>Improved bioavailability of microemulsion form</td>
<td>Oral dose 5 mg/kg per day (given in two divided doses)</td>
</tr>
<tr>
<td>Tacrolimus (FK506)</td>
<td>Binds to FKBP Inhibits calcineurin and IL-2 synthesis</td>
<td>Nephrotoxicity Hypertension Neurotoxicity GI toxicity (nausea, diarrhea)</td>
<td>Improved patient and graft survival in (liver) primary immunosuppression and rescue therapy Used as mainstay of maintenance protocols</td>
<td>IV 0.015 mg/kg per day as continuous infusion PO 0.05 mg/kg per day (given every 12 h)</td>
</tr>
<tr>
<td>Mycophenolate mofetil</td>
<td>Antimetabolite Inhibits enzyme necessary for de novo purine synthesis</td>
<td>Leukopenia GI toxicity</td>
<td>Effective for primary immunosuppression in combination with tacrolimus</td>
<td>1 g bid PO</td>
</tr>
<tr>
<td>Sirolimus</td>
<td>Inhibits lymphocyte effects driven by IL-2 receptor</td>
<td>Thrombocytopenia Increased serum cholesterol/LDL Poor wound healing</td>
<td>May allow early withdrawal of steroids and decreased calcineurin doses</td>
<td>2–4 mg/d, adjusted to trough drug levels</td>
</tr>
<tr>
<td>Corticosteroids</td>
<td>Multiple actions Anti-inflammatory Inhibits lymphokine production</td>
<td>Cushingoid state Glucose intolerance Osteoporosis</td>
<td>Used in induction, maintenance, and treatment of acute rejection</td>
<td>Varies from milligrams to several grams per day Maintenance doses, 5–10 mg/d</td>
</tr>
<tr>
<td>Azathioprine</td>
<td>Antimetabolite Interferes with DNA and RNA synthesis</td>
<td>Thrombocytopenia Neutropenia Liver dysfunction</td>
<td>Used in maintenance protocols or if intolerance to mycophenolate mofetil</td>
<td>1–3 mg/kg per day for maintenance</td>
</tr>
<tr>
<td>Belatacept</td>
<td>T-cell blocker</td>
<td>Increased risk of bacterial infections</td>
<td>New drug for maintenance immunosuppression in renal transplants only</td>
<td>5–10 mg/kg per day infusion</td>
</tr>
</tbody>
</table>

FKBP = FK506-binding protein; GI = gastrointestinal; IL = interleukin; IV = intravenous; LDL = low-density lipoprotein; PO = oral

Attempts to limit or discontinue their use, they remain an integral component of most immunosuppressive protocols, for both induction and maintenance. Moreover, they are often the first-line agents in the treatment of acute rejection. Steroids bind to glucocorticoid-responsive elements in DNA that prevent the transcription of cytokine genes and cytokine receptors. In addition, steroids have an impact on lymphocyte depletion, on the transcription of cytokine genes and cytokine receptors. In glucocorticoid-responsive elements in DNA that prevent calcineurin activity and decreases antibody production. It has been used as a first-line agent in transplant recipients for more than 40 years, but it became an adjunctive agent after the introduction of cyclosporine. With the development of newer agents such as MMF, the use of AZA has decreased significantly. However, it is preferred in recipients who are considering conceiving a child because MMF is teratogenic and can cause birth defects. Use of AZA remains an option for recipients who cannot tolerate the gastrointestinal (GI) side effects of MMF.

**Azathioprine**

An antimetabolite, azathioprine (AZA) is converted to 6-mercaptopurine and inhibits both the de novo purine synthesis and salvage purine synthesis. AZA decreases T-lymphocyte activity and decreases antibody production. It has been used as a first-line agent in transplant recipients for more than 40 years, but it became an adjunctive agent after the introduction of cyclosporine. With the development of newer agents such as MMF, the use of AZA has decreased significantly. However, it is preferred in recipients who are considering conceiving a child because MMF is teratogenic and can cause birth defects. Use of AZA remains an option for recipients who cannot tolerate the gastrointestinal (GI) side effects of MMF.

The most significant side effect of AZA, often dose-related, is bone marrow suppression. Leukopenia is often reversible with dose reduction or temporary cessation of the drug. Other significant side effects include hepatotoxicity, pancreatitis, neoplasia, anemia, and pulmonary fibrosis. Its most significant drug interaction is with allopurinol, which blocks AZA metabolism, increasing the risk of pancytopenia. Recommendations are to not use AZA and allopurinol together, or if doing so is unavoidable, to decrease the dose of AZA by 75%.
Mycophenolate Mofetil

Approved in May 1995 by the U.S. Food and Drug Administration (FDA) for preventing acute rejection after kidney transplants, MMF has now been incorporated into routine maintenance regimens after many solid organ transplants. Mycophenolate is the prodrug of mycophenolate acid, derived from Penicillium fungi. Mycophenolate acid is an inhibitor of inosine monophosphate dehydrogenase (IMPDH) involved in the de novo pathway of purine synthesis. MMF is available in capsules (250 and 500 mg); the starting dose is 1 g twice daily. In hopes of decreasing the GI side effects, an enteric-coated formulation called Myfortic was developed; its benefits have not been clearly demonstrated in studies, but in some conversion studies patients did report less GI intolerance. The pharmacokinetics of MMF are complex; mycophenolic acid (MPA) levels are not routinely performed at most transplant centers. Studies have shown that MPA levels and the incidence of rejection are not significantly correlated. The most common side effects of MMF are GI in nature, most commonly diarrhea, nausea, dyspepsia, and bloating. Esophagitis and gastritis occur in roughly 5% of recipients and may represent a cytomegalovirus (CMV) or herpesvirus family infection. The other important side effects are leukopenia, anemia, and thrombocytopenia (Table 11-4). Leukopenia can sometimes be reversed by lowering the MMF dose and discontinuing other agents like valganciclovir. MMF does not have any significant drug interactions, but clinicians should be careful to avoid additive toxicities with other medications that might lead to leukopenia and thrombocytopenia.

Sirolimus

The first mammalian target of rapamycin (mTOR) inhibitors to enter clinical use was sirolimus (Rapamune). A key regulatory kinase, mTOR changes cells from the G1 to S phase in the cell cycle, in response to proliferation signals provided by cytokines like IL-2. The mTOR inhibitors bind to FK506-binding protein (FKBP), and the sirolimus-FKBP complex binds to mTOR. Sirolimus also inhibits proliferation of vascular smooth muscle cells, possibly easing the vasculopathy and progressive fibrosis that can affect allografts. Sirolimus is a substrate for CYP3A4/4 and has many significant drug interactions (see Table 11-4).

To date, sirolimus has been used in a variety of combinations for maintenance immunosuppression, alone or in

<table>
<thead>
<tr>
<th>Table 11-4</th>
<th>Side effects and drug interactions of the main immunosuppressive drugs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>COMMON SIDE EFFECTS</strong></td>
<td><strong>OTHER MEDICATIONS THAT INCREASE BLOOD LEVELS</strong></td>
</tr>
<tr>
<td>Cyclosporine (CSA)</td>
<td>Hypertension, nephrotoxicity, hirsutism, neurotoxicity, gingival hyperplasia, hypomagnesemia, hyperkalemia</td>
</tr>
<tr>
<td>Tacrolimus (FK506)</td>
<td>Hypertension, nephrotoxicity, alopecia, hyperglycemia, neurotoxicity, hypomagnesemia, hyperkalemia</td>
</tr>
<tr>
<td>Sirolimus</td>
<td>Thrombocytopenia and neutropenia, elevated cholesterol, extremity edema, impaired wound healing</td>
</tr>
<tr>
<td>Mycophenolate mofetil</td>
<td>Leukopenia, thrombocytopenia, GI upset</td>
</tr>
<tr>
<td>Corticosteroids</td>
<td>Hyperglycemia, osteoporosis, cataracts, myopathy, weight gain</td>
</tr>
<tr>
<td>Azathioprine</td>
<td>Leukopenia, anemia, thrombocytopenia, neoplasia, hepatitis, cholestasis</td>
</tr>
</tbody>
</table>

ACE-I = angiotensin-converting enzyme inhibitor; ARB = angiotensin receptor blocker; NSAID = nonsteroidal anti-inflammatory drug; TMP-SMX = trimethoprim-sulfamethoxazole
conjunction with one of the calcineurin inhibitors. In such combinations, sirolimus usually is used to help withdraw from, or completely avoid, the use of steroids. It also has been used as an alternative to tacrolimus or cyclosporine, in a calcineurin-sparing protocol. One of the most significant side effects of sirolimus is hypertriglyceridemia, a condition that may be resistant to statins and fibrates. Impaired wound healing (immediately posttransplant in particular), thrombocytopenia, leukopenia, and anemia also are associated with sirolimus, and these problems are exacerbated when it is used in combination with MMF.25,26

### Cyclosporine

The introduction of cyclosporine in the early 1980s dramatically altered the field of transplantation by significantly improving outcomes after kidney transplantation. Cyclosporine binds with its cytoplasmic receptor protein, cyclophilin, which subsequently inhibits the activity of calcineurin, thereby decreasing the expression of several critical T-cell activation genes, the most important being for IL-2. As a result, T-cell activation is suppressed.27

Many formulations of cyclosporine exist, so it is important to know which one the transplant recipient is taking. Sandimmune, an older, oil-based formulation, has poor bioavailability and variable absorption. The newer formulations, Gengraf and Neoral, are microemulsified with improved bioavailability. Cyclosporine can be given intravenously or orally to maintain trough levels of 250 to 350 ng/mL for the first 3 months posttransplant; then it can be tapered to 150 to 250 ng/mL.28

The metabolism of cyclosporine is via the cytochrome P450 system, resulting in many significant drug interactions (see Table 11-4). Calcineurin inhibitors are nephrotoxic and constrict the afferent arteriole in a dose-dependent, reversible manner (Table 11-5). They also can cause hyperkalemia and hypomagnesemia. Several neurologic complications, including headaches, tremor, and seizures, also have been reported.29

Cyclosporine has several undesirable cosmetic effects, including hirsutism and gingival hyperplasia. It is associated with a higher incidence of hypertension and hyperlipidemia than is tacrolimus.

### Tacrolimus

The calcineurin inhibitor tacrolimus (Prograf) is now the backbone of most immunosuppressive regimens. Tacrolimus acts by binding FKBP, causing roughly 10 to 100 times more potent inhibition of IL-2 production than cyclosporine (which acts by binding cyclophilins). It can be given intravenously, orally, or sublingually to maintain trough levels of 8 to 12 ng/mL for the first 3 months posttransplant; then it can be tapered to 6 to 10 ng/mL. The metabolism of tacrolimus is via the cytochrome P450 system, resulting in many significant drug interactions (see Table 11-4).

Tacrolimus causes a higher incidence of new-onset diabetes posttransplant than does cyclosporine. Other side effects include alopecia, nephrotoxicity, neurotoxicity, hypertension, hyperkalemia, hypomagnesemia, and an increased incidence of certain types of infection.30

### Belatacept

The best-characterized pathway of T-cell costimulation includes CD28; its homologue, the cytotoxic T-lymphocyte–associated protein 4 (CTLA4); and their ligands, CD80 and CD86. Belatacept (also known as LEA29Y) was developed through two amino acid substitutions to abatacept (also known as CTLA4-Ig), a fusion protein consisting of the extracellular domain of CTLA4 and the Fc domain of immunoglobulin G (IgG). It is a high-avidity molecule with slower dissociation rates.

Clinical trials have compared the use of belatacept vs. a standard cyclosporine protocol in recipients of living donor, deceased donor, and extended-criteria donor kidneys. Belatacept was not inferior to cyclosporine in both patient and allograft survival rates, but was associated with a higher rate of biopsy-proven acute cellular rejection.

In terms of adverse effects, belatacept differs from standard calcineurin-based regimens because of an increased risk of posttransplant lymphoproliferative disorder (PTLD); the greatest risk is in recipients who are Epstein-Barr virus (EBV)-seronegative pretransplant. The FDA recommends the use of belatacept only in seropositive recipients. Studies in liver transplant recipients were halted early because of increased mortality rates.

However, belatacept does have a lower incidence of cardiovascular risk factors including metabolic lipid disorders, hypertension, neurotoxicity, glucose abnormalities, and adverse cosmetic effects. Except for the increased risk of malignancy, the more favorable adverse effect profile of belatacept and its convenient monthly dosing schedule may make it an attractive option for maintenance of immunosuppression, possibly improving compliance.31,32

### HUMORAL REJECTION

#### Rituximab

A chimeric anti-CD20 (anti-B cell) monoclonal antibody, rituximab is currently FDA approved for treating several types of lymphoma. The CD20 antigen is expressed early in the B-cell cycle but is absent on mature plasma cells. The variable region binds to CD20 through three different mechanisms: (a) antibody-dependent cell cytotoxicity, (b) complement-dependent cell killing, and (c) induction of apoptotic cell death. The use of

---

**Table 11-5**

<table>
<thead>
<tr>
<th>Drug interactions and side effects associated with calcineurin inhibitors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INTERACTIONS</strong></td>
</tr>
<tr>
<td>Inhibition of metabolism</td>
</tr>
<tr>
<td>Induction of metabolism</td>
</tr>
<tr>
<td>Hyperkalemia</td>
</tr>
<tr>
<td>Nephrotoxicity</td>
</tr>
</tbody>
</table>
rituximab has grown to include the treatment of antibody-mediated rejection and use in desensitization protocols. Studies so far have been small, with rituximab usually used in conjunction with plasmapheresis, steroids, and intravenous immunoglobulin (IVIG).\textsuperscript{33,35}

**Bortezomib**

A proteasome inhibitor, bortezomib is FDA approved for treating multiple myeloma. It can directly target plasma cells. Traditional treatments have been successful in removing antibodies, inhibiting antibody activity, or lowering antibody production; however, targeting mature antibody production in plasma cells has not met with success. Bortezomib has been shown to cause apoptosis of normal plasma cells, thereby decreasing alloantibody production in sensitized patients. Several case reports and series have described the use of bortezomib for the treatment of antibody-mediated rejection and in desensitization protocols.\textsuperscript{34,36,37}

**Eculizumab**

A humanized monoclonal antibody with high affinity for C5, eculizumab is a first-in-class, FDA-approved agent for treating paroxysmal nocturnal hemoglobinuria, hemolytic uremic syndrome, and generalized myasthenia gravis. It blocks the activation of the terminal complement cascade. Most antibody-mediated rejection episodes are associated with early complement activation as evidenced on renal transplant biopsies by the presence of C4d\textsuperscript{+} staining of the peritubular capillaries. Given its highly selective mechanism of action, this agent is predicted to be useful to treat antibody-mediated rejection and to desensitize patients pretransplant. However, its serious adverse effects include an increased risk of infections, especially due to encapsulated bacteria such as Neisseria meningitidis. Patients should be immunized with meningococcal vaccine at least 2 weeks before the administration of eculizumab.\textsuperscript{34,38,39}

**INFECTIONS AND MALIGNANCIES**

Advances in immunosuppression have led to improved graft survival rates. However, the growing population of immunosuppressed patients, in turn, has led to an increased incidence of opportunistic infections and malignancies. Such posttransplant complications have become important barriers to long-term disease-free survival.

**Infections**

Transplant recipients are predisposed to a variety of infections. Immunosuppression is the obvious reason. Moreover, such patients have already endured end-stage organ disease pretransplant and then the stress of an invasive transplant operation. Posttransplant, they continue to have significant comorbid conditions.

**Early.** Early infections (i.e., infections occurring within 1 month posttransplant) can be due to a wide spectrum of pathogens (bacterial, viral, and fungal). In the immediate postoperative period, recipients are significantly compromised from the stress of the operation, from induction immunosuppression, and often from initially impaired graft function. Infections during this period can be devastating.

It is imperative to differentiate between medical and surgical infections. Surgical infections are the most common and require expedient surgical intervention. Typical examples include generalized peritonitis, intra-abdominal abscesses, and wound infections.

In liver and pancreas recipients, surgical infections are most severe. The incidence of intra-abdominal infections is decreasing, but they remain a significant problem: they are the second most common reason (after vascular thrombosis) for graft loss in pancreas recipients.

Lengthy operations with significant blood loss, prolonged warm and cold ischemic times, and spillage of contaminated fluid (bile, urine, or bowel contents) predispose patients to intra-abdominal infections. Other prominent risk factors are the high level of induction immunosuppression immediately posttransplant and anastomotic leaks. Furthermore, pretransplant infections can reemerge or worsen.

The signs and symptoms of intra-abdominal infections are those of peritonitis: fever, hypotension, ileus, and abdominal pain, although the latter can be masked by immunosuppression. Treatment entails a prompt return to the operating room. Intra-abdominal infections are usually polymicrobial, involving several bacterial and fungal species. Common bacterial isolates include Escherichia coli, as well as Enterococcus, Klebsiella, and Pseudomonas species. Common fungal isolates are Candida albicans, Candida krusei, and Candida glabrata. Localized infections or abscesses can be treated with percutaneous drainage and antibiotics.

Medical infections include respiratory, urinary tract, and bloodstream infections. Medical treatment should also be aggressive, often including empiric antibiotics and antifungal medications even before culture results are available. Recipients of organs from infected donors should be treated per the results of donor culture speciation and the antibiotic sensitivity profile.

**Late.** Late infections primarily are due to chronic immunosuppression, specifically the depression of cell-mediated immunity that renders recipients susceptible to viruses, fungi, and parasites.

Members of the herpesvirus group are the most common etiologic agents of viral infections posttransplantation, with herpes simplex virus (HSV), CMV, and EBV being the most prominent. Pretransplant exposure to viruses may confer immunity. Recipients who are seronegative for HSV, CMV, and/or EBV have a higher incidence of posttransplant infections, especially if they receive donor allografts from seropositive donors.

CMV is a latent infection that can be transmitted to seronegative recipients by donor organs from seropositive individuals, can reactivate during immunosuppression, or both. Infections usually occur 3 to 6 months posttransplant or during treatment for rejection. The incidence of CMV has been greatly reduced with 12-week acyclovir prophylaxis.\textsuperscript{40} CMV infections range from an asymptomatic or mild flu-like syndrome to tissue-invasive disease resulting in pneumonitis, hepatitis, and GI ulcerations. Symptomatic infections and all tissue-invasive CMV disease should be treated with intravenous (IV) ganciclovir, a reduction in immunosuppression, or both, although successful treatment of mild to moderate rejection and concurrent mild to moderate CMV disease has been described.

EBV infections range from a mild mononucleosis syndrome to severe hepatitis and highly morbid PTLD. PTLD ranges from a localized tumor to a progressive, diffuse infiltration of various organs including the brain. It results from the proliferation of EBV-positive B cells in immunosuppressed patients. The main risk factors are a high degree of immunosuppression and...
a predisposing EBV serostatus (seronaive recipient, seropositive donor). Among patients with early lesions, the first line of treatment is to reduce immunosuppression. For those with more advanced PTLD, rituximab is used.

After 6 months posttransplant, the risk of invasive fungal infections is closely associated with environmental exposures. Blastomyces dermatitidis grows in moist soil in the Midwest and Southeast regions of the United States. Diagnosis is confirmed by biopsy; the preferred treatment is IV amphotericin B.

Coccidioides immitis can cause invasive coccidioidomycosis after inhalation of aerosolized infectious particles. It is endemic in the Southwest, Northern Mexico, and various parts of Central and South America. This infection can be resilient and difficult to treat. The first line of treatment is high-dose amphotericin B.

Histoplasma capsulatum is found in chicken, pidgeon, and bat droppings in the Ohio River and Mississippi River valleys. Dissemination is commonplace; up to a quarter of patients have central nervous system (CNS) involvement. Treatment consists of prolonged (3 to 13 months) administration of oral itraconazole.

Opportunistic infections with Aspergillus, Cryptococcus, Mucor, and Rhizopus species are rare but can cause serious infections. Patients with invasive Candida or Aspergillus infections exhibit a 20% mortality rate. Prophylaxis with fluconazole has been shown to reduce invasive fungal infections in liver recipients.41

Pneumocystis jiroveci (also known as PCP) is ubiquitous and can cause pulmonary disease in immunocompromised patients. However, trimethoprim-sulfamethoxazole (TMP-SMX) is effective prophylaxis against PCP, and daily, lifelong administration has virtually eliminated this infection among transplant recipients.

Malignancies
Chronic immunosuppression increases the risk of developing certain types of malignancies. The most extensive data, from a cohort study involving more than 175,000 solid organ transplant recipients, showed that 10,656 of them developed malignancies. The standardized incidence ratio was 2:10 (as compared with the general population). Recipients had at least a fivefold increase (as compared with the general population) in these types of malignancies: Kaposi’s sarcoma, nonmelanoma skin cancer, non-Hodgkin’s lymphoma, and cancer of the liver, anus, vulva, and lip. In addition, recipients had a statistically significant increase (as compared with the general population) in melanoma, Hodgkin’s lymphoma, and cancer of the lung, kidney, colon, rectum, and pancreas.42

ORGAN PROCUREMENT AND PRESERVATION
Organ procurement is a key element in organ transplantation. Currently, over 100 organ procurement organizations (OPOs) exist in the United States, all members of the Organ Procurement and Transplantation Network (OPTN), which is a federally mandated network created by and overseen by UNOS. Each OPO is responsible for evaluating and procuring deceased donor organs for transplantation in a specific geographic region. Hospitals receiving any type of federal reimbursement for their services (whether transplant-related or not) are required to report all deaths to their OPO in a timely manner. Each OPO then determines the medical suitability of the deceased for organ donation; requests consent for donation from family members; if consent is given, contacts the OPTN to analyze and identify potential recipients whose HLA antigens most closely match those of the donor; and arranges for the recovery and transport of any donated organs.

Strategies to increase organ donation and utilization have been successfully implemented in the last 10 to 15 years. The nationwide “Organ Donation Breakthrough Collaborative,” sponsored by the U.S. Department of Health and Human Services in 2003, brought the OPOs and transplant communities into a single concerted program to develop best practices guidelines. However, a severe donor shortage remains. The number of living organ donors peaked in 2007 and has declined since.

Alternative options include tissue engineering and stem cell research, but those fields are in their infancy in terms of producing fully functional and vascularized human organs. With the development of genetic “knockout” pigs, xenotransplantation still shows promise, but two problems in particular—immunologic barriers and xenosis (also known as zoonosis) of endogenous porcine retroviruses—have yet to be satisfactorily addressed.

Today, the gap between patients waiting for organ transplants and the number of organs available continues to widen. More than 118,000 patients are on the waiting list for solid organ transplants, but only 33,611 transplants were performed in 2016.

Deceased Donors
Most transplants today utilize organs from deceased donors. Formerly, death was determined by the cessation of both cardiac and respiratory function.

Donation After Brain Death. In 1968, the concept of “irreversible coma” was introduced by an ad hoc committee report at Harvard Medical School; that concept was pivotal to the final acceptance, in 1981, of “brain death” as a legal definition in the United States. The legal language states that the declaration of brain death should be in accordance with acceptable medical standards but does not specify clinical methodology. It is customary for hospitals to establish their own policies to declare brain death, according to their standards of care and local regulations.

Typically, brain death is defined as the irreversible cessation of brain function, including the brainstem. The presence of medical conditions that mimic brain death—such as drug overdose, medication side effects, severe hypothermia, hypoglycemia, induced coma, and chronic vegetative state—need to be excluded. The latest evidence-based guideline on determining brain death in adults reaffirmed the validity of current clinical practice.43 Briefly, the clinical diagnosis of brain death consists of four essential steps: (a) establishment of the proximate cause of the neurologic insult; (b) clinical examinations to determine coma, absence of brainstem reflexes, and apnea; (c) utilization of ancillary tests, such as electroencephalography (EEG), cerebral angiography, or nuclear scans, in patients who do not meet clinical criteria; and (d) appropriate documentation. A similar guideline on determining brain death in pediatric patients was recently developed.44

Once the diagnosis of brain death has been established, the local OPO assumes the care of the potential donor and initiates the process of donor evaluation and organ donation, and the potential donor is screened for contraindications to donation. The medical history and social history are obtained from the available family members. A battery of tests, including serologic
or molecular detection of human immunodeficiency virus (HIV) and viral hepatitis, are performed. The exact medical conditions that preclude donation vary; nonetheless, in the United States, infections and other medical conditions that determine eligibility are dictated by UNOS bylaws and routinely reviewed and updated.

The OPO focuses on preserving organ function and optimizing peripheral oxygen delivery until organ procurement commences. In all deceased donors, core temperature, systemic arterial blood pressure, arterial oxygen saturation, and urine output must be determined routinely and frequently. Arterial blood gases, serum electrolytes, blood urea nitrogen, serum creatinine, liver enzyme, hemoglobin, and coagulation tests need to be monitored regularly. In all brain-dead donors, elevated intracranial pressure triggers a compensatory catecholamine response to maintain cerebral perfusion pressure. Ischemic injury to the spinal cord and the sympathetic system may lead to a profound vasodilation. As a result, brain-dead donors frequently have severe hemodynamic and metabolic derangements, so aggressive monitoring and intervention are required to prevent loss of precious organs.

Previous studies of deceased donor care focused on organ-specific resuscitation protocols that resulted in only marginal gains in the number of organs transplanted. The latest developments center on multisystem protocols to increase the number of organs transplanted per donor (OTPD). The goals are to maintain a core temperature between 36.0°C and 37.5°C, a mean arterial pressure >70 mmHg or a systolic pressure >100 mmHg, and a hemoglobin level between 7 and 10 g/dL; hormonal therapy and aggressive treatment of arrhythmias and metabolic derangements are also called for.

Surgical Technique. Procurement of multiple organs (heart, lungs, kidney, liver, pancreas, and/or small bowel), or multivisceral procurement, was first described by the Pittsburgh group in 1987. Since then, most centers have incorporated changes, especially with regard to the timing and location of dissection and flushing. The basic steps involve a long incision to provide wide exposure of all thoracic and abdominal organs (Fig. 11-3). A Cattell-Braasch maneuver (complete mobilization of the distal small bowel, right colon, and duodenum) is performed to allow for identification of the distal aorta, iliac bifurcation, and distal inferior vena cava (IVC). The infrarenal aorta is the site for inserting the cannula that will allow for flushing of the organs with cold preservation solution. Sometimes, division of the inferior mesenteric artery is necessary to facilitate the exposure of the distal aorta. The third portion of the duodenum is retracted cephalad to expose the root of the superior mesenteric artery (SMA). Limited dissection is performed at the root of the SMA, which is encircled with a vessel loop to enable its temporary occlusion at the time of flushing, thus reducing the incidence of overperfusion injury to the pancreas.

A large anomalous or replaced right hepatic artery typically rises from the SMA, and this should be identified and preserved. Lateral to the SMA is the inferior mesenteric vein (IMV), which can be cannulated for portal flushing. Dissection of the hepatic hilum and the pancreas should be limited to the common hepatic artery (CHA), and branches of the CHA (e.g., splenic, left gastric, and gastroduodenal arteries) are exposed. The gastrohepatic ligament is carefully examined to preserve a large anomalous or replaced left hepatic artery, if present. The supraceliac aorta can be exposed by dividing the left triangular ligament of the liver and the gastrohepatic ligament.

The common bile duct is transected at the superior margin of the head of the pancreas. The gallbladder is incised and flushed with ice-cold saline to clear the bile and sludge. If the pancreas is to be procured, the duodenum is flushed with antimicrobial solution. Before the cannulation of the distal aorta, systemic heparinization (300 units/kg) is administered. The supraceliac aorta is clamped; cold preservation fluid is infused via the aortic (systemic) and IMV (portal) cannulas. The thoracic organs, liver, pancreas, and kidneys are then removed.

Donation After Cardiac Death. Given the severe shortage of donor organs, donation after cardiac death (DCD)—also known as donation by non–heart-beating donors (NHBDs)—was introduced to the transplant community in the 1990s. The category of DCD (Maastricht classification) was initially proposed at an international workshop and is now widely adopted for organ procurement. Currently, most NHBDs in the United States meet Maastricht classification III; that is, they have suffered a devastating injury with no chance of a meaningful recovery but do not meet the criteria for brain death. After consent for donation is obtained from the next of kin, the donor’s life support is removed. After the cessation of cardiac and respiratory function, organ procurement commences. DCD procurement protocols vary between states; religious and cultural differences need to be taken into consideration. The surgical team must be familiar with, and respect, the local protocol.

With cardiac death (as opposed to brain death), warm ischemic injury to organs can occur during the period between circulatory cessation and rapid core cooling through perfusion of preservation solution. However, the difference in long-term outcomes is negligible for recipients of organs from either type of donor. Still, a significant percentage of liver grafts procured after cardiac death, especially those with more than 25 minutes of warm ischemic time, develop devastating ischemic cholangiopathy and fail.

A new development to minimize ischemic injury to organs procured after cardiac death has been the application of extracorporeal membrane oxygenation (ECMO). With ECMO, DCD differs in two key ways: (a) cannulation occurs before withdrawal of life support and (b) organs are perfused via ECMO with warm oxygenated blood after declaration of cardiac death.

Figure 11-3. Exposure for thoracic and abdominal organ procurement.
The initial experience with organs procured using ECMO has been encouraging.

**Surgical Technique.** Surgeons who perform multiple organ retrieval should be familiar and experienced with the super-rapid technique described by the Pittsburgh group. Preferably, NHBDs undergo withdrawal of life support in the operating room after the surgical site is prepped and draped, as soon as the surgical team is ready. Alternatively, the NHBD is transported to the operating room after declaration of cardiac death.

A midline incision is used to rapidly gain entry into the abdominal cavity. An assistant retracts the small bowel and the sigmoid colon laterally, so that the bifurcation of the aorta can be easily identified on the left side of the vertebral column. A short segment of the distal aorta is dissected out from the retroperitoneum. A moist umbilical tape is passed around the aorta, which is used to secure a cannula. The distal aorta is clamped. Next, a cannula is passed cephalad through an aortotomy and secured. Flushing with cold preservation solution is started at once, followed by cross-clamping the aorta proximally (thoracic aorta) and venting through the vena cava. The portal flush is then instituted.

The rest of the procedure is similar to procurement after brain death, with two noticeable differences. First, to avoid injury to a large anomalous or replaced left hepatic artery, the gastrohepatic ligament and the left gastric artery are separated from the stomach at the lesser curvature. Second, to avoid injury to a large anomalous or replaced right hepatic artery, the SMA is examined before it is divided. If the pancreas is not procured, a common aortic patch encompassing both the SMA and the celiac artery can be procured with the liver.

**Living Donors**

The maxim of medical ethics is “primum non nocere” (first, do no harm), and for that reason, living organ donation presents unique ethical and legal challenges. Performing potentially harmful operations to remove organs from healthy individuals seems, at first glance, to contradict that maxim. But in fact, the ethical framework of living organ donation rests on three guiding principles respected in all discussions of medical practice: beneficence to the recipient, nonmaleficence to the donor, and the donor’s right to autonomy. In order to achieve optimal outcomes (the common good), transplant professionals should focus on maximizing the benefits for the recipient and minimizing the damage to the donor. The Uniform Anatomical Gift Act adopted by all states in the United States (with slight variations) provides the legal framework for competent adult living donors to decide whether or not to donate. It is the fiduciary duty of transplant professionals to explain the risks of organ donation. Any decision to donate should be uncoerced, and no enticements should be offered.

The use of living donors offers numerous advantages for recipients in need. First and foremost is the availability of lifesaving organs for those who would otherwise succumb to the progression of their end-stage disease. In certain parts of the world, such as East Asia, the concept of brain death and the use of deceased donors conflict with the prevailing culture or religion. Even in countries where the use of deceased donors is accepted, the use of living donors may significantly shorten the waiting time for recipients. A shorter waiting time generally implies a healthier recipient—one whose body has not been ravaged by prolonged end-stage organ failure. Moreover, with the use of living donors, transplants are planned (rather than emergency) procedures, allowing for better preoperative preparation of the recipient. Receiving an organ from a closely matched relative may also have immunologic benefits. And long-term results may be superior with the use of living donors, as is certainly the case with kidney transplants.

The major disadvantage is the risk to the living donor. Medically, there is no possibility of benefit to the donor, only the potential for harm. The risk of death associated with donation depends on the organ being removed. For a nephrectomy, the estimated mortality risk is less than 0.05%; for a partial hepatectomy, about 0.2%. The risk of surgical and medical complications also depends on the procedure being performed. In addition, long-term complications may be associated with a partial loss of organ function after donation. The guiding principle should be minimization of risk to the donor. All potential risks must be carefully explained to the potential donor, and written informed consent must be obtained.

**Surgical Technique.** The kidney, the first organ to be transplanted from living donors, is still the most common organ donated by these individuals. The donor’s left kidney is usually preferable because of the long vascular pedicle. Use of living donor kidneys with multiple renal arteries should be avoided in order to decrease the complexity of the vascular reconstruction and to help avoid graft thrombosis. Most donor nephrectomies are now performed via minimally invasive techniques, that is, laparoscopically, whether hand-assisted or not. With laparoscopic techniques, an intraperitoneal approach is most common: it involves mobilizing the colon, isolating the ureter and renal vessels, mobilizing the kidney, dividing the renal vessels and the distal ureter, and removing the kidney (Fig. 11-4). Extensive dissection around the ureter should be avoided, and the surgeon should strive to preserve as much length of the renal artery and vein as possible.

Liver transplants with living donors are not as commonly performed, given the significantly higher rates of donor mortality and morbidity. Initially, only adult donors for pediatric recipients were selected, but now, living donor liver transplants also involve adult donors for adult recipients. In dual graft living donor liver transplants, segmental grafts from two living donors augment the recipient’s graft size. The donor hepatectomy is similar to a major lobar hepatectomy, except that it is important to preserve the integrity of the vascular structure until graft resection (Fig. 11-5).

Living donor transplants of organs other than the kidney and liver are fairly uncommon, but certain centers do perform such transplants. Living donor pancreas transplants involve performing a distal pancreatectomy, with the graft consisting of the body and tail of the pancreas; vascular inflow and outflow are provided by the splenic artery and splenic vein. Living donor intestinal transplants usually involve removal of about 200 cm of the donor’s ileum, with inflow and outflow provided by the ileocolic vessels. Living donor lung transplants involve removal of one lobe of one lung from each of two donors; both grafts are then transplanted into the recipient.

**Organ Preservation**

The development and continuing refinement of organ preservation methods have completely revolutionized the transplant field. Extending the time that organs can be safely stored after procurement has enabled better organ utilization and better recipient outcomes. Hypothermia and pharmacologic inhibition are the two most frequent methods. Both slow—yet cannot
Figure 11-4. Laparoscopic left donor nephroureterectomy. A. Takedown of splenic flexure of colon to expose the left renal hilum. B. Dissection of left ureter off the psoas muscle. C. Dissection of left renal vein and gonadal vein. Left ureter seen lateral to the dissection. D. Dissection of left renal artery. Lumbar veins clipped and divided. E. Endo-TA stapler transection of the left renal artery. F. Placement of ports and Pfannenstiel incision for the donor kidney extraction.

Figure 11-5. Donor hepatectomy (right hepatectomy). A. The liver parenchymal transection line (c, the Cantlie line) marked with cautery. Right portal vein (p) and right hepatic artery (a) isolated. b = bile duct. Cystic duct was cannulated for intraoperative cholangiography. B. Exposure of hepatic veins after transection of the parenchyma. IVC = inferior vena cava; L = left hepatic vein; M = middle hepatic vein; R = right hepatic vein.
completely shut down—the removed organ’s metabolic activity, so both have adverse effects, such as cellular swelling and degradation. Cold storage solutions were introduced to mitigate some of the adverse effects of hypothermia or pharmacologic inhibition alone. Such solutions help prevent cellular swelling and the loss of cellular potassium.

One, and perhaps the most effective, preservation solution was developed at the University of Wisconsin and remains in wide use. Its ingredients include lactobionate (which helps prevent cellular swelling and reperfusion injury), raffinose, and hydroxyethyl starch (which helps reduce swelling of endothelial cells, thereby decreasing edema). Histidine-tryptophan-ketoglutarate solution is also currently in wide use.

Despite enhancements in preservation methods, the amount of time that an organ can be safely stored remains relatively short (hours, not days), particularly with organs from marginal donors. Among kidney recipients, delayed graft function becomes significantly more frequent after cold ischemic times of more than 24 hours, necessitating temporary dialysis, which is associated with increased risks of graft loss and higher costs. Among liver recipients, primary nonfunction and biliary complications ensue after prolonged cold ischemic times. In the case of heart and lung recipients, ischemic times should be under 6 hours. All of those times assume the use of normal donors.

There is revived interest in the use of the pulsatile perfusion pump, a kidney graft preservation method that has been available for more than 40 years. With the increasing shortage of available donor organs and the rise in the use of organs after cardiac death, the pulsatile perfusion pump is garnering renewed enthusiasm as an adjunct method of preservation, even for donor organs other than kidneys.

**KIDNEY TRANSPLANTATION**

**Introduction**

Ullman reported the first attempted human kidney transplant in 1902. For the next 50 years, sporadic attempts all ended in either technical failure or in graft failure from rejection. Joseph Murray performed the first successful kidney transplant in 1954, an epochal event in the history of organ transplantation. In that first case, the immunologic barrier was circumvented by transplanting a kidney between identical twins. For his pivotal contribution, Murray shared the Nobel Prize in Physiology or Medicine in 1990 with E. Donnall Thomas for their discoveries concerning “organ and cell transplantation in the treatment of human disease.”

The introduction of AZA (Imuran) in 1960 marked the beginning of a new era in kidney transplantation. With the combination of steroids and AZA for maintenance immunosuppression, the 1-year graft survival rate with a living related donor kidney approached 80%; with a deceased donor kidney, the rate was 65%. In the ensuing years, major milestones included the introduction of more effective immunosuppressive medications with lower toxicity profiles, such as polyclonal antilymphocyte globulin in the 1970s, cyclosporine in the 1980s, tacrolimus in the 1990s, and biologics in the first decade of the 21st century, as previously mentioned.

Parallel to the developments in medical science were the transplant community’s concerted efforts to improve use of healthcare resources. In the United States, the Social Security amendments of 1972 provided Medicare coverage for patients with end-stage renal disease (ESRD). The National Organ Transplant Act of 1984 initiated the process of creating what later became UNOS, an umbrella organization to ensure access to organs by patients in need, to enhance organ procurement and allocation, and to improve posttransplant outcomes. This infrastructure later became the blueprint for other countries to follow. As a result, organ transplantation is the most transparent field of medicine. Data such as transplant center performance are readily available on public websites; penalties for violation of regulations and for underperformance often result in transplant programs being shut down.

Today, a kidney transplant remains the most definitive and durable renal replacement therapy for patients with ESRD. It offers better survival and improved quality of life and is considerably more cost-effective than dialysis. According to the 2016 Scientific Registry of Transplant Recipients (SRTR) annual report, nearly 100,000 adult patients were on the kidney transplant waiting list, while nearly 20,000 patients underwent renal transplantation. Trends over the past decade indicated that living related transplants remained relatively stable, while the number of deceased donor transplants rose. Posttransplant outcomes have continued to improve: in 2015, the 1-year graft survival rate with a living donor kidney was nearly 98%; with a deceased donor kidney, the rate was approximately 95.0%.

The advantages of a living donor kidney transplant include better posttransplant outcomes, avoidance of prolonged waiting time and dialysis, and the ability to coordinate the donor and recipient procedures in a timely fashion. Living donor kidney recipients enjoy better long-term outcomes, a low incidence of delayed graft function, and reduced risks of posttransplant complications. Furthermore, the elective nature of living donor kidney transplants provides unique opportunities for recipient desensitization treatment if the donor and recipient are ABO-incompatible or if the HLA cross-match results are positive.

Some of the challenges transplant professionals face today are closing the growing gap between supply and demand and thereby reducing the current prolonged waiting times; refining immunosuppressive medications to achieve better outcomes with reduced toxicity; and caring for patients who develop rejection, especially antibody-mediated rejection.

**Pretransplant Evaluation**

Active infection or the presence of a malignancy, active substance abuse, and poorly controlled psychiatric illness are the few absolute contraindications to a kidney transplant. Studies have demonstrated the overwhelming benefits of kidney transplants in terms of patient survival, quality of life, and cost-effectiveness, so most patients with ESRD are referred for consideration of a kidney transplant. However, to achieve optimal transplant outcomes, the many risks (such as the surgical stress to the cardiovascular system, the development of infections or malignancies with long-term immunosuppression, and the psychosocial and financial impacts on compliance) must be carefully balanced.

Any problems detected during the evaluation of transplant candidates are communicated to their referring physician and/or to a specialist if advanced evaluation and treatment are needed, ultimately improving overall care. Essentially, the pretransplant evaluation is a multifaceted approach to patient education and disease management.

Before the pretransplant medical evaluation begins, kidney transplant candidates are encouraged to attend a group meeting.
focused on patient education. The meeting is coordinated by a transplant physician or surgeon. The intent is to familiarize patients with the pretransplant evaluation process and with pertinent medical concepts and terms. In an open forum format, important decisions such as type of donor (living vs. deceased) are discussed. The group meeting empowers patients to fully participate in their care and serves as an impetus for a meaningful dialogue with healthcare professionals.

Medical Evaluation

Cardiovascular Disease. Diabetes and hypertension are the leading causes of chronic renal disease. Concomitant cardiovascular disease (CVD) is a common finding in this population. An estimated 30% to 42% of deaths with a functioning kidney graft are due to CVD.72,73 Therefore, assessment of the potential kidney transplant candidate’s cardiovascular status is an important part of the pretransplant evaluation.

In fact, the American Heart Association and the American College of Cardiology Foundation recently published their expert consensus on CVD evaluation and management for solid organ transplant candidates.74 The process should focus on careful screening for the presence of significant cardiac conditions (e.g., angina, valvular disease, and arrhythmias) and for a prior history of congestive heart failure, coronary interventions, or valvular surgery. The perioperative risk assessment is based on patient symptoms and exercise tolerance. For all kidney transplant candidates, a resting 12-lead electrocardiogram (ECG) should be obtained. In addition, in this population, the use of echocardiography to analyze left ventricular function and to assess for pulmonary hypertension is useful.

Stress testing may be considered in patients with no active cardiac condition but with risk factors such as diabetes, hemodialysis for more than 1 year, left ventricular hypertrophy, age greater than 60 years, smoking, hypertension, and dyslipidemia. The utility of noninvasive stress testing (as compared with angiographic studies) for evaluating coronary artery disease is controversial; an additional prognostic marker is the troponin T (cTN T) level.

Malignancies. Because of the long-term use of immunosuppressive medications, transplant recipients are at increased risk for development of malignancies. Untreated and/or active malignancies are absolute contraindications to a transplant (with two exceptions: nonmelanocytic skin cancer and incidental renal cell cancer identified at the time of concurrent nephrectomy [i.e., for polycystic kidney disease] and renal transplantation). For most patients who have undergone treatment of low-grade tumors with a low risk of recurrence (e.g., completely locally excised low-grade squamous cell cancer of the skin, colon cancer in a polyp absent stalk invasion), a wait of at least 2 years after successful treatment is recommended before a kidney transplant can be considered. However, for certain types of tumors, especially at advanced stages or those with a high risk of recurrence (e.g., melanoma, lymphoma, renal cell cancer, breast cancer, colon cancer), a delay of at least 5 years is advisable. According to the Israel Penn International Transplant Tumor Registry, tumor recurrence posttransplant is not infrequent: the recurrence rate is 67% in patients with multiple myeloma, 53% in nonmelanocytic skin cancer, 29% in bladder cancer, and 23% in breast cancer.75

Infections. A thorough history of infections and immunizations should be obtained from transplant candidates, who need all recommended age-appropriate vaccinations according to the Centers for Disease Control and Prevention (CDC) guidelines. Ideally, vaccinations should be completed at least 4 to 6 weeks before the kidney transplant takes place. Immunosuppressive medications blunt the immune response and reduce the effectiveness of vaccinations; even more important, with attenuated vaccines, vaccine-derived infections could occur. If a splenectomy is anticipated (e.g., in recipients whose donor is ABO-incompatible or whose HLA cross-match results are positive), then they should be immunized against encapsulated organisms (such as Neisseria meningitidis, Haemophilus influenzae, and Streptococcus pneumoniae) well in advance of the splenectomy.

Transplant candidates should undergo routine tuberculosis (TB) screening. According to the Centers for Disease Control (CDC), in 2016, 9272 TB cases were diagnosed in the United States with 68.5% of cases occurring in foreign-born persons.76 Serologic screening combined with a chest roentgenogram for fungal infections such as coccidioidomycosis or histoplasmosis, in patients who either have a history of those infections or are from an endemic area, are recommended. Chronic infections such as osteomyelitis or endocarditis must be fully treated; a suitable waiting period after successful treatment must occur, in order to ensure that relapse does not occur.

Hepatitis can be caused by five different type of viruses, hepatitis virus A, B, C, D, and E, with the first three being the most common. Acute viral hepatitis is a contraindication to a kidney transplant; however, chronic viral hepatitis (most commonly caused by hepatitis B [HBV] or C [HCV]) does not preclude a recipient from undergoing a kidney transplant. In such candidates, obtaining a liver biopsy is essential to assess the disease severity. Recipients infected with HBV should undergo antiviral treatment (e.g., lamivudine) to prevent reactivation and progression of liver disease. Note that HBV is a noncytopathic virus; the liver damage is the result of an immune-mediated process.77 Moreover, the presence of normal liver enzymes in patients with HBV antigenemia does not predict the severity of parenchymal damage.

Transplant candidates with chronic HCV infection often have HCV-related glomerulonephritis. As with HBV infection, the clinical presentation and biochemical findings with HCV infection are often unreliable in predicting liver damage. In such patients who also exhibit evidence of cirrhosis, a combined liver-kidney transplant should be considered. In appropriate candidates, pretransplant antiviral treatment with interferon-α may be considered. However, after a kidney transplant, interferon treatment is not recommended because it is an immunosuppressive agent, and thus HIV may precipitate graft rejection.

Thanks to the excellent outcomes of highly active antiretroviral therapy (HAART), infection with HIV is no longer considered a contraindication to a kidney transplant. Kidney transplant candidates with HIV must have an undetectable HIV viral load and a CD4 lymphocyte count greater than 200/mm3; in addition, they must not have developed any opportunistic infection in the previous year.78 Latent viral infections such as CMV and EBV are of particular interest in the field of transplantation, given the risks of reactivation posttransplant and their detrimental effects on graft and patient survival. Knowing the CMV and EBV serologic status of the recipient and donor helps transplant professionals gauge the risk of immunosuppressive regimens in relation to potential infection, thereby guiding plans for posttransplant antiviral prophylaxis treatment or, as noted earlier,
avoiding transplants between a seropositive donor and a seronegative recipient.

**Kidney Disease.** The third most common cause of graft loss in kidney transplant recipients is recurrence of glomerular diseases such as focal segmental glomerulosclerosis (FSGS), immunoglobulin A (IgA) nephropathy, hemolytic uremic syndrome, systemic lupus erythematosus, and membranoproliferative glomerulonephritis. FSGS deserves special mention due to its frequent occurrence and dramatic presentation of early graft loss. An estimated 30% to 40% of FSGS patients develop recurrent disease posttransplant; of those, up to half eventually lose their graft. In recipients with a history of FSGS, posttransplant nephrotic proteinuria should be promptly investigated; if diagnosis is confirmed by transplant kidney biopsy, rescue plasmapheresis should be instituted at once. Adjuvant therapy with rituximab has been proposed.

**Hypercoagulopathy.** Kidney transplant candidates with a history of thrombotic events, repeated miscarriages, or a family history of thrombophilia should be screened for the following coagulopathic disorders: activated protein C resistance ratio, factor V Leiden mutation, factor II 20210 gene mutation, antiphospholipid antibody, lupus anticoagulant, protein C or S deficiency, antithrombin III deficiency, and hyperhomocysteinemia. In recipients at risk for hypercoagulopathy, pediatric kidney grafts and any kidney allografts with a complex vascular anatomy should be avoided. A perioperative anticoagulation protocol is recommended in this population.

**Surgical Evaluation**

**Urologic Evaluation.** Kidney transplant candidates (pediatric patients, in particular) with chronic kidney disease as a result of congenital or genitourinary abnormalities should undergo a thorough urologic evaluation. A voiding cystourethrogram and a complete lower urinary tract evaluation to rule out outlet obstruction are essential. Indications for a native nephrectomy include chronic pyelonephritis, large polycystic kidneys with loss of intra-abdominal domain, significant vesicoureteral reflux, or uncontrollable renovascular hypertension.

**Vascular Evaluation.** The potential implant sites for a kidney graft include the recipient’s iliac vessels and, less commonly, the aorta and vena cava. Careful physical examination often reveals significant central and/or peripheral vascular disease. Findings such as a pulsatile intra-abdominal mass, diminished or absent peripheral pulse, claudication, rest pain, and tissue loss in lower extremities should be further evaluated by abdominal computed tomography scan or ultrasound, Doppler studies, and/or angiography. With the popularity of endovascular interventions, transplant surgeons also should be familiar with such technology and obtain detailed anatomic studies of patients with vascular stents.

**Immunologic Evaluation.** ABO blood typing and HLA typing (HLA-A, -B, and -DR) are required before a kidney transplant. The method of screening for preformed antibodies against HLA antigens (because of prior transplants, blood transfusions, or pregnancies) continues to evolve. The panel-reactive antibody (PRA) assay is a screening test that examines the ability of serum from a kidney transplant candidate to lyse lymphocytes from a panel of HLA-typed donors. A numeric value, expressed as a percentage, indicates the likelihood of a positive crossmatch with a donor. A higher PRA level identifies patients at high risk for a positive cross-match and therefore serves as a surrogate marker to measure the difficulty of finding a suitable donor and the subsequent risk of graft rejection.

An important development in anti-HLA antibody screening is Luminex technology, using HLA-coated fluorescent microbeads and flow cytometry, which is considered the “gold standard.” This technology pinpoints donor-specific antibodies (DSAs) in the serum of a kidney transplant candidate with a high PRA level. Since all organ donors must undergo HLA typing, a negative cross-match for recipients with a high PRA level can be ensured by avoiding the selection of donors carrying unacceptable antigens (i.e., a virtual cross-match). Kidney transplant candidate data (including ABO blood types, HLA types, and DSAs) are entered into a nationwide central database to facilitate deceased donor kidney allocation, as described earlier.

**Psychosocial Evaluation.** Psychiatric disorders have been recognized as important contributing factors to poor outcomes posttransplant. Patients with uncontrolled psychiatric disorders are at high risk for noncompliance with drug treatment, impaired cognitive function, and the development of substance abuse. A robust psychosocial evaluation is essential to ensure that transplant candidates understand the risks and benefits of the procedure and that they adhere to the lifetime immunosuppressive medication regimen.

**Recipient Operation**

Kidney allografts usually are transplanted heterotopically. The iliac fossa is recognized as the ideal position because of its proximity to the recipient’s bladder and iliac vessels.

Retroperitoneal allograft placement also allows easy access for percutaneous biopsies and interventions for ureteral complications. The right iliac fossa is the preferred site because of its easy access to the recipient’s iliac vessels. However, if a pancreas transplant is anticipated in the future or if now failed kidney grafts have been placed at the right iliac fossa, then the left iliac fossa is used for implantation. The current surgical technique for kidney transplants was developed and popularized in the 1950s and 1960s and has changed little since.

A large-bore three-lumen urinary catheter is inserted after the recipient is anesthetized, and it is occluded with a clamp beneath the surgical drapes. Recipients whose native kidneys produce urine will naturally fill up the urinary bladder; those individuals whose kidneys do not will require insufflation of saline prior to creation of the ureteral anastomosis.

Exposure of the operative field starts with a curvilinear skin incision, one to two finger widths above the midline pubic bone and the lateral edge of the rectus sheath. Superiorly, the extension of the incision depends on the recipient’s body habitus and the size of the donor kidney. The anterior rectus sheath is incised, medially to laterally, until the lateral edge of the rectus sheath is exposed. The posterior rectus sheath is missing below the arcuate line, thus providing direct access to the extraperitoneal space. The rectus muscle can be easily mobilized medially without being divided. The remainder of the fascial incision is along the lateral edge of the rectus sheath until the desired exposure is achieved (Fig. 11-6).

The retroperitoneal space of the iliac fossa is entered by mobilizing the peritoneum medially. The inferior epigastric vessels, the round ligament (in females), and the spermatic cord and its vasculature (in males) are encountered in this space; the former two structures are divided, while the latter is retracted with a vascular loop. A self-retained retractor is used to expose
the surgical field. The iliac vessels should be dissected with great care. To minimize the risk of lymphocele development postoperatively, dissection of the iliac artery should be limited; the intertwining lymphatics around the iliac vessels should be ligated. In general, the donor’s renal artery and vein are anastomosed to the recipient’s external iliac vessels in an end-to-side fashion (Fig. 11-7). In recipients with a severely calcified iliac artery, the internal iliac artery can be used as an alternative, and in select cases, an endarterectomy must be performed.

After restoring the circulation to the donor’s kidney, urinary continuity can be established via several approaches. The approach chosen depends on such factors as the length of the donor ureter and a recipient history of bladder surgery, native nephrectomy, or pelvic radiation. The two most common procedures to restore urinary continuity are the Leadbetter-Politano and a modification of the Lich (e.g., extravesical) ureteroneocystostomy.

During the former procedure, a large cystotomy is created in the dome of the bladder, and the donor ureter is brought through a lateral and somewhat inferior 1-cm submucosal tunnel into the bladder, the end of which is spatulated and then sewn in place without tension with interrupted absorbable sutures placed through the mucosa and submucosa on the inside of the bladder.

An extravesical ureteroneocystostomy is performed by careful dissection of a 1-cm portion of the muscular layers on the anterolateral portion of the bladder until a “bubble” of mucosa is exposed. The donor ureter is spatulated in a diamond-shaped fashion, the bladder mucosa is incised, absorbable interrupted sutures are placed in four quadrants, and a mucosa-to-mucosa anastomosis is created using running absorbable sutures with a temporary ureteral stent in place of the first three-quarters of the anastomosis. The muscular layers of the bladder are then carefully approximated over the anastomosis to prevent reflux.

The decision to place a ureteral stent depends on the surgeon, who must try to balance the risk of infectious complications with the possible technical complications of a ureteral anastomosis, but in general, this is not required except during the rarely performed donor ureter to recipient ureter anastomosis or in the case of a pediatric kidney transplant. Fixation of the donor’s kidneys is not necessary, except in the case of small kidneys (usually from a pediatric donor) or en bloc kidneys.

Figure 11-6. Incision and exposure for kidney transplant. A. Mark for the skin incision. B. Anterior rectus sheath incised obliquely. The abdominal muscle transected lateral to the rectus muscle. C. External iliac artery and vein dissected.

Figure 11-7. Vascular anastomoses of kidney transplant. A. Arterial anastomosis: donor renal artery with Carrel patch to recipient external iliac artery, end-to-side. B. Venous anastomosis: donor renal vein with caval extension conduit to recipient external iliac vein, end-to-side.
Figure 11-8. Arterial and venous reconstruction. A. Two renal arteries combined into a single Carrel patch (arrow). Right renal vein extension conduit constructed with stapled caval patch. IVC = inferior vena cava; R = right renal vein. B. Three renal arteries anastomosed to external iliac artery separately.

Grafts With Multiple Renal Arteries
In 10% to 30% of donor kidneys, multiple renal arteries are encountered. Unless kidney transplant candidates have hypercoagulopathy, grafts with multiple renal arteries fare as well as those with single vessels. Vascular reconstruction options include implanting the donor’s arteries separately, reconstructing the multiple arteries into a common channel, or combining multiple arteries into a common Carrel patch (Fig. 11-8).

En Bloc Grafts
Debate persists about whether to implant kidneys obtained from young donors (<5 years or whose body weight is under 20 kg) as a single en bloc unit into one recipient or separately into two recipients. The underlying issues are the shortage of donor organs, the complexity of the surgical procedure, the risks of graft thrombosis, ureteral complications, and long-term outcomes.

In en bloc kidney transplants, the donor aorta and vena cava are used as the vascular inflow and outflow conduits. Therefore, reconstruction of the en bloc graft pretransplant is key to a successful transplant. The donor’s suprarenal vena cava and aorta are oversewn. The lumbar branches of the cava and aorta are ligated. Dissection around the renal hilum should be avoided. The orientation of the cava and aorta should be clearly marked, in order to avoid torsion of the anastomosis. If the color of the two kidneys looks different after reperfusion, repositioning should be attempted to rule out vascular torsion; fixation of the en bloc kidneys to the retroperitoneum is often necessary. The donor’s ureters are implanted to the recipient’s bladder, either as two separate anastomoses or as a common patch (Fig. 11-9). Only a handful of centers have performed en bloc kidney transplants, but the long-term outcomes are encouraging.

Figure 11-9. En bloc kidney transplant (3-month-old donor kidneys). A. En bloc kidneys benched. Vascular integrity tested with methylene blue (blue hue look of the kidneys). B. En bloc kidneys transplanted into a 62-year-old woman. Donor aorta anastomosed to recipient’s external iliac artery; donor cava, to recipient’s external iliac vein.
Perioperative Care

Preoperatively, a thorough history and physical examination should be performed. Any changes in transplant candidates’ recent medical history should be investigated in great detail. In those recipients with a historically negative PRA level who have recently undergone blood transfusions, a prospective tissue cross-match is necessary to avoid graft rejection. Electrolyte panels should be checked. Emergency dialysis may be necessary for transplant candidates experiencing hyperkalemia or fluid overload.

For dialysis-dependent transplant candidates, the catheter sites should be examined preoperatively to rule out infections. Vascular access for hemodialysis is essential to avoid complications related to posttransplant acute tubular necrosis (ATN). Vascular evaluation is mandatory; any changes in results should be investigated by appropriate imaging studies.

As is routine for other major surgical procedures, transplant candidates should preoperatively undergo a chest X-ray, a 12-lead ECG, blood typing, cross-match tests, and prophylaxis against surgical site infection (by administration of a nonnephrotoxic antibiotic with activity against both common skin microflora and gram-negative pathogens); candidates should receive nothing to eat or drink.

Intraoperatively, transplant recipients should be kept well hydrated to avoid ATN and should receive heparin prior to vascular occlusion. Before reperfusion of the transplanted kidney, the desired central venous pressure should be maintained at around 10 mmHg, and the systolic blood pressure should be above 120 mmHg. In pediatric recipients of an adult graft, a superphysiologic condition may be necessary to avoid ATN or graft thrombosis. Mannitol often is administered before reperfusion as a radical scavenger and diuretic agent, and a diuretic such as furosemide is administered as well.

Postoperatively, the guiding principles for the care of kidney transplant recipients are the same as for other surgical patients. The crucial elements include hemodynamic stability and fluid and electrolyte balance. To achieve a euvolemic state, the recipient’s urine output is replaced with either an equal or a reduced volume of IV fluid on an hourly basis, depending on the medical status. In recipients undergoing brisk diuresis, aggressive replacement of electrolytes (including calcium, magnesium, and potassium) may be necessary. In recipients experiencing ATN, fluid overload, or hyperkalemia, however, fluid restriction, treatment for hyperkalemia, and even hemodialysis may be necessary.

Hypotension is an unusual event immediately posttransplant. The differential diagnoses include hypovolemia, vasodilation, and myocardial infarction with cardiac failure. Immediate action should be taken to avoid life-threatening complications. Posttransplant hypertension can be mediated by catecholamines, fluid overload, or immunosuppressive agents.

Postoperatively, urine output is used as a surrogate marker to monitor graft function. Among recipients whose native kidneys produce significant amounts of urine, normal or increased urine output can be misleading; for them, serum blood urea nitrogen and creatinine levels are more reliable indicators of kidney graft function.

Suddenly decreased or minimal urine output requires immediate attention. A change in volume status is the most common cause, but other culprits include blockage of the urinary catheter, urinary leak, vascular thrombosis, hypotension, drug-related nephrotoxicity, ATN, and rejection (all of which must be thoroughly investigated). Diagnostic studies such as Doppler ultrasound, nuclear renograms, or biopsies should be considered.

Postoperative bleeding is an uncommon event after a kidney transplant. Recipients on anticoagulation or antiplatelet treatments are at increased risk. Signs and symptoms (such as an expanding hematoma over the surgical site, increased pain over the graft, a falling hemoglobin level, hypotension, and tachycardia) should arouse suspicion of hemorrhage. Doppler ultrasound is useful to establish the underlying cause. Surgical exploration seldom is required because the accumulated hematoma tamponades the bleed. Indications for surgical exploration include ongoing transfusion requirement, hemodynamic instability, and graft dysfunction from hematoma compression. For recipients on anticoagulation or antiplatelet treatments, the threshold for surgical exploration is lower. Small unligated vessels at the donor’s renal hilum or recipient’s retroperitoneum are likely sources of bleeding.

One of the most devastating postoperative complications in kidney recipients is graft thrombosis. It is rare, occurring in fewer than 1% of recipients. The recipient risk factors include a history of recipient hypercoagulopathy and severe peripheral vascular disease; donor-related risk factors include the use of en bloc or pediatric donor kidneys, procurement damage, technical factors such as intimal dissection or torsion of vessels, and hyperacute rejection. Graft thrombosis usually occurs within the first several days posttransplant. Acute cessation of urine output in recipients with brittle posttransplant diuresis and the sudden onset of hematuria or graft pain should arouse suspicion of graft thrombosis. Doppler ultrasound may help confirm the diagnosis. In cases of graft thrombosis, an urgent thrombectomy is indicated; however, it rarely results in graft salvage.

Urologic complications are seen in up to 5% of recipients. The cause is often related to ureteral ischemia, damage during procurement of the donor’s distal ureter, or technical errors. Symptoms of urine leak include fever, pain, swelling at the graft site, increased creatinine level, decreased urine output, and cutaneous urinary drainage. Diagnosis can be confirmed by a combination of ultrasound, nuclear renography, drainage of perinephric fluid collection, and comparison of serum and fluid creatinine levels. Depending on the location and volume of the urine leak, satisfactory results can be achieved by surgical exploration and repair or by percutaneous placement of a nephrostomy and ureteral stenting.

Early urinary obstruction can be due to edema, blood clots, torsion of the ureter, or compression from a hematoma. Late urinary obstruction is often related to ischemia. The appearance of hydronephrosis on ultrasound is a good initial indicator. Treatment includes percutaneous placement of a nephrostomy and ureteral stenting. If transmural intervention fails, surgical intervention (such as ureteral reimplantation or a ureteropyelostomy) can be undertaken.

Results

A kidney transplant remains the most common solid organ transplant in the world today. With the introduction of induction immunosuppressive therapy and ever-improving, less toxic immunosuppressive medications, posttransplant outcomes have become better and better. And, as noted above, posttransplant outcomes have continued to improve: in 2014 allograft and patient survival rates were well over 90%, and in 2015, the 1-year graft survival rate with a living donor kidney was nearly
98%; with a deceased donor kidney, the rate was approximately 95%. The biggest improvements have been in the reduction of 1-year graft failure. With a deceased donor kidney, the 1-year graft failure rate dropped from approximately 20% in 1989 to less than 7% in 2009 to 4.8% in 2015; with a living donor kidney, the rate dropped from 8.5% in 1989 to less than 3% in 2015. Furthermore, steroid-free protocols and calcineurin-free protocols have been validated and implemented in the last several decades, further reducing medication-related side effects and vastly improving the quality of life for tens of thousands of recipients.

Currently, the most common cause of graft loss is recipient death (usually from cardiovascular causes) with a functioning graft. The second most common cause is chronic allograft nephropathy; characterized by a slow, unrelenting deterioration of graft function, it likely has multiple causes (both immunologic and nonimmunologic). The graft failure rate due to complications related to surgical technique has remained at about 1% to 2%.

**PANCREAS TRANSPLANTATION**

A successful pancreas transplant currently is the only definitive long-term treatment for patients with insulin-dependent diabetes mellitus (IDDM) that (a) restores normal glucose hemostasis without exposing patients to the risk of severe hypoglycemia and (b) prevents, halts, or, in some cases, reverses the development or progression of secondary complications of diabetes.

Given its vast medical, social, and financial implications, diabetes mellitus is a huge burden to patients and to society as a whole. An estimated 10% to 15% of the U.S. population is affected by it; of all diabetic patients, 10% have early-onset disease. In the United States, diabetes mellitus is the most common cause of end-stage kidney disease, blindness, impotence, major limb amputations, and coronary or peripheral vascular bypass procedures. It is one of the most common causes of death, along with myocardial infarction and stroke. Diabetes significantly decreases not only the quality of life but also life expectancy.

Despite improvements in exogenous insulin administration (including the use of devices such as insulin pumps), wide fluctuations in glucose levels and the risk of hypoglycemic episodes are common. The Diabetes Control and Complications Trial (DCCT) demonstrated in the late 1990s that intensive insulin therapy may slow the rate of secondary complications of diabetes—yet at the expense of (life-threatening) iatrogenic hypoglycemia. The annual mortality rate of patients with insulin-induced inadvertent hypoglycemia is estimated to be as high as 2% to 3%.

Since the first pancreas transplant in December 1966, performed by William Kelly and Richard Lillehei at the University of Minnesota, more than 25,000 pancreas transplants in the United States and more than 10,000 pancreas transplants from all over the world have been reported to the International Pancreas Transplant Registry (IPTR).

Pancreas transplants are performed in three recipient categories:

- **Simultaneous pancreas and kidney (SPK) transplant in diabetic and uremic patients.** Almost 80% of pancreas transplants are performed in this category. The recipient is already obligated to lifelong immunosuppressive therapy, due to the need for a kidney transplant, so only the surgical risk of a pancreas transplant is added. A successful SPK transplant renders the recipient dialysis-free and insulin-independent.
- **Pancreas after kidney (PAK) transplant in diabetic and posturemic patients.** Approximately 15% of all pancreas transplants fall into this category. These patients previously underwent a kidney transplant with either a living or deceased donor, but are candidates for a subsequent pancreas transplant because of poor glucose control or because of progression of secondary diabetic complications (which may include the development of diabetic nephropathy in the transplanted kidney).
- **Pancreas transplant alone (PTA) in nonuremic patients with brittle diabetes mellitus.** Only about 5% of all pancreas transplants are in this category. These patients have not yet developed advanced diabetic nephropathy, but their glucose levels are extremely labile despite best efforts of control. Because of the lifelong need for immunosuppressive therapy, the surgical risk has to be balanced with the medical risks of brittle diabetes (e.g., frequent episodes of hypoglycemia and hypoglycemic unawareness).

In SPK recipients, a plethora of literature exists that demonstrates significant improvements in secondary diabetic complications (across all organ systems) posttransplant. Improvements have been reported in diabetic nephropathy, neuropathy (autonomic and peripheral), micro- and macrovascular disease, retinopathy, gastroparesis, and other secondary complications. Currently, more than 1000 pancreas transplants are performed annually in the United States, with the goal of conferring the following benefits: excellent glucose control (similar to that of a functioning native pancreas), prevention or improvement of secondary diabetic complications, and increased quality of life and life expectancy. In addition, pancreas transplants can be successfully performed in patients who have undergone a total pancreatectomy for benign disease (such as chronic pancreatitis) to treat both endocrine and exocrine deficiency after surgery.

**Donor Operation**

The general criteria for selecting deceased donors for pancreas procurement are similar to those for other solid organs; a history of type 1 diabetes mellitus obviously is a contraindication. Relative contraindications include previous pancreatic procedure(s), as well as pancreatic disorders, such as chronic pancreatitis and intraductal papillary mucinous neoplasm. Hyperglycemia in itself is not a contraindication to pancreas procurement because its cause in brain-dead donors usually is severe insulin resistance, which is rarely observed in recipients.

In light of better anatomic understanding and improved surgical skills, all three abdominal organs that share a common blood supply (pancreas, liver, and intestine) can be procured at the same time and transplanted into three different recipients (Fig. 11-10). During pancreas procurement, a “no-touch” technique of the gland is preferred; dissection of the pancreas is carried out in a way that avoids direct manipulation of the organ such that simultaneous procurement of the spleen, duodenum, and surrounding connective tissues occurs.

In contrast to the liver and kidneys, the pancreas should not be extensively flushed at the end of the procurement. To minimize the amount of preservation fluid that reaches the pancreas, the splenic artery and SMA can be temporarily clamped...
at their origin from the aorta. Usually, the celiac axis with an aortic Carrel patch is retained with the liver. The splenic artery is divided close to its origin and is retained with the pancreas. The SMA is also procured with an aortic Carrel patch and is retained with the pancreas.

In case of a replaced or aberrant right hepatic artery, this first branch off of the SMA is carefully dissected out from the posterior surface of the pancreas. A replaced or aberrant right hepatic artery does not transverse the pancreas and is not a contraindication to combined pancreas and liver procurement. But with this anatomic variant, an aortic Carrel patch with the proximal SMA and replaced or aberrant right hepatic artery remains with the liver; the distal SMA with the inferior pancreaticoduodenal artery remains with the pancreas.

In the relatively rare event that the liver is not procured, then neither the splenic nor the gastroduodenal arteries need to be divided at their respective takeoff; the donor’s celiac axis and the SMA are included on a common Carrel patch. This technique allows a single arterial anastomosis to be performed in the recipient without reconstruction. At the end of the procurement, the pancreas is attached to the spleen, duodenum, and proximal jejunum, which is stapled at both ends.

**Back Table Preparation of the Pancreas Graft**

Back table preparation of the pancreas graft consists of four steps: (a) removal of the spleen; (b) shortening, restapling, and/or suture reinforcement of the mesenteric root; (c) trimming of any excess distal and proximal duodenum, along with reinforcement of the proximal staple line; and (d) arterial reconstruction.

Back table preparation is carried out in a basin filled with chilled preservation solution. The most common technique to create a single arterial inflow to the pancreas graft is the “Y-graft” reconstruction, using a resected segment of the donor iliac artery bifurcation. In this technique, the donor external iliac artery is anastomosed end-to-end to the donor SMA, and the donor internal iliac artery is anastomosed end-to-end to the splenic artery (Fig. 11-11). This procedure allows the donor common iliac artery to be anastomosed as a single vessel to the recipient’s common iliac artery. For venous outflow, the portal vein is kept relatively short, in order to avoid the risk of venous thrombosis by kinking or impingement.

**Recipient Operation**

Over the years, different surgical techniques have been described for (a) the management of exocrine pancreatic secretions and (b) the type of venous drainage. For the secretions, the two most common techniques are drainage of the duodenal segment to the bladder (bladder drainage) or to the small bowel (enteric drainage) (Figs. 11-12 and 11-13). For venous drainage, systemic venous drainage is preferred over portal venous drainage.

The pancreas graft is usually placed intra-abdominally and preferably on the right side because the iliac vessels are...
in a more shallow position on the right than on the left side; moreover, the vessels are already appropriately aligned for the vascular anastomoses (i.e., a lateral position for the common iliac vein, a medial position for the common iliac artery). Venous and arterial anastomoses are performed end-to-side. After restoration of blood flow to the graft, hemostasis must be meticulously maintained. Because the donor portal vein purposely is kept short, ligation and transection of all of the recipient’s internal iliac vein branches are frequently performed in order to prevent tension on the venous anastomosis. The pancreas usually is placed with the pancreatic head and duodenum pointing caudally.

Bladder drainage is performed using either a hand-sewn or a stapled anastomosis in which the antimesenteric side of the donor duodenum is sewn to the superior portion of the dome of the bladder. The stapled technique requires that a circular cutting stapler be inserted through the open distal end of the donor duodenum, which is subsequently closed. Bladder drainage has two main advantages. First, rejection of the exocrine pancreas precedes rejection of the endocrine pancreas by 5 to 7 days. Amylase levels are measured routinely in the recipient’s urine. With bladder drainage, antirejection treatment can successfully be implemented when the recipient is still normoglycemic and only hypoamylasuric. In the absence of hyperglycemia, more than 90% of pancreas rejection episodes are reversible. Second, bladder drainage avoids the bacterial contamination that occurs with enteric drainage. If an anastomotic leak occurs, it is easier to treat because the infection usually remains localized to the right lower quadrant.

Enteric drainage is more physiologic and has advantages as well. The antimesenteric side of the donor’s duodenum is anastomosed to the antimesenteric portion of the recipient’s jejunum in a side-to-side fashion. The enteric anastomosis can also involve a defunctionalized Roux-en-Y loop, which minimizes the potential complications if an enteric leak occurs. Currently, in the United States, more than 80% of all pancreas transplants are performed with enteric drainage for the exocrine pancreatic secretions, and more than 90% employ systemic venous drainage.
Complications

The technical complication rate for pancreas transplants is higher than for any other solid organ transplant. Four factors contribute to the high surgical complication rate: (a) the nature of the organ itself with inherent organ-specific surgical complications (e.g., peritonitis, abscesses, necrosis, fistulas, and pseudocysts) and its low blood flow (which significantly increases the risk of thrombosis, as compared with a kidney or liver transplant); (b) the risk of a leak or infection after connecting two hollow viscera (the duodenum and either the bladder or small intestine); (c) the increased incidence of rejection episodes because the pancreas is one of the most immunogenic solid organs; and (d) the underlying disease of diabetes mellitus, predisposing patients not only to infections but also to cardiovascular and other complications.

The most common surgical complications are thrombosis (an incidence of 5%–15%), intra-abdominal abscesses (5%–10%), and bleeding (6%–8%). Other pancreas-specific complications include graft pancreatitis (frequently due to procurement or reperfusion injury), pancreatic fistulas, and pancreatic pseudocysts. Anastomotic leaks do not always require a graft pancreatectomy, but arterial pseudoaneurysms, arteriovenous fistulas, and wound dehiscence may. Bleeding frequently requires relaparotomy.

Thrombosis usually occurs within the first week posttransplant. It manifests as a sudden increase in insulin requirements or as a sharp drop in urinary amylase levels. Venous thrombosis, which is more common than arterial thrombosis, is associated with distinct clinical symptoms, including a swollen and tender graft, hematuria, lower extremity edema, and deep vein thrombosis, the latter two occurring ipsilaterally. Arterial thrombosis is less symptomatic and may not initially cause pain; its diagnosis is usually confirmed by Doppler ultrasonography. Surgical exploration in recipients with thrombosis usually requires a graft pancreatectomy.

With the advent of advanced interventional radiologic procedures to drain intra-abdominal abscesses, the reoperation rate has markedly decreased. Pancreas transplant recipients are usually kept on broad-spectrum antimicrobial agents for the first 7 days posttransplant.

The most common nonsurgical complication posttransplant is rejection. The incidence of rejection is about 30% within the first year. The diagnosis is usually based on an increase in serum amylase and lipase levels and, in bladder-drained recipients, a decrease in urinary amylase levels. A sustained drop in urinary amylase levels greater than 25% from baseline should prompt a pancreas graft biopsy to rule out rejection. In enteric-drained recipients, one must rely on serum amylase and lipase levels only. Other signs and symptoms of rejection include tenderness over the graft, unexplained fever, and hyperglycemia, which usually is a late finding; fewer than 5% of all rejection episodes can be reversed in its presence. The diagnosis of rejection should be confirmed by a percutaneous pancreas graft biopsy.

Other nonsurgical complications include infections with CMV, HCV, or extra-abdominal bacteria or fungi; malignancies, such as PTLD; and, rarely, graft-versus-host disease. For such complications, the diagnosis and treatment are similar to what is recommended after other solid organ transplants.

Bladder-drained pancreas recipients may experience an array of unique urologic complications. Usually the result of the irritating nature of pancreatic enzymes on the urothelium in the bladder and urethra, these urologic complications can lead to cystitis, hematuria, and dysuria. With the loss of bicarbonate from pancreatic secretions, dehydration and metabolic acidosis are not uncommon. Many of these complications are chronic, such that approximately 20% to 30% of all bladder-drained recipients require conversion to enteric drainage within the first 5 years posttransplant.

Living Donor Pancreas Transplants

Pancreas transplants using living donors also can be performed safely and successfully in select donors and recipients. Since 1979, about 150 such transplants have been performed worldwide, with 1-year graft survival rates in excess of 85% over the last decade. A meticulous donor evaluation using standard criteria remains key to a low donor metabolic and surgical complication rate. The concept of procuring the distal pancreas from a living donor is based on the observation that patients with benign or malignant pancreatic disorders can undergo a distal hemipancreatectomy without any serious change in endocrine function.

Living donor pancreas transplants are ideal for patients with an identical twin, but other relatives can be suitable donors as well. In particular, patients with high PRA levels should be considered for a living donor transplant.

Living donor pancreas transplants decrease the number of deaths of diabetic patients on the waiting list, help overcome the organ shortage, reduce mortality and morbidity, and improve the quality of life for patients with debilitating side effects of diabetes. The use of living donors also reduces the risk of graft rejection, as compared with the use of deceased donors. Yet living donor pancreas transplants remain relatively rare, performed under very selective circumstances. In terms of surgical technique, the donor splenic artery and vein are anastomosed to the recipient’s external iliac artery and vein in an end-to-side fashion, and exocrine drainage can occur via an anastomosis.

new donor pancreases is an appealing option for patients with type 1 diabetes. An islet transplant involves the procurement of a solitary pancreas for one islet recipient. If the primary goal of current islet transplant trials is not insulin independence but rather a reduction in the incidence and severity of hypoglycemic events, a reduction in exogenous insulin requirements, and an amelioration of hemoglobin A1c levels. Islet transplants rarely maintain long-term insulin independence. A recent study showed a higher rate of insulin independence in PTA recipients than in recipients of an islet transplant alone, despite the use of up to three donor pancreases in each of the islet recipients. Until islet transplant results significantly improve and include long-term insulin independence, a pancreas transplant remains the treatment of choice for β-cell replacement therapy in patients with IDDM.

**Islet versus Pancreas Transplants**

Pancreas transplants are frequently compared with islet transplants (vide infra), which are less invasive and, therefore, more appealing. It is important to emphasize that these two types of transplants are not mutually exclusive but rather complementary. The results of islet transplants have improved over the past decade, but overall islet graft function, specifically long-term function, is still significantly trailing overall pancreas graft function.

Islet transplants involve pancreas procurement (as described earlier) and then separation of islets from the exocrine pancreatic tissues using proteolytic enzymes (as described later). The human pancreas contains about one million islets, of which half are lost during the isolation process. About 10,000 islets per kilogram of body weight are needed to achieve insulin independence when transplanted into the liver. Frequently, one donor pancreas does not suffice; in fact, up to four donor pancreases have been used for one islet recipient.

Because of the relatively disappointing long-term outcomes, insurance providers in the United States do not provide reimbursement for islet transplants. Transplant centers with both pancreas and islet transplant programs follow an algorithm that favors islet transplants in patients with a high surgical risk and pancreas transplants in patients with a low surgical risk. Although solitary donor pancreases are not in short supply, only one donor pancreas is required for a successful pancreas transplant; in contrast, two to four donor pancreases are commonly used for one islet recipient with less favorable long-term outcomes.

Of note, the primary goal of current islet transplant trials is not insulin independence but rather a reduction in the incidence and severity of hypoglycemic events, a reduction in exogenous insulin requirements, and an amelioration of hemoglobin A1c levels. Islet transplants rarely maintain long-term insulin independence. A recent study showed a higher rate of insulin independence in PTA recipients than in recipients of an islet transplant alone, despite the use of up to three donor pancreases in each of the islet recipients. Until islet transplant results significantly improve and include long-term insulin independence, a pancreas transplant remains the treatment of choice for β-cell replacement therapy in patients with IDDM.

**Islet Transplantation**

Transplanting islets of Langerhans isolated from deceased donor pancreases is an appealing option for patients with type 1 diabetes. An islet transplant involves the procurement of a simultaneous kidney transplant and excellent results. Transplant outcomes in SPK improved significantly because of a decrease in the rates of technical and immunologic graft loss. In 2010 to 2014 vs. 2005 to 2009, U.S. SPK transplant patient survival at 1 year posttransplant increased from 95.7% to 97.4%, pancreas graft function increased from 88.3% to 91.3%, and kidney function increased from 93.6% to 95.5%. A significant improvement was also noted in PAK transplants. One-year patient survival increased from 96.4% to 97.9%, and pancreas graft function increased from 81.0% to 86.0%. PTA 1-year patient survival remained constant at 97%, and pancreas 1-year graft survival improved from 81.0% to 85.7%. IPTR data show significant improvements in patient survival and pancreas graft function rates since the inception of UNOS, over a course of 24 years. Clearly, pancreas transplants now offer excellent outcomes for patients with IDDM.

**Results**

As of December 2010, more than 35,000 pancreas transplants had been reported to the IPTR: more than 25,000 transplants in the United States and more than 10,000 in other countries. According to IPTR data, recipient age at the time of the transplant has increased significantly, and so has the number of transplants for patients with type 2 diabetes. The trend over time has been toward stricter donor criteria, with a concentration on younger donors, preferably trauma victims, and on short pancreas graft preservation time.

Drainage techniques have changed over time, too: enteric drainage of exocrine pancreatic secretions is now predominant, in combination with systemic drainage of the venous effluent of the pancreas graft. Immunosuppressive protocols have developed toward antibody induction therapy, followed by administration of tacrolimus and MMF for maintenance. Steroid avoidance has increased over time in all three recipient categories.

Between 2005 and 2009 and 2010 and 2014, the number of U.S. pancreas transplants declined by over 20%, while the overall number of pancreas transplants performed outside the United States has increased. The decline in U.S. numbers is predominantly due to the decline in primary and secondary pancreas after kidney transplants (PAK). During the time period studied, the number of PAK transplants dropped by 50%. In contrast, the number of simultaneous pancreas/kidney transplants (SPK) declined by only 10%, and the number of pancreas transplants alone (PTA) by 20%. Over 90% of pancreas transplants worldwide were performed, with a simultaneous
donor pancreas and its transportation to a specialized islet isolation facility, where the pancreas is enzymatically digested; then, the islets are purified from the rest of the digested pancreas using density gradients. The purified islets are then cultured and evaluated for their identity, viability, and potency, before being infused into the portal vein of a diabetic recipient. When the procedure is successful, these islet cells engraft into the recipient and secrete insulin, providing excellent moment-to-moment control of blood glucose, as is seen with a whole-pancreas transplant.

A successful islet transplant offers advantages over exogenous insulin injections—advantages that are similar to those of a whole-pancreas transplant. These advantages include restoring β-cell secretory capacity, improving glucose counterregulation, restoring hypoglycemia awareness, providing perfect or near-perfect glucose homeostasis, and, potentially, preventing secondary diabetic complications.

Unlike a whole-pancreas transplant, an islet transplant does not involve a major surgical procedure with its associated mortality and morbidity. Instead, it can generally be performed as an outpatient procedure using percutaneous catheter-based therapy to cannulate a branch of the portal vein, with minimal recovery time for the recipient. Potential complications associated with islet injection include portal hypertension, portal vein thrombosis, hepatic abscesses, and bacteremia. Theoretically, islet transplants could have wider application (as compared with current practice and with whole-pancreas transplants), given the significantly lower surgical risk, the relatively small tissue volume transplanted, and the potential for islet immunomodulation or immunoisolation, which could minimize or eliminate the need for immunosuppression.

The first reported attempt at an islet transplant was in 1893 by Watson-Williams and Harsant: they transplanted a sheep’s minced pancreas into the subcutaneous tissue of a young boy with ketoacidosis. The discovery of insulin may have reduced interest in islet transplants as a treatment for diabetes, at least until the realization that insulin could not provide perfect glycemic control and that, therefore, patients ultimately suffered devastating secondary complications. Several milestones ensued: the first whole-pancreas transplants, early success with rodent islet transplants, and then, in the 1970s, human islet autotransplants after pancreatectomy, in order to address the intractable pain associated with chronic pancreatitis, by Sutherland, Najarian, and colleagues in Minnesota.108

Until recently, attempts to extend these trailblazing findings of clinical islet autotransplants to clinical islet allografts in patients with type 1 diabetes met with generally very poor success. For example, in 1995, a report of the International Islet Transplant Registry indicated that of 270 recipients, only 5% were insulin-independent at 1 year posttransplant.

In 2000, Shapiro and colleagues reported the results of the Edmonton protocol, which enabled consistent diabetes reversal and short-term (<1 year) insulin independence. The Edmonton protocol prescribed transplanting a large number of freshly isolated islets (>10,000 islet equivalents per kilogram body weight, typically requiring the use of two to four pancreases) with a specialized “islet-sparing,” steroid-free immunosuppressive protocol consisting of low-dose tacrolimus, sirolimus, and IL-2 receptor antibody induction. Those results were replicated at other experienced transplant centers, but the rates of long-term (>5 year) insulin independence remained poor, well below those of whole-pancreas transplants.114 Still, despite the low rates of long-term insulin independence, most islet recipients were C-peptide positive and retained hypoglycemia awareness, indicating residual islet function and benefit. In fact, at 9 years posttransplant, 15% remained insulin-independent, and 73% had hypoglycemia awareness and corrected hemoglobin A1c levels.

In the mid-2000s, new trials began with the goal of establishing protocols that enable insulin independence, using islets from a single donor pancreas; the results were good, especially with strict donor and recipient selection. In the most experienced centers, long-term rates of diabetes reversal are now about 50% at 5 years posttransplant. The reasons include refinements in pancreas preservation, islet isolation, and culture protocols, as well as the use of newer induction immunosuppressive agent combinations, such as a T-cell-depleting antibody (anti-CD3 antibody, alemtuzumab, or antithymocyte globulin) and a tumor necrosis factor-alpha (TNF-α) inhibitor (etanercept or infliximab). Presumably, viable β-cell mass is now preserved, both pre- and posttransplant. Thus, islet transplant results are approaching those of whole-pancreas transplants; however, because islets from more than one pancreas are typically needed, those results cannot be directly compared with the results of whole-pancreas transplants.

In the United States, an islet transplant is still officially deemed an experimental procedure. In contrast, since 2001, it has been considered a standard of care and is fully reimbursed in Canada and, more recently, in the United Kingdom, Sweden, Switzerland, France, and Italy as well. The full potential of islet transplants remains to be realized, but the future is exciting. As the latest improvements in pancreas preservation, islet isolation and purification, islet culture, and islet immunosolation are implemented clinically, the hope is that sustained insulin independence may become consistently possible with a single pancreas donor and without the need for systemic immunosuppression.

**LIVER TRANSPLANTATION**

The first attempts at liver transplants in the late 1960s through the 1980s were largely experimental endeavors, with a 1-year survival rate of only 30%. But breakthroughs in immunosuppression, surgical technique, organ preservation, anesthesia, and critical care have improved that rate to approximately 85% today. Liver transplants remain daunting, especially in the face of an organ shortage that results in sicker potential candidates. Unfortunately, the perioperative mortality rate and the 1-year mortality rate are among the highest of any surgical operation currently performed.

**History**

The first experimental liver transplants in dogs are often attributed to C. Stuart Welch in 1955 and then Jack Cannon in 1956. However, current scholarship reveals that Vittorio Staudacher first described the technique in 1952. A series of canine experiments followed, which refined the surgical technique to ensure perioperative survival.

The next obstacle—immunologic rejection—was addressed by drug immunosuppression with AZA and prednisone. The first human liver transplant trials started in 1963 with Thomas Starzl, but a series of deaths led to a voluntary moratorium for 3.5 years. With the resumption of clinical transplants in 1967, Starzl performed the first successful liver transplant.
Still, for the next decade, survival rates were dismal: only 20% of the 170 liver transplant recipients in Starzl’s program at the University of Colorado survived more than 5 years.\(^{124}\)

Several innovations dramatically improved outcomes. The advent of better immunosuppressive drugs was instrumental. In 1978, cyclosporine was introduced clinically in England. It was soon combined with prednisone to great effect. The arrival of tacrolimus in the 1990s further improved graft survival.

Technical advances were also significant. Donor procurement techniques and cold organ preservation protocols were standardized, and the recipient operation was also refined. Choledochocystic and choledochojunostomy to a Roux-en-Y limb became standard and significantly decreased the frequency of biliary complications. Innovations, including living donor liver transplants and deceased donor split-liver transplants, enabled more pediatric recipients to be transplanted. Improvements in portosystemic shunting and perioperative critical care also were contributory.

**Indications**

In general, any form of irreversible liver disease is an indication for a liver transplant. Chronic alcoholic disease and HCV are the most common indications in the United States. An extensive list of acute and chronic diseases of the liver that are treatable by a liver transplant is provided in Table 11-6.

Offering transplants to alcoholic patients has always drawn some opposition because of the perception of it being a self-inflicted illness, as well as concerns about recidivism and the recipient’s possible inability to maintain postoperative immunosuppression and care. Yet studies have shown that such patients have excellent outcomes and that liver transplants for them are cost-effective.\(^{125-127}\) Because patients who drink 4 to 8 ounces of liquor daily for 10 to 15 years have an increased risk of developing cirrhosis, the general requirement for acceptance as a transplant candidate is 6 months of abstinence. Furthermore, most transplant centers recommend rehabilitation and Alcoholics Anonymous programs.

Transplants for HCV have yielded worse outcomes than transplants for other diseases.\(^{128}\) The reason is the universal recurrence of the virus posttransplant. Viral levels reach pretransplant levels as early as 72 hours posttransplant.\(^{129}\) The course of the viral infection is often accelerated posttransplant: 10% to 20% of recipients develop cirrhosis after just 5 years.\(^{130}\) Studies have suggested that use of older donors may increase the chance of aggressive recurrence.\(^{131}\) The best method to prevent recurrence would be to eradicate the infection pretransplant, but doing so is not always possible because patients with decompensated cirrhosis often cannot tolerate treatment. Once recurrence occurs, treatment methods are limited. One study found that pegylated interferon and ribavirin therapy achieved a sustained viral response in 44% of patients.\(^{132}\)

A substantial number of patients undergo liver transplants for cholestatic disorders. Primary biliary cirrhosis, an autoimmune disease, is characterized by damage to the intralobular bile ducts that progresses to liver cirrhosis. Trends toward earlier treatment may explain the slight decrease in liver transplants for this disorder.\(^{133}\) Posttransplant outcomes in patients with this disorder have been excellent, with many centers achieving 1-year survival rates of 90% to 95%. Recurrence is relatively uncommon: a large series reported a 30% recurrence rate at 10 years posttransplant.\(^{134}\)

The second most common cholestatic disorder among liver transplant candidates is primary sclerosing cholangitis. It is characterized by inflammation and fibrosis of large intra- and extrahepatic biliary ducts; 70% of such patients also have inflammatory bowel disease. Recurrent cholangitis is common and increases mortality rates beyond what would be expected on the basis of laboratory values. On behalf of such patients, appeals can often be made for priority in allocation to the UNOS regional review boards. Posttransplant outcomes for such patients have been excellent. Primary sclerosing cholangitis is a significant risk factor for cholangiocarcinoma, so annual screenings (including imaging and measurement of serum CA 19-9 levels) should be carried out. Recurrence is fairly uncommon: studies have reported a recurrence rate of up to 20% at 10 years posttransplant.\(^{135}\)

Progressive metabolic disorders also are treatable with liver transplants. Hemochromatosis, an inherited disorder, results in excessive intestinal iron absorption. Iron deposition can cause cirrhosis and severe cardiomyopathy. Careful cardiac evaluation is necessary pretransplant.

Another metabolic disorder, \(\alpha_1\)-antitrypsin deficiency, is characterized by insufficient levels of a protease inhibitor, resulting in early-onset emphysema and cirrhosis. Careful pulmonary evaluation is necessary pretransplant.

Wilson’s disease, an autosomal recessive disorder characterized by impaired cellular copper transport, leads to copper accumulation in the liver, brain, and cornea. Patients can develop significant neurologic complications and cirrhosis. Several reports suggest improvement of neurologic deficiencies posttransplant.\(^{136,137}\)

Transplants can also be performed in patients with hepatic malignancies, but only in accordance with strict criteria.

---

**Table 11-6**

<table>
<thead>
<tr>
<th>Diseases amenable to treatment by a liver transplant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autoimmune liver disease</td>
</tr>
<tr>
<td>Autoimmune hepatitis</td>
</tr>
<tr>
<td>Primary biliary cirrhosis</td>
</tr>
<tr>
<td>Primary sclerosing cholangitis</td>
</tr>
<tr>
<td>Congenital</td>
</tr>
<tr>
<td>Biliary atresia</td>
</tr>
<tr>
<td>Viral hepatitis</td>
</tr>
<tr>
<td>Hepatitis B</td>
</tr>
<tr>
<td>Hepatitis C</td>
</tr>
<tr>
<td>Alcoholic liver disease</td>
</tr>
<tr>
<td>Metabolic diseases</td>
</tr>
<tr>
<td>(\alpha_1)-Antitrypsin deficiency</td>
</tr>
<tr>
<td>Cystic fibrosis</td>
</tr>
<tr>
<td>Hemochromatosis</td>
</tr>
<tr>
<td>Tyrosinemia</td>
</tr>
<tr>
<td>Wilson’s disease</td>
</tr>
<tr>
<td>Hepatic malignancy</td>
</tr>
<tr>
<td>Hepatocellular carcinoma</td>
</tr>
<tr>
<td>Neuroendocrine tumor metastatic to liver</td>
</tr>
<tr>
<td>Fulminant hepatic failure</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Alagille syndrome</td>
</tr>
<tr>
<td>Cryptogenic cirrhosis</td>
</tr>
<tr>
<td>Budd-Chiari syndrome</td>
</tr>
<tr>
<td>Polycystic liver disease</td>
</tr>
<tr>
<td>Amyloidosis</td>
</tr>
</tbody>
</table>
Hepatocellular carcinoma (HCC), a complication of cirrhosis, is the most common type of hepatic malignancy. Resection is the first line of treatment if possible, but often, cirrhosis is too advanced. If the tumor meets the Milan criteria, a liver transplant can be performed. These criteria were established by a landmark paper in 1996 showing that patients with a single tumor under 5 cm in diameter, or with three tumors under 3 cm in diameter, in the absence of vascular invasion, had a 4-year survival rate of 85%. Patients with such tumors receive exception points, based on their UNOS region, allowing for a timely transplant before their tumors spread.

Transplants for cholangiocarcinoma remains controversial but may be performed if the center has an experimental protocol in place that entails strict recipient selection. The use of a multimodality oncologic approach including neoadjuvant chemo radiotherapy with subsequent OLT achieves excellent results for patients with localized, regional lymph node-negative pCCA. Patient survival after OLT is comparable to the results of OLT for other causes.

Acute fulminant hepatic failure also is an indication for a liver transplant; in fact, such patients are the highest priority for the next available liver in their UNOS region. This devastating illness is defined by acute and severe liver injury with impaired synthetic function and encephalopathy in a person who had normal liver function. It is often caused by acetaminophen overdose; acute fulminant viral hepatitis A, B, and E; other viral infections; drug toxicity; ingestion of Amanita mushrooms; acute fatty liver of pregnancy; or Wilson’s disease. A significant number of patients will recover with supportive care. The difficulty lies in predicting who will not recover and therefore would benefit from a liver transplant. The King’s College criteria were developed for this purpose: patients with acetaminophen-induced disease, a pH <7.3 or grade III/IV encephalopathy, a prothrombin time >100 seconds, and serum creatinine >3.4 mg/dL meet those criteria. Management of acute liver failure is very intensive. Such patients suffer from severe coagulopathy, hypoglycemia, lactic acidosis, and renal dysfunction. They are susceptible to infections, which are frequently overwhelming. Cerebral edema, a serious complication of acute liver failure, is a leading cause of death from brain herniation. Intracranial pressure monitoring and serial imaging are often necessary; if a patient develops irreversible brain damage, a transplant is not performed.

**Recipient Selection**

The diagnosis of cirrhosis itself is not an indication for a transplant. Patients may have compensated cirrhosis for years such that the traditional indication for a transplant is decompensated cirrhosis, manifested by hepatic encephalopathy, ascites, spontaneous bacterial peritonitis, portal hypertensive bleeding, and hepatorenal syndrome (each described below).

Hepatic encephalopathy is an altered neuropsychiatric state caused by metabolic abnormalities resulting from liver failure. The early stages result in sleep disturbances and depression. As the liver disease progresses, patients can become somnolent and confused and, in the end stages, comatose. Ammonia is produced by enterocytes from glutamine and from colonic bacterial catabolism, and the use of serum ammonia levels as a marker of encephalopathy is controversial because a variety of factors can influence levels. Hyperammonemia suggests worsening liver function and bypass of portal blood flow around the liver. GI bleeding and infection can exacerbate hepatic encephalopathy.

Ascites (the accumulation of fluid in the abdominal cavity) that is caused by cirrhosis is a transudate with a high serum-ascites gradient (>1.1 g/dL). Associated with portal hypertension, it is treated initially with sodium restriction and diuretics. Refractory ascites necessitates large-volume paracentesis and eventually a transjugular intrahepatic portosystemic shunt (TIPS). Contraindications to TIPS placement include significant hepatic encephalopathy, advanced liver disease, congestive heart failure, renal insufficiency, and severe pulmonary hypertension.

Spontaneous bacterial peritonitis, an infection of the ascitic fluid without an evident intra-abdominal source, is characterized by fever, abdominal pain, and an ascitic fluid polymorphonuclear count ≥250 cell/mm³ on paracentesis. The first line of empiric treatment is with a third-generation cephalosporin because the majority of cases are caused by aerobic gram-negative microbes such as *E. coli*, although Gram stain and culture results should be used to guide therapy.

Portal hypertensive bleeding can be a devastating event for patients with cirrhosis. Each bleeding event carries a 30% mortality rate and accounts for a third of all deaths related to cirrhosis. Only 50% of bleeding events cease spontaneously, so treatment must be expedient. The initial medical treatment is with vasopressin and octreotide. The initial intervention is endoscopy with sclerotherapy and band ligation of bleeding varices. If these initial attempts fail, more aggressive treatment is necessary with a balloon tamponade (using a Sengstaken-Blakemore tube) and with emergent TIPS placement. The last line of treatment is emergency surgery to place a portosystemic shunt, transect the esophagus, or devascularize the gastroesophageal junction (Sugiura procedure). Preventing variceal bleeding is essential and can be achieved, with some success, using β-blockers.

Hepatorenal syndrome is a form of acute renal failure that develops as liver disease worsens. The etiology is unclear, but splanchnic vasodilation from portal hypertension and increased production of circulating vasodilators result in a decline in renal perfusion. Characterized by oliguria (<500 mL of urine/day) and low urine sodium levels (<10 mEq/L), hepatorenal syndrome is often reversed by a liver transplant, even after dialysis dependence. Pretransplant, other causes of renal failure need to be excluded, including ATN, drug nephrotoxicity, and chronic renal disease. The initial medical therapy includes octreotide, midodrine, and vasopressin analogs, but the syndrome often progresses to dialysis dependence.

The Model for End-Stage Liver Disease (MELD) was originally developed to assess risk for TIPS placement. Later analysis revealed it to be an excellent model to predict survival among patients with cirrhosis, especially those on the waiting list for a liver transplant. In 2002, liver graft allocation was restructured to be based on the MELD score.

Although the historic indication for a liver transplant is decompensated cirrhosis, a landmark analysis comparing waiting list mortality with posttransplant mortality established that a minimum MELD score of 18 is necessary to have a survival benefit posttransplant. A MELD score between 15 and 18 does not confer a survival advantage, but a transplant may be justified if the patient has significant morbidity from cirrhosis.

Acute liver failure itself is an indication for a liver transplant. To qualify for Status 1 (first priority for a donor liver within the UNOS region), the transplant candidate must meet the following criteria: (a) onset of hepatic encephalopathy within 8 weeks after the first symptoms of liver disease;
(b) absence of preexisting liver disease; and (c) ventilator dependence, dialysis, or an international normalized ratio (INR) >2.0.

**Contraindications**

In general terms, contraindications to a liver transplant include insufficient cardiopulmonary reserve, uncontrolled malignancy or infection, and refractory noncompliance. Older age is only a relative contraindication: carefully selected recipients over the age of 70 years can achieve satisfactory outcomes.  

Patients with reduced cardiopulmonary reserve are unlikely to survive a liver transplant. Candidates should have a normal ejection fraction. If coronary arterial disease is present, they should undergo revascularization pretransplant. Severe chronic obstructive pulmonary disease (COPD) with oxygen dependence is a contraindication. Severe pulmonary hypertension with a mean pulmonary artery pressure greater than 35 mmHg that is refractory to medical therapy is also a contraindication. Candidates with pulmonary hypertension should be evaluated with a right heart catheterization.

For candidates with alcoholic liver disease, few reliable predictors of posttransplant relapse exist. Most centers require 6 months of abstinence from drugs and alcohol. Insurance companies often make more stringent demands, including random drug screening and 1 year of abstinence.

Uncontrolled infections pretransplant are a substantial risk posttransplant when the patient becomes significantly immunosuppressed. Fungal and multidrug-resistant bacterial infections are relative contraindications. Some centers require an extended period of treatment and documented eradication pretransplant. HIV infection is a relative contraindication; some centers have strict protocols that exclude patients with a history of acquired immunodeficiency syndrome (AIDS)-related illnesses as well as those who are coinfected with HCV.

Ideally, patients with a history of malignancy (with the exception of HCC) should be cured of the cancer pretransplant. In most cases, this means eradication, completion of curative therapy, and absence of recurrence over a certain period of time, which varies by the tumor type, but can be up to 5 years or longer for aggressive tumors (see “Malignancies”).

**Surgical Procedure**

A liver transplant is among the most extensive operations performed, and it can be associated with considerable blood loss. A bilateral subcostal incision with midline extension is used. Mechanical retraction spreads the rib cage to allow access. The ligamentous attachments of the liver are dissected free. The vascular structures are isolated, including the supracheliac and infrahepatic vena cava, the portal vein, and hepatic artery (Fig. 11-15). The bile duct, portal structures, and vena cava are divided, completing the hepatectomy (Fig. 11-16)—often the bloodiest and most difficult part of the operation, particularly in the presence of extensive varices and severe coagulopathy.

After the liver is removed, the anhepatic phase begins. This phase is characterized by the absence of inferior vena cava return to the heart and by portal congestion due to clamping of the portal vein. Significant hemodynamic instability and increased variceal bleeding can occur. Patients who are unable to tolerate this phase can be placed on venovenous bypass, with cannulas drawing blood from the IVC via the femoral vein and via the portal vein, returning it to the systemic circulation via the subclavian vein. Venovenous bypass itself can cause complications, including air embolism, thromboembolism, and trauma to the cannulated vessels.

The donor liver is placed in the orthotopic position. The supracheliac vena caval anastomosis is performed first in an end-to-end fashion, followed by the infrahepatic vena caval and portal anastomosis, both also end-to-end. The liver is then reperfused, often leading to a period of hemodynamic instability and cardiac arrhythmias due to the release of byproducts of ischemia from the donor liver. Coagulopathy also can worsen because of these byproducts as well as fibrinolysis.

The arterial anastomosis between the donor common hepatic or celiac trunk is most often performed with the recipient CHA in an end-to-end fashion. Of course, many variations are possible. After arterial reperfusion, the bile duct anastomosis is performed between the donor and recipient common ducts, also in an end-to-end fashion. If necessary for technical reasons, the recipient common duct can be joined to a Roux-en-Y limb. Some surgeons choose to insert a T-tube or place internal stents in the common bile duct to protect the anastomosis.

The piggyback technique is a common variation of the standard technique. The recipient’s IVC is preserved by carefully dissecting off the posterior aspect of the liver. This added dissection is a disadvantage of this variation, often increasing hepatectomy time and blood loss. The recipient’s liver is
removed by dividing it at the confluence of the hepatic veins. The preserved IVC is an advantage of this variation, allowing venous return from the lower body to the heart during the anhepatic phase and improving renal perfusion. No randomized studies, however, have demonstrated the superiority of the piggyback technique over the standard technique.

**Pediatric Transplants**

Outcomes after pediatric liver transplants are among the best after any type of transplant, with a 1-year survival rate of 90%. The most common indication is biliary atresia. After diagnosis is confirmed, a Kasai procedure is promptly carried out: a Roux-en-Y loop of bowel is directly anastomosed to the hilum of the liver. The Kasai procedure often allows time for the children to grow in size, reducing the risk of a transplant when it is required, as it eventually is in 75% of such children.

The other common indication for a pediatric liver transplant is a metabolic disorder, such as α₁-antitrypsin deficiency, tyrosine metabolism deficiencies, and primary oxalosis. Since the MELD score was developed for adults, pediatric liver allocation is based on an analogous model, the Pediatric End-Stage Liver Disease (PELD) score, which incorporates bilirubin levels, INR, albumin levels, age, and growth failure.

The surgical procedure is similar to the adult procedure. Graft implantation is more challenging, given the pediatric recipient’s smaller vascular structures. As a result, surgical complications are much more common in pediatric recipients. Hepatic artery thrombosis is about three times more common. Donor size matching is very important in the pediatric population and often limits the donor pool for pediatric recipients. To address this issue, deceased donor split-liver transplants and living donor transplants (both described in the following sections) have been developed.

**Deceased Donor Split-Liver Transplants**

A deceased donor allograft can be split into two grafts, most frequently into a left lateral segment for a child and an extended right segment for an adult (Fig. 11-17). It can be done in vivo (during the donor operation) or ex vivo (on the back table after the donor liver is removed). Both techniques have similar outcomes. Increased morbidity is associated with splitting allografts, whether for adult or pediatric recipients; however, the technique is justified given the donor shortage and has been important for improving access to transplants for pediatric recipients.¹⁴⁷

**Living Donor Transplants**

Donation by an adult living donor to an adult recipient requires either the right or left lobe of the liver (Fig. 11-18). Donation by an adult living donor to a pediatric recipient requires the left lateral lobe (Fig. 11-19). Donor safety is paramount. The overall donor mortality rate after donation was 0.4%, and the overall complication rate was 40%, with multiple complications occurring in 19% of the patients. The rate of serious complications resulting in lasting disability was 1.1%, with liver failure or death in 0.4%.¹⁴⁸ Careful donor selection is vital. Potential donors should be medically and psychologically healthy, their hepatic anatomy should be amenable to donation, and absolutely no coercion can occur. A separate donor team should serve as the donor advocate and thoroughly explain all risks.

Careful recipient selection is essential. Transplant candidates also must qualify for a deceased donor liver transplant because a significant number of living donor transplant recipients will eventually require a retransplant. Transplant candidates should be medically fit enough to withstand the rigors of the operation and of the postoperative course with a partial graft. An absolute contraindication is a critical illness: the limited success of such transplants does not justify the risks to the living donor. The obvious advantages of a living donor transplant are that it can be done expeditiously (avoiding the waiting list mortality associated with candidates for a deceased donor transplant) and that it can be planned.

**Postoperative Care**

A liver transplant imposes significant trauma on the major organ systems. Immediately posttransplant, the first goal is to stabilize those systems. Acid-base equilibrium and hemodynamic stability are often difficult to maintain but are essential. Periods of hypotension can increase the risk of hepatic artery thrombosis. Careful attention needs to be paid to ongoing bleeding. Appropriate hemoglobin levels should be maintained. Ongoing bleeding mandates a return trip to the operating room; the rate of reoperation can be as high as 25% among high-risk patients. Transfusion of platelets and fresh frozen plasma must be done prudently because theoretically their administration can increase the risk of hepatic artery thrombosis. Graft function should be evaluated frequently; if it is impaired, an ultrasound is urgently required to assess for the presence of vascular complications.
Evaluation of Graft Function

Evaluation of the graft begins in the operating room. Its appearance overall, any swelling, and the quantity and quality of bile production after reperfusion can help assess function. In the intensive care unit, hemodynamic stability, correction of coagulopathy, euglycemia, successful temperature regulation, clearance of lactic acid, and restoration of neurologic status are all signs of a functioning graft, even before the first set of liver function test results are obtained. Transaminases usually peak by postoperative day 2. An aspartate transaminase (AST) level greater than 2500 IU/L is suggestive of significant injury. Cholestasis usually peaks from

Figure 11-18. A. Hepatic transection completed for right lobe removal. CA = cystic artery; CBD = common bile duct; CD = cystic duct; FL = falciform ligament; IVC = inferior vena cava; LHD = left hepatic duct; LHV = left hepatic vein; MHV = middle hepatic vein; MPV = main portal vein; PHA = proper hepatic artery; RHA = right hepatic artery; RHV = right hepatic vein; RPV = right portal vein; S2, S3, S4 = segments 2, 3, and 4. B. Implantation of the donor right lobe with the MHV. CHA = common hepatic artery. (Reproduced with permission from Gruessner RWG, Benedetti E: Living Donor Organ Transplantation. New York, NY: McGraw-Hill Education; 2008.)
postoperative day 7 to 12. The INR should improve shortly after reperfusion.

In 3% to 4% of patients undergoing a liver transplant, the graft does not function for any identifiable reason, a condition termed primary nonfunction; in such cases, a retransplant is the only option. Some studies suggest that a peak AST level of 5000 IU/L may be predictive of primary nonfunction. Factors associated with primary nonfunction include donor macrosteatosis, prolonged cold and warm ischemic times, and prolonged donor hospital stay.

Complications

Vascular complications occur in about 8% to 12% of recipients and include thrombosis, stenosis, and pseudoaneurysm formation.

The most common vascular complication is hepatic artery thrombosis. Initial reviews suggest that its incidence is between 1.6% and 4%;152 the mortality rate is 50%, even after definitive therapy.153 Early presentation can be quite dramatic, with fulminant hepatic necrosis, primary nonfunction, transaminisis, or fever. Late presentation, however, can be asymptomatic or subtle, with cholangitis, bile leak, mild transaminisis, hepatic abscesses, or failure to thrive. Diagnostic imaging with ultrasound has more than 90% sensitivity and specificity. If hepatic artery thrombosis is identified, urgent reexploration is needed. A thrombectomy or revision of an anastomosis may be successful, but with significant hepatic necrosis, a retransplant is necessary.

Thrombosis of the portal vein is very uncommon. Signs of early thrombosis include liver dysfunction, ascites, and variceal bleeding. Upon diagnosis, an operative thrombectomy should be attempted.

Biliary complications remain the “Achilles’ heel” of liver transplantation, affecting 10% to 35% of these organ recipients. Signs include fever and abdominal pain, with bilious drainage from surgical drains. Diagnosis is made with cholangiography.

Complications manifest themselves as leaks or strictures. Leaks require a reoperation and surgical correction, whereas strictures can most often be managed with radiologic or endoscopic interventions. Two common reconstructions are choledochostomy and choledochojejunostomy. Some centers also routinely use T-tube stents or internal stents. Consensus has not been reached as to which reconstruction technique is superior. Early infectious complications are often associated with initial graft function and pretransplant risk factors. Intra-abdominal infections should raise concerns of a possible bile leak. Fungal infections are often associated with poor graft function. Given the immunosuppressed and compromised state of liver recipients, early infectious complications can be devastating.

The types of opportunistic infections that occur in liver transplant recipients are similar to those that occur in other types of solid organ transplant recipients and are due to suppression of cell-mediated immunity by chronic immunosuppressive drug administration.

Acute rejection occurs in approximately 20% of liver recipients. The first line of treatment is with a high dose of a corticosteroid, which is usually effective; if not, antilymphocyte therapy is initiated. Rejection of the liver (unlike other transplanted organs) does not adversely affect patient or graft survival rates. Maintenance immunosuppression consists of a corticosteroid, tacrolimus, and mycophenolate.

INTESTINE AND MULTIVISCERAL TRANSPLANTATION

After the introduction of long-term total parenteral nutrition (TPN) in the late 1970s and the early success of liver, kidney, and heart transplants, the first attempts at intestine transplants were made. Over the first two decades, the results were dismal. But the introduction of the immunosuppressive drug tacrolimus in the late 1980s led to significant improvement in graft and patient survival rates. Nonetheless, intestine transplants remain the least frequently performed of all transplants, with the lowest graft survival rates.

The main obstacle is the high immunogenicity of the intestine, caused by its abundant lymphoid tissue. High levels of immunosuppression are needed, yet the rejection rate is still high. The microbial colonization of the intestine confers the risk of translocation of pathogenic microorganisms into the recipient’s circulation, causing severe systemic infections. Through the first decade of the 21st century, the survival of patients on long-term TPN was superior to the survival of intestine transplant recipients, so a transplant was considered only as rescue therapy for patients with life-threatening TPN-related complications.
Over the last several years, improvements in surgical techniques, in perioperative and postoperative care, and particularly in immunosuppressive protocols have led to significantly better patient and graft survival rates posttransplant. Recent data indicate that survival rates after an intestine transplant often are better than, or at least similar to, survival rates among patients receiving chronic TPN in the home setting with improved quality of life in selected patients. Today, an intestine or multivisceral transplant is recognized as a feasible treatment.

**Indications and Recipient Selection**

An intestine transplant is indicated for patients with irreversible intestine failure in combination with TPN failure. The definition of intestine failure does not specify the exact length of the remaining intestine. Intestine failure is typically multifactorial. Variables include what part of the small intestine is absent, whether or not the ileocecal valve is present, whether or not the patient underwent an ostomy, and how long the remaining colon is. TPN failure is defined as significant biochemical or pathologic evidence of liver injury, loss of central vein access with thrombosis of at least two central veins, frequent indwelling catheter infection or a single episode of fungal infection, and recurrent episodes of severe dehydration despite IV fluid supplementation.

Indications for a transplant differ between the adult and pediatric population. The leading causes of intestine failure are summarized in Table 11-7. The disease involvement of organs other than the intestine dictates the extent of the operation required. Liver failure is often seen in patients on long-term TPN. If pathologic or biochemical evidence of severe liver damage is combined with signs of portal hypertension, then a combined liver-intestine transplant is the treatment of choice. However, a multivisceral transplant (liver, pancreas, stomach, duodenum, and/or small intestine) might be necessary among children who suffer diffuse intestinal dysmotility syndromes and adults who develop diffuse portomesenteric thrombosis, extensive intra-abdominal desmoid disease encasing the main visceral vascular structures with concurrent short gut syndrome, or massive abdominal trauma.

**Surgical Procedure**

For both the donor and recipient surgery, the key decision is which organs will be transplanted. For an isolated intestine transplant, the blood supply is based on the arterial inflow from the SMA and the venous outflow from the superior mesenteric vein (SMV). Both vessels are isolated at the root of the mesentery.

For a combined liver-intestine transplant, the blood supply is based on the arterial inflow from the celiac axis and SMA, which are procured en bloc with an aortic patch. The liver, duodenum, pancreas, and small intestine—because of their close anatomic relationship—are procured en bloc. If the hepatoduodenal ligament is left intact, no biliary reconstruction is necessary, which virtually eliminates the risk of postoperative biliary complications. Because the entire splanchnic system drains into the liver, venous drainage is achieved by Anastomosis of the hepatic veins to the recipient’s vena cava.

For both an isolated intestine transplant and a combined liver-intestine transplant, the proximal transection of the GI tract occurs at the first portion of the duodenum. For a multivisceral transplant, the stomach is part of the graft; hence, the transection of the GI tract occurs at the distal esophagus. Figures 11-20 to 11-22 show these three main types of transplants.

The vast majority of intestine transplants use a deceased donor organ. However, advances in surgical techniques have made the use of living donors a feasible alternative for either an isolated intestine transplant or a combined liver-intestine transplant. With a living donor, the donor operation is slightly different: for an isolated intestine transplant, 150 to 200 cm of the donor’s ileum, on a vascular pedicle comprising the ileocolic artery and vein, are used (Fig. 11-23); for a combined liver-intestine transplant, performed almost exclusively for pediatric recipients, segments II and III of the donor’s liver are used, in addition to the intestine (Fig. 11-24).

### Table 11-7

**Leading causes of intestine failure**

<table>
<thead>
<tr>
<th>CHILDREN</th>
<th>ADULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gastrochisis</td>
<td>Visceral ischemia secondary to SMA/SMV thrombosis</td>
</tr>
<tr>
<td>Midgut volvulus</td>
<td>Crohn’s disease</td>
</tr>
<tr>
<td>Intestinal atresia</td>
<td>Trauma</td>
</tr>
<tr>
<td>Necrotizing enterocolitis</td>
<td>Mesenteric desmoid tumors</td>
</tr>
<tr>
<td>Microvillus involution disease</td>
<td>Radiation enteritis</td>
</tr>
<tr>
<td>Hirschsprung’s disease</td>
<td>Massive resection secondary to tumors</td>
</tr>
<tr>
<td>Crohn’s disease</td>
<td>Chronic intestinal pseudo-obstruction</td>
</tr>
<tr>
<td>Pseudo-obstruction</td>
<td>Autoimmune enteropathy</td>
</tr>
</tbody>
</table>

*SMA = superior mesenteric artery; SMV = superior mesenteric vein*
Figure 11-21. Combined liver-intestine transplant.

Figure 11-22. Multivisceral transplant.

Figure 11-23. A. Donor operation. About 180 to 200 cm of distal ileum on a vascular pedicle comprising the ileocolic artery and vein are removed. B. Recipient operation. The donor’s ileocolic artery and vein (or the terminal branches of the donor’s superior mesenteric artery and vein) are anastomosed end-to-side to the recipient’s infrarenal aorta and vena cava. (Reproduced with permission from Gruebner RWG, Benedetti E: Living Donor Organ Transplantation. New York, NY: McGraw-Hill Education; 2008.)
BASIC CONSIDERATIONS

HEART AND LUNG TRANSPLANTATION

History

The first successful heterotopic heart transplant, in an animal model, was performed by Carrel and Guthrie in 1905. Subsequent progress with cardiopulmonary bypass and immunologic modulation facilitated the first successful adult human heart transplant, performed by Christiaan Barnard in 1967 in Cape Town, South Africa. However, it was Norman Shumway at Stanford who persisted with heart transplants, in the face of disappointing patient outcomes at a number of early centers.

Postoperative Care

Initial postoperative care for intestine transplant recipients does not significantly differ from that for other organ transplant recipients. In the intensive care unit, each recipient’s cardiovascular, pulmonary, and renal function is closely monitored; aggressive resuscitation with fluid, electrolytes, and blood products is performed. Broad-spectrum antibiotics are an integral component of care.

Of all solid organ transplants, intestine transplants have the highest rate of rejection. With intestine transplants, no serologic marker of rejection is available, so frequent biopsies and histologic evaluation of the intestinal mucosa are of utmost importance. Rejection leads to structural damage of the intestinal mucosa. Translocation of endoluminal pathogens into the circulation can cause systemic infections.

Thanks to the introduction of new immunosuppressive protocols, the rejection rates and the overall patient and graft survival rates have improved significantly. Variations between the protocols exist, but the general concept is to induce immunosuppression with polyclonal T-cell antibody and high doses of a corticosteroid, followed by maintenance doses of corticosteroids and the calcineurin inhibitor tacrolimus.

Immediately, posttransplant, recipients are maintained on TPN. Enteral nutrition is initiated as early as possible, but it is advanced very cautiously. It can take several weeks for the transplanted intestine to achieve structural integrity and functionality and for the recipient to tolerate the full strength of tube feeds.

Despite all the recent advances, the complication rate posttransplant remains high. The most common complications include intra-abdominal abscesses, enteric leaks, intra-abdominal sepsis, the need for a reoperation, graft thrombosis, life-threatening bleeding, and central line problems. Immunosuppression-specific complications include rejection, PTLD, graft-versus-host disease (GVHD), infections, and malignancies. Tailoring the recipient’s immunosuppression plays a critical role in preventing these complications; a low level of immunosuppression leads to graft rejection, but too much confers a high risk of infectious complications, PTLD, and, less commonly, GVHD—all of which are associated with a significantly increased risk of graft failure and mortality.

The long-term results of intestine transplants have improved significantly, even though they still remain inferior to the results of other abdominal organ transplants.159,160

Figure 11-24. Recipient operation. For a combined living donor liver-intestine transplant in a pediatric recipient, liver segments 2 and 3 are implanted in standard fashion (the donor’s left hepatic vein to the recipient’s vena cava, the donor’s left hepatic artery to the recipient’s proper or common hepatic artery, the donor’s left portal vein branch to the recipient’s portal vein trunk). The donor’s ileocolic artery and vein are anastomosed to the recipient’s infra-renal aorta and cava. In the recipient, a duodenum-to-donor ileum anastomosis and a distal Bishop-Coop ileostomy are constructed to reestablish bowel continuity. A very short Roux-en-Y loop (10 to 20 cm) is anastomosed to the donor’s bile duct(s). (Reproduced with permission from Gruessner RWG, Benedetti E: Living Donor Organ Transplantation. New York, NY: McGraw-Hill Education; 2008.)

Similarly, the recipient operation also varies by the organs transplanted. Generally, the recipient’s infrarenal aorta is used to achieve the arterial inflow to the graft. For an isolated intestine transplant, venous drainage is achieved via systemic or portomesenteric drainage; for a combined liver-intestine transplant or a multivisceral transplant, venous drainage is achieved via the hepatic veins. Systemic venous drainage, given its lesser technical difficulty, is preferred over portomesenteric drainage. The diversion of splanchnic flow into the systemic venous circulation can cause several metabolic abnormalities, but no hard evidence shows any negative impact clinically on the recipient.

After the organs are perfused, the continuity of the recipient’s GI tract is restored, which includes the placement of a gastrostomy or jejunostomy feeding tube and an ileostomy. In the early postoperative period, the ileostomy enables regular endoscopic surveillance and biopsy of the intestinal mucosa. Once the recipient recovers, the ileostomy can be taken down.

The last, but often the most difficult, part of the recipient operation is abdominal wall closure. It is especially challenging in intestine transplant recipients because they have usually undergone multiple previous procedures, resulting in many scars, ostomies, feeding tubes, and the loss of abdominal domain. To provide sufficient coverage of the transplanted organs, the use of prosthetic mesh often is necessary.
Thanks to the diligence of Shumway and colleagues in perfecting heart transplant techniques, along with the development, by Caves, of endomyocardial biopsy as a method of allograft rejection surveillance, human heart transplants began to reappear in the 1980s as a viable solution to end-stage heart failure. By 1981, the introduction of cyclosporin A finally created the necessary clinical immunologic modulation necessary to make long-term survival of heart recipients a reality.

Lung transplants have a similar history. In the 1950s, Metras in France and Hardin and Kittle in the United States performed canine lung transplants, demonstrating that meticulous anastomotic technique could produce normal pulmonary pressures. Hardy performed the first human lung transplant in 1963, although the patient lived only 18 days. The first successful long-term lung transplant was performed in 1983 in Toronto. These early lung recipients, however, were plagued by infection, rejection, and, most significantly, bronchial anastomotic dehiscence. Cooper and colleagues soon determined that the high-dose corticosteroids used for immunosuppression were responsible for the frequent occurrence of dehiscence. The combination of high-dose corticosteroids and ischemic donor bronchi was deadly to lung recipients. Cooper, Morgan, and colleagues showed that the bronchial anastomosis could be protected by wrapping it with a vascular omental pedicle, which not only provided neovascularity but also offered a buttress against any partial dehiscence.

Once cyclosporine became available for lung recipients, corticosteroid doses could be quickly tapered and stopped; cyclosporine poses no danger to the integrity of the bronchial anastomosis. In fact, the introduction of cyclosporine allowed the success of the first combined heart-lung transplant at Stanford in 1981 (after unsuccessful attempts by Cooley in 1969, Lillehei in 1970, and Barnard in 1981, all of whom used only high-dose corticosteroids for immunosuppression). The 1980s marked the start of the modern age of thoracic transplants.

Heart Transplants

Indications. The most common diagnosis leading to a heart transplant is ischemic dilated cardiomyopathy, which stems from coronary artery disease, followed by idiopathic dilated myopathy and congenital heart disease. About 3000 patients are added to the waiting list each year.

Evaluation. Pretransplant, both candidates and potential donors are evaluated to ensure their suitability for the procedure. Transplant candidates undergo echocardiography, right and left heart catheterization, evaluation for any undiagnosed malignancies, laboratory testing to assess the function of other organs (such as the liver, kidneys, and endocrine system), a dental examination, psychosocial evaluation, and appropriate screening (such as mammography, colonoscopy, and prostate-specific antigen testing). Once the evaluation is complete, the selection committee determines, at a multidisciplinary conference, whether or not a heart transplant is needed and is likely to be successful. Transplant candidates who meet all of the center’s criteria are added to the waiting list, according to the UNOS criteria, which are based on health status.

Once a potential deceased donor is identified, the surgeon reviews the status report and screening examination results. The donor is initially matched to the recipient per the recipient’s status on the UNOS waiting list, the size match, and the blood type. Results of the donor’s serologic testing, echocardiography, chest X-ray, hemodynamic testing, and possibly coronary artery evaluation are assessed, in order to determine whether or not the donor’s heart can withstand up to 4 hours of cold ischemic time during procurement, transport, and surgery.

Procedure. Heart transplants are most often performed orthotopically (Fig. 11-25). The recipient’s native heart is removed, leaving the superior vena cava, the IVC, the left atrial cuff, the aorta, and the pulmonary artery in situ, in order to allow for anastomosis of the donor’s heart. Usually the left atrial cuff is anastomosed first, providing left heart inflow. Right heart inflow is achieved using a bicaval technique, by directly sewing the donor’s superior vena cava and IVC to the recipient’s venae cavae or by creating an anastomosis of the right atrium to a right atrial cuff. The donor’s main pulmonary artery is connected to the recipient’s pulmonary artery, and finally, the aortic anastomosis is completed (Fig. 11-26).

Once the cross-clamp is removed, the heart is allowed to receive circulation from the recipient and begins to function normally. Inotropic support with isoproterenol, dobutamine, or epinephrine is often required for 3 to 5 days, in order to support recovery from the cold ischemia.

On rare occasions, a heterotopic or “piggyback” heart can be transplanted, leaving the native heart in place. But this scenario is becoming very uncommon with the increasing use of mechanical circulatory support for single-ventricle failure.

Posttransplant Care. Patient survival rates for heart recipients differ slightly after primary transplants vs. retransplants. In 2016, 3209 heart transplants were performed in the United States. New, active listings increased 57% since 2005. Overall 1-year survival for patients who underwent heart transplant in 2009 to 2011 was 90.1%, 3-year survival was 83.5%, and 5-year survival was 76.2%.
survival was 78.3%, and the most common cause of death within the first year after transplant was infection. An increasing number of heart recipients have now survived more than 15 to 20 years with their first graft, especially those with no significant history of either cellular or antibody-mediated rejection.

Heart recipients must be monitored for both early and late complications. Early complications include primary graft dysfunction, acute cellular or antibody-mediated rejection, right heart failure secondary to pulmonary hypertension, and infection. Hemodynamic values are monitored to assess early graft function; pharmacologic and sometimes mechanical support is instituted if needed.

The goal of immunosuppression is to prevent rejection, which is assessed by immunosuppressive levels and, early on, by endomyocardial biopsy. Both T-cell–mediated (cellular) and B-cell–mediated (antibody-mediated) rejection are monitored. Most of the immunosuppression used is aimed at T cells; however, if the recipient has many preformed antibodies or develops donor-specific antibodies, other strategies (such as plasmapheresis or rituximab) are used to reduce the antibody load. Immunosuppressive regimens can vary by center, but most often consist of three categories of medications: a calcineurin inhibitor (usually tacrolimus or cyclosporine), an antiproliferative agent (MMF or AZA), and a corticosteroid (prednisone). Other immunosuppressive agents can be used, depending on the needs of individual recipients.

Recipients are also assessed for any infections, with visual inspection of wound healing and with monitoring of the complete blood count and cultures as needed. Other common early sequelae include drug-induced nephrotoxicity, glucose intolerance, hypertension, hyperlipidemia, osteoporosis, malignancies, and biliary disease.

Late complications include acquired transplant vasculopathy, progressive renal failure, and, most commonly, malignancies, especially skin cancer and PTLD. Accelerated coronary artery disease is the third most common cause of death posttransplant (after infections and acute rejection) and the most common cause after the first year. Coronary artery disease can begin to develop as early as 1 year posttransplant. Its pathogenesis is unknown, but it is believed to be immunologic. Because of these late complications, most transplant centers continue to perform screening tests and recipient examinations at least annually after the first year.

### Lung Transplants

**Indications.** The indications for a lung transplant include congenital disease, emphysema, COPD, cystic fibrosis, idiopathic pulmonary fibrosis, primary pulmonary hypertension, α1-antitrypsin deficiency, and the need for a retransplant after primary graft failure. Each year in the United States, about 1600 patients are added to the waiting list; nearly a third of them have COPD and/or emphysema. The next most common diagnosis among patients on the waiting list is cystic fibrosis. A lung allocation score (LAS) was instituted in 2005. The average lung transplant candidate requires oxygen (often 4 L/min or more at rest) and has an extensively compromised quality of life, as documented by the results of pulmonary function and 6-minute walk tests.

**Evaluation.** Evaluation for a lung transplant is very similar to evaluation for a heart transplant, except that lung transplant candidates undergo more extensive pulmonary function testing, a 6-minute walk test, chest computed tomography, ventilation-perfusion (V-Q) scanning, and arterial blood gas assessment. In addition, all lung transplant candidates must have adequate cardiac function and must meet psychosocial requirements.

Potential lung donors are also screened for blood type and size match. Larger lungs are accepted for COPD patients; smaller lungs are chosen for the restricted chest cavity of fibrotic patients. Donors should have a partial pressure of oxygen in arterial blood (Pao2) value >300 mmHg on a fraction of inspired oxygen (FiO2) of 100% and a positive end-expiratory pressure (PEEP) value of 5. Ideally, donors will have normal chest X-ray results, but exceptions for isolated abnormalities that will not affect subsequent graft function can be made. Living donors can donate a single lobe to a smaller recipient, such as a child. Single-lung transplants are common in many centers and can serve to increase the availability of lungs for multiple recipients. Newer concepts, such as “lung in the box” extracorporeal lung perfusion and stem cell technologies, may further improve the availability of donor lungs by optimizing the use of otherwise marginal grafts.

**Procedure.** Lung transplants can be done either as (a) single-lung transplants (to either side via thoracotomy) or as (b) sequential bilateral-lung transplants (via bilateral thoracotomies or via a single clamshell incision that divides the sternum; Fig. 11-27). They can be done absent extracorporeal mechanical cardiopulmonary perfusion (bypass), with the lung with the worst function (as predicted by preoperative ventilation and perfusion scanning) transplanted first. Despite careful surgical technique and excellent anesthesia, the poor pulmonary reserve of some lung recipients may require the institution of cardiopulmonary bypass to complete the transplant. Bypass is initiated through the chest by direct cardiac cannulation or peripherally via the femoral vessels.

Once the thoracotomy is made, a recipient pneumonectomy is performed with care, in order to avoid injury to the phrenic or recurrent laryngeal nerves. The pulmonary veins and main pulmonary artery are encircled outside the pericardium. At this point, once the main pulmonary vessels are occluded, the need for cardiopulmonary bypass can be assessed.
Posttransplant Care. Patient survival rates for lung recipients vary significantly after primary vs. redo transplants. After primary transplants, the patient survival rates at 1, 3, and 5 years are 83%, 62%, and 46%, respectively; after retransplants, the rates are 64%, 38%, and 28%.

Postoperative care of lung recipients can be very labor-intensive. These patients require meticulous ventilator management, in order to maintain Fio₂ at a minimum and to keep Pao₂ at 70 mmHg. Most patients are extubated within the first 24 to 48 hours. Recipients can require multiple bronchoscopies for both airway management and surveillance biopsies. Diuretics are used generously to counteract any positive fluid balance from the operation and to help with pulmonary recovery.

Early complications include technical complications, graft dysfunction, infections, and rejection. Technical complications often involve stenosis of one or more anastomoses leading to graft dysfunction. Bronchoscopy, V-Q scanning, echocardiography, and radiologic imaging are useful in identifying the causes of graft dysfunction. In up to 20% of recipients, primary early graft dysfunction can occur with no obvious cause. Such dysfunction may be due to some pathology from the donor, perhaps an unknown aspiration, infection, or contusion; or it could result from poor graft preservation at the time of organ procurement. In the intensive care unit, aggressive ventilator and pharmacologic management can help, but recipients can nonetheless progress to the need for mechanical support in the form of ECMO. Infections are treated with appropriate antibiotics, which can be challenging in patients with cystic fibrosis and a history of multidrug-resistant organisms. Rejection is monitored by biopsies and treated as needed.

Late complications include airway complications, such as strictures and, rarely, dehiscence, bronchiolitis obliterans, and malignancies. Strictures are treated with bronchosopic dilation and intervention. Bronchiolitis obliterans often is a sequel of chronic rejection, but can be due to aspiration, chronic infections, or various other causes. In recipients with a progressive fall in their forced expiratory volume in 1 second (FEV₁), bronchiolitis obliterans is suspected. All recipients should be taught to perform microspirometry at home as a screening tool posttransplant. Biopsies are performed to confirm the diagnosis of any complication and, if possible, the cause. Despite aggressive screening and treatment, more than 50% of recipients will develop graft dysfunction. Most if not all of the sequelae of chronic immunosuppression that occur in lung transplant recipients are similar to those occurring in other groups of solid organ transplants.

Heart-Lung Transplants

Every year in the United States, 30 to 50 patients are added to the list of patients waiting to receive a simultaneous heart-lung transplant. The most common diagnosis is idiopathic pulmonary fibrosis, followed by primary pulmonary hypertension. Heart-lung candidates are often younger than their single-organ counterparts. The patient survival rates at 1, 3, and 5 years are 66%, 48%, and 39%, respectively. Often, lung complications ultimately lead to graft failure. The immunosuppression is the same as that for single thoracic organ recipients, with emphasis on weaning the patient off corticosteroids as early as possible.

XENOTRANSPLANTS

Xenotransplants (i.e., cross-species transplants of organs, tissues, or cells) have immense, yet untapped, potential to solve the critical shortage of available grafts. A primary hurdle is the formidable immunologic barrier between species, especially
with vascularized whole organs. Other problems include the potential risk of transmitting infections (known as zoonoses or xenoses) and the ethical problems of using animals for widespread human transplants, even though great progress has been made in the past few years in efforts to overcome these problems.

Pigs are generally accepted as the most likely donor species for xenotransplants into human beings. Pigs would also be easier to raise on a large-scale basis. Guidelines for raising pigs in specialized facilities designated as pathogen-free have been established; in anticipation of clinical trials, such facilities have already been created and populated.

The immunologic barrier in pig-to-human xenotransplants is highly complex, but generally involves four subtypes of rejection. The first is hyperacute rejection (HAR), which is mediated by the presence of natural (preformed) xenoblood factors in humans. These antibodies bind to antigens found mainly on the vascular endothelial cells of porcine donor organs, leading to complement activation, intravascular coagulation, and rapid graft ischemia soon after the transplant. The second subtype is acute humoral xenograft rejection (AHXR), a delayed form of antibody-mediated rejection seen in pig-to-nonhuman-primate transplants after steps to prevent HAR—steps such as depletion of antipig antibodies or complement from nonhuman primates’ serum. Alternative names for AHXR include acute vascular rejection or delayed xenograft rejection. The third subtype is an acute cellular rejection process (similar to the classic T-cell–mediated acute rejection seen in allograft recipients). The fourth subtype is chronic rejection in grafts that survive for more than a few weeks (similar to the chronic rejection seen in long-surviving allograft recipients, with features of chronic vasculopathy).

Many different options are being tested to overcome this immunologic barrier, including the genetic engineering of pigs, the use of agents to inhibit platelet aggregation and complement activation, and the administration of powerful immunosuppressive drugs.

During the first decade of the 21st century, the field of whole-organ xenotransplantation progressed significantly, thanks to the increasing availability of genetically engineered pigs and new immunosuppressive protocols. At a recent symposium organized by the International Xenotransplantation Association, data presented demonstrated extended survival time of porcine solid organs in nonhuman primates: from about 30 days to an average of 60 days and even up to 250 days (depending on the model). However, clinical application is still limited by thrombotic microangiopathy and consumptive coagulopathy; novel methods to prevent those complications will be required for further progress.

Cellular xenotransplants have made great strides and are currently in the early stages of clinical trials. Porcine islet xenotransplants are the most advanced form; five independent groups have now demonstrated survival and function of porcine islets in nonhuman primates for more than 100 days. For the clinical trials, cost-benefit models have been developed, and the regulatory framework has been established. One trial of particular interest involves transplanting encapsulated porcine islets without immunosuppression. Early results are encouraging. But the efficacy of that approach may be limited until further genetic engineering enables proper oxygenation and nourishment of islet grafts, thereby supporting their viability and function.

The future of xenotransplantation is exciting. Continued active research will focus on further genetic engineering of pigs, newer immunosuppressive drugs, and tissue engineering approaches that will minimize or eliminate the need for immunosuppression. Given recent progress, routine clinical application of cellular xenotransplants is likely within the next decade.

REFERENCES

Entries highlighted in bright blue are key references.


BACKGROUND

Patient harm due to medical mistakes can be catastrophic, resulting in high-profile consequences for the patient, surgeon, and institution. A single error can even destroy a surgeon’s career. While mistakes are inherent to human nature, it is becoming more recognized that many mistakes are preventable.

Patient safety is a science that promotes the use of evidence-based medicine and local wisdom to minimize the impact of human error on quality patient care. Wrong-site/wrong-procedure surgeries, retained sponges, unchecked blood transfusions, mismatched organ transplants, and overlooked allergies are all examples of potentially catastrophic events that can be prevented by implementing safer hospital systems. This chapter provides an overview of the modern-day field of patient safety by reviewing key measures of safety and quality, components of culture, interventions and tools, assessment methods, risk management strategies, and a selected review of common complications in surgery.

Medical Care Gone Wrong

Today, there are more medications, diagnoses, procedures, and handoffs performed than ever in the history of medicine. Moreover, overtreatment is now an endemic problem in some areas of healthcare. With more medical care being delivered, there are naturally more opportunities for things to go wrong. In fact, harm may be associated with complexity. The Commonwealth Fund reported that the United States leads the world in medical errors, observing that 34% of patients with health problems in the United States report experiencing medical, medication, or test errors—the highest rate of any nation, and an analysis suggests that the problem of medical care gone wrong, i.e., medical errors including systems errors, may rank as the third leading cause of death in the United States.¹

Medical error is defined as an unintended act (either omission or commission) or one that does not achieve its intended outcome, the failure of a planned action to be completed as intended (an error of execution) or the use of a wrong plan to achieve an aim (an error of planning), and a deviation from the process of care, which may or may not cause harm to the patient. Medical error can occur at the individual provider level or at the system level. An expanding taxonomy is maturing to better categorize the types of factors and events that are avoidable. The role of error may be complex; error can sometimes tragically end the life of a thriving person with a long life expectancy, or it can also accelerate an imminent death.

The most commonly cited report on the incidence of deaths due to medical error, the 1999 Institute of Medicine (IOM) report, describes an incidence of 44,000 to 98,000 deaths annually.² However, this estimate by the IOM was not based on primary research conducted by the IOM; rather, it was based on two older studies conducted in 1984 and 1992. Both studies were small and limited. In 2013, after compiling more recent evidence from multiple sources, James estimated an incidence range of 210,000 to 400,000 deaths a year associated with
Key Points

1. Medical error ranks as the third leading cause of death in the United States when defined to include system errors.
2. One form of medical error is unnecessary or excessive medical care, which represents 21% of medical care administered in the United States.
3. New peer-comparison metrics evaluate appropriateness of surgical care by measuring a physician’s practice pattern among all the physician’s patients benchmarked to the physician’s peers.
4. Judicious opioid prescribing upon discharge after surgery is critical given the magnitude of the opioid crisis.
5. The structure-process-outcome framework within the context of an organization’s culture helps to clarify how risks and hazards embedded within the organization’s structure may potentially lead to error and injure or harm patients.
6. Poor communication contributes to approximately 60% of the sentinel events reported to The Joint Commission.
7. Operating room briefings are team discussions of critical issues and potential hazards that can improve the safety of the operation and have been shown to improve operating room culture and decrease operating room delays.
9. The most important determinant of malpractice claims against a surgeon is patient rapport, not undertesting.

Unnecessary Medical Care

Increasingly preventable complications and complications from unnecessary procedures are considered to be forms of medical error. Unnecessary medical care accounts for an estimated $210 billion in excess spending each year, according to the National Academy of Medicine. The issue represents a significant opportunity to make improve patient safety and lower healthcare costs.

Based on our estimate, medical error is the 3rd most common cause of death in the US

Figure 12-1. Causes of death in the United States 2013. (Reproduced with permission from Makary MA, Daniel M. Medical error—the third leading cause of death in the US, BMJ. 2016 May 3;353:i2139.)
costs. In a Johns Hopkins study, surveying over 2000 physicians in the United States, unnecessary medical care was reported to be common. On average, these authors reported that 21% of medical care is unnecessary. Breaking the problem down by type of medical care, the doctors reported that 22% of prescription medications, 25% of medical tests, and 11% of procedures are unnecessary. These perceptions by U.S. physicians validate previous estimates of the National Academy of Medicine that suggest that one-third of healthcare spending is wasteful and does not result in better health. Addressing avoidable medical and surgical care is a topic gaining increasing recognition in healthcare.

One example of overtreatment in surgical care is opioid over-prescribing. In the United States in 2015 alone, clinicians handed out 249 million opioid prescriptions, almost one for every American adult. And in 2016, the United States produced 14 billion opioid pills (40 for every American citizen). With the exception of pain specialists treating patients with pain syndromes, surgeons are the most common prescribers of opioids. Judicious opioid prescribing is important because of the addictive potential of these medications. Moreover, many patients can recover comfortably after hospital discharge with nonopioid or nonaddictive pain regimens.

**THE SCIENCE OF PATIENT SAFETY**

Medicine is considered a high-risk system with a high error rate, but these two characteristics are not always correlated. Other high-risk industries have managed to maintain an impeccably low error rate. For example, one of the highest risk systems in existence today, the U.S. Navy’s nuclear submarine program, has an unmatched safety record.

Much of the credit for their safety record is due to the culture of the nuclear submarine program, with its insistence on individual ownership, responsibility, attention to detail, professionalism, moral integrity, and mutual respect. These characteristics have created the cultural context necessary for high-quality communications under high-risk, high-stress conditions. Each reactor operator is aware of what is going on at all times and is responsible for understanding the implications and possible consequences of any action. Communication flows freely between crewmen and officers, and information about any mistakes that occur are dispersed rapidly through the entire system so that other workers can learn how to prevent similar mistakes in the future.

**High Reliability Organizations**

The nuclear submarine program is an example of an organization that has achieved the distinction of being considered a “high reliability organization.” High reliability organization theory recognizes that there are certain high-risk industries and organizations that have achieved very low accident and error rates compared to what would be expected given the inherent risks involved in their daily operations. Other high reliability industries and organizations include aircraft carrier flight decks, nuclear power plants, and the Federal Aviation Administration’s air traffic control system. In fact, one reason why nuclear power plants have such an excellent reliability record may be that their operators are often former naval submarine officers whose previous experience and training within one highly reliable organization are easily transferable to other organizations.

One of the assumptions underlying the science of high reliability organizations is that humans who operate and manage complex systems are themselves not sufficiently complex to sense and anticipate the problems generated by the system. This introduces another important idea undergirding the science of patient safety: the concept of normal accident theory. Instead of attributing accidents to individual error, this theory states that accidents are intrinsic to high-volume activities and even inevitable in some settings. Accidents should not be used merely to identify and punish the person at fault, but should be seen as a systems problem and addressed at a broader level. As Ruchlin states, even the “best people can make the worst errors as a result of latent conditions.”

High-risk systems, as defined by Perrow in 1984:

- Have the potential to create a catastrophe, loosely defined as an event leading to loss of human or animal life, despoiling of the environment, or some other situation that gives rise to the sense of “dread.”
- Are complex, in that they have large numbers of highly interdependent subsystems with many possible combinations that are nonlinear and poorly understood.
- Are tightly coupled, so that any perturbation in the system is transmitted rapidly between subsystems with little attenuation.

However, high reliability organization theory suggests that proper oversight of people, processes, and technology can handle complex and hazardous activities and keep error rates acceptably low. Studies of multiple high reliability organizations show that they share the following common characteristics:

- People are supportive of one another.
- People trust one another.
- People have friendly, open relationships emphasizing credibility and attentiveness.
- The work environment is resilient and emphasizes creativity and goal achievement, providing strong feelings of credibility and personal trust.

Developing these characteristics is an important step toward achieving a low error rate in any organization. For this reason, safety culture is a measure used by hospitals nationwide to improve outcomes and is increasingly recognized as a metric of hospital quality.

**The Conceptual Model**

The Donabedian model of measuring quality identifies three main types of improvements: changes to organizational structure, changes in organizational processes, and changes in outcomes. Structure refers to the physical and organizational tools, equipment, and policies that improve safety. Structural measures ask, “Do the right tools, equipment, and policies exist?” Process is the application of these tools, equipment, and policies/procedures to patients (good practices and evidence-based medicine). Process measures ask, “Are the right tools, policies, and equipment being used?” Outcome is the result on patients. Outcome measures ask, “How often are patients harmed?” In this model, structure (how care is organized) plus process (what we do) influences patient outcomes (the results achieved).

The structure, process, and outcome components of quality measurement all occur within the context of an organization’s overall culture. The local culture impacts all aspects of the delivery of care because it affects how front-line personnel
understand and deliver safe patient care. In fact, culture (collective attitudes and beliefs of caregivers) is increasingly being recognized to be the fourth measurable component to the structure-process-outcome model. This recognition is based on growing evidence that local culture is linked to a variety of important clinical outcomes. For any new patient safety initiative to be deemed successful, any change in structure or process must lead to a corresponding positive change in patient outcomes.

### CREATING A CULTURE OF SAFETY

Culture is to an organization what personality is to the individual—a hidden, yet unifying theme that provides meaning, direction, and mobilization. Organizations with effective safety cultures share a constant commitment to safety as a top-level priority that permeates the entire organization. These organizations frequently share the following characteristics:

- An acknowledgment of the high-risk, error-prone nature of an organization’s activities
- A nonpunitive environment where individuals are able to report errors or close calls without fear of punishment or retaliation
- An expectation of collaboration across ranks to seek solutions to vulnerabilities
- A willingness on the part of the organization to direct resources to address safety concerns

Traditional surgical culture stands almost in direct opposition to the values upheld by organizations with effective safety cultures for several reasons. Surgeons are less likely to acknowledge their propensity to make mistakes or to admit these mistakes to others. Surgeons tend to minimize the effect of stress on their ability to make decisions. The surgical culture, especially in the operating room (OR), is traditionally rife with hierarchy. Intimidation of other OR personnel by surgeons was historically accepted as the norm. This can prevent nurses and other OR staff from pointing out potential errors or

### Table 12-1  Systems change resulting from medical error

<table>
<thead>
<tr>
<th>Types of medical error</th>
<th>Adverse event</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Injury caused by medical management rather than the underlying condition of the patient</td>
<td></td>
</tr>
<tr>
<td>• Prolongs hospitalization, produces a disability at discharge, or both</td>
<td></td>
</tr>
<tr>
<td>• Classified as preventable or unpreventable</td>
<td></td>
</tr>
<tr>
<td>Negligence</td>
<td></td>
</tr>
<tr>
<td>• Care that falls below a recognized standard of care</td>
<td></td>
</tr>
<tr>
<td>• Standard of care is considered to be care a reasonable physician of similar knowledge, training, and experience would use in similar circumstances</td>
<td></td>
</tr>
<tr>
<td>Near miss</td>
<td></td>
</tr>
<tr>
<td>• An error that does not result in patient harm</td>
<td></td>
</tr>
<tr>
<td>• Analysis of near misses provides the opportunity to identify and remedy system failures before the occurrence of harm</td>
<td></td>
</tr>
<tr>
<td>Sentinel event</td>
<td></td>
</tr>
<tr>
<td>• An unexpected occurrence involving death or serious physical or psychological injury</td>
<td></td>
</tr>
<tr>
<td>• The injury involves loss of limb or function</td>
<td></td>
</tr>
<tr>
<td>• This type of event requires immediate investigation and response</td>
<td></td>
</tr>
<tr>
<td>• Other examples</td>
<td></td>
</tr>
<tr>
<td>• Hemolytic transfusion reaction involving administration of blood or blood products having major blood group incompatibilities</td>
<td></td>
</tr>
<tr>
<td>• Wrong-site, wrong-procedure, or wrong-patient surgery</td>
<td></td>
</tr>
<tr>
<td>• A medication error or other treatment-related error resulting in death</td>
<td></td>
</tr>
<tr>
<td>• Unintentional retention of a foreign body in a patient after surgery</td>
<td></td>
</tr>
</tbody>
</table>

Case 12-2  High-profile sentinel event

On December 3, 1994, Betsy Lehman, a *Boston Globe* health columnist, died as a result of receiving four times the intended dose of chemotherapy for breast cancer. Remarkably, 2 days later, Maureen Bateman, a teacher being treated for cancer, also received a chemotherapy overdose and suffered irreversible heart damage. After investigating the medication errors, the prescribing doctor, three druggists, and 15 nurses were disciplined by state regulators. The hospital was sued by the two women’s families and by one of the doctors disciplined.

As a result of this widely publicized event, the Dana-Farber Cancer Institute invested more than $11 million to overhaul their safety programs, including providing new training for their employees and giving doctors more time to meet with patients. The hospital adopted a full disclosure policy so that patients would be informed anytime a mistake had affected their care. Dana-Farber also started a patient committee providing advice and feedback on ways to improve care at the hospital.

Although perceptions of teamwork climate can differ as a function of one’s role in the OR, perceptions of safety climate are relatively consistent across OR providers in a given hospital. Validated in over 300 hospitals, the SAQ is used to establish benchmark safety culture scores by healthcare worker type, department, and hospital. Using this survey, hospitals can compare culture between different types of healthcare workers within a department as well as culture between departments throughout the institution. Scores can be compared to those of other participating institutions to compare safety climates. This allows hospitals to participate with one another to implement programs to improve safety culture. In addition, scores are used to evaluate the effectiveness of safety interventions by comparing the SAQ safety climate scores after implementation to baseline scores.

Strong teamwork is at the core of any effective organization and is a key element to ensuring patient safety in the OR. Teamwork is dependent on the underlying culture and patterns of communication. The ability for all team members, to “speak up” about patient safety concerns is one of the most important elements of creating a culture of patient safety.

### TEAMWORK AND COMMUNICATION

According to The Joint Commission, communication breakdown is one of the top three root causes of sentinel events such as wrong-site surgery (Fig. 12-2). Poor communication contributed to over 60% of sentinel events reported to The Joint Commission in 2011. Good communication is an essential component of teamwork and is especially important in the OR, one of the most complex work environments in healthcare.

Within the realm of patient care, there are enormous amounts of information being exchanged between healthcare providers on a daily basis. Much of this information, if prioritized correctly, has the potential to prevent unintended medical errors and serious harm to patients. The importance of good communication in preventing medical errors is undeniable; however, it is difficult to achieve. The traditional surgical hierarchy can prevent OR personnel from sharing important patient data and expressing safety concerns. One perioperative field study showed a 30% rate of communication failure in the OR, with 36% of these breakdowns having a substantial impact on patient safety.

In addition to overcoming the cultural barrier to better teamwork and communication, the prospective study by Christian and associates of patient safety in the OR demonstrated that the standard workflow of the OR itself presents many opportunities for the loss or degradation of critical information. Hand-offs of patient care from the OR to other locations or providers are particularly prone to information loss, which has been demonstrated in other clinical settings. Hand-offs and auxiliary tasks, such as surgical sponge and instrument counts, frequently take place during critical portions of the case and place competing demands on provider attention from primary patient-centered activities. Communication between the surgeon and pathologist also is vulnerable because the communication often occurs through secondary messengers such as nurses or technicians. This information loss can lead to delays, overuse of
staff and resources, uncertainty in clinical decision making and planning, and oversights in patient preparation.

**Measuring Teamwork**

Research in commercial aviation has demonstrated a strong correlation between better teamwork and improved safety performance. Cockpit crew members’ reluctance to question a captain’s judgment has been identified as a root cause of aviation accidents. Good attitudes about teamwork are associated with error-reduction behaviors in aviation, improved patient outcomes in ICUs, and decreased nurse turnover in the OR. It is also associated with higher job satisfaction ratings and less sick time taken from work.

The SAQ can be used to measure teamwork and provide benchmarks for departments or hospitals seeking to measure and improve their teamwork climate. The SAQ teamwork scores are responsive to interventions that aim to improve teamwork among operating teams, such as the implementation of ICU checklists, executive walk rounds, and preoperative briefing team discussions. The communication and collaboration sections of the SAQ reflect OR caregiver views on teamwork and can be used to distinguish meaningful interventions from impractical and ineffective programs.

In a survey of OR personnel across 60 hospitals, the SAQ identified substantial differences in the perception of teamwork in the OR depending on one’s role. Physicians frequently rated the teamwork of others as good, while nurses at the same institutions perceived teamwork as poor (Fig. 12-3). Similiar

**COMMUNICATION TOOLS**

High reliability organizations such as aviation frequently use tools such as prompts, checks, standard operating protocols, and communication interventions such as team briefings and debriefings. These tools identify and mitigate hazards and allow an organization to complete tasks more efficiently. They also foster a culture of open communication and speaking up if a team member senses a safety concern. Safety checks and standardized team discussions serve as prompts to help “engineer out” human error, providing quality assurance and improving information flow. They also can prevent errors related to omissions, which are more likely to occur when there is information overload, multiple steps in a process, repetitions in steps, and planned departures from routine processes, and when there are other interruptions and distractions present while the process is being executed. These same interventions have been shown to improve patient safety in ORs and ICUs.

**Operating Room Briefings (A Surgical Checklist)**

Preoperative briefings and checklists, when used appropriately, help to facilitate transfer of information between team members (Table 12-3). A briefing, or checklist, is any preprocedure discussion of requirements, needs, and special issues of the procedure. Briefings often are locally adapted to the specific needs of the specialty. They have been associated with an improved safety culture, including increased awareness of wrong-site/wrong-procedure errors, early reporting of equipment problems, reduced operational costs and fewer unexpected delays.

In one study, 30.9% of OR personnel reported a delay before the
institution of OR briefings, and only 23.3% reported delays after briefings were instituted. OR briefings are increasingly being used to ensure evidence-based measures are used, such as the appropriate administration of preoperative antibiotics and deep vein thrombosis (DVT) prophylaxis. Briefings allow personnel to discuss potential problems, before they become a “near miss” or cause actual harm.

Operating Room Debriefings

Postprocedural debriefings improve patient safety by allowing for discussion and reflection on causes for errors and critical incidents that occurred during the case. Errors or critical incidents are regarded as learning opportunities rather than cause for punishment. During the debriefing, the team also can discuss what went well during the case and designate a point person to follow up on any proposed actions that result from the discussion. In addition, most debriefings include a verification of the sponge, needle, and instrument counts and confirmation of the correct labeling of the surgical specimen.

Errors in surgical specimen labeling have not received as much attention as incorrect sponge or instrument counts as an indicator of the quality of communication in the OR. However, an error in communication or during the hand-off process increases the risk of mislabeling a surgical specimen before its arrival in a pathology laboratory. In one study, this type of identification error occurred in 4.3 per 1000 surgical specimens, which implies an annualized rate of occurrence of 182 mislabeled specimens per year (Fig. 12-5). Errors involving specimen identification can result in delays in care, the need for an additional biopsy or therapy, failure to use appropriate therapy, or therapy administered to the wrong body site, side, or patient. These errors can lead to significant harm to the patient, costs to the institution, and distrust by a community. Given the frequency of occurrence and the feasibility and validity of measuring them, mislabeled surgical specimens may serve as a useful indicator of patient safety and should be included in any postprocedural debriefing checklist.

Sign Outs

In healthcare, information frequently passes to covering providers without prioritizing potential concerns. This makes sign outs a very vulnerable process of care, which can lead to catastrophic events.

The term sign out can refer to either the verbal or written communication of patient information to familiarize oncoming physicians about patients who will be under their care. Sign outs should occur whenever a patient’s care setting or provider is changing. When performed well, sign outs help to ensure the transfer of pertinent information. However, previous studies have shown the hand-off process to be variable, unstructured, and prone to error. Common categories of communication failure during sign outs include content omissions, such as failure to mention active medical problems, and failures in the actual communication process, such as leaving illegible or unclear notes (Case 12-3). These failures lead to confusion and uncertainty by the covering physician during patient care decisions, resulting in the delivery of inefficient and suboptimal care.

The use of more structured verbal communication such as the Situational Debriefing Model, otherwise known as SBAR (situation, background, assessment, and recommendation), used by the U.S. Navy, can be applied to healthcare to improve the communication of critical information in a timely and orderly fashion. In addition, all sign outs should begin with the statement, “In this patient, I am most concerned about . . .” to signal to the healthcare provider on the receiving end the most important safety concerns regarding that specific patient.

Implementation

Tools such as checklists, sign outs, briefings, and debriefings improve communication between healthcare providers and create a safer patient environment (Fig. 12-6). Although their use in healthcare is still highly variable, specialties that

---

**Table 12-2**

<table>
<thead>
<tr>
<th>CAREGIVER POSITION PERFORMING RATING</th>
<th>SURGEON</th>
<th>ANESTHESIOLOGIST</th>
<th>NURSE</th>
<th>CRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgeon</td>
<td>85</td>
<td>84</td>
<td>88</td>
<td>87</td>
</tr>
<tr>
<td>Anesthesiologist</td>
<td>70</td>
<td><strong>96</strong></td>
<td>89</td>
<td>92</td>
</tr>
<tr>
<td>Nurse</td>
<td><strong>48</strong></td>
<td>63</td>
<td>81</td>
<td>68</td>
</tr>
<tr>
<td>CRNA</td>
<td>58</td>
<td>75</td>
<td>76</td>
<td>93</td>
</tr>
</tbody>
</table>

The best teamwork scores were recorded by anesthesiologists when they rated their teamwork with other anesthesiologists (“high” or “very high” 96% of the time). The lowest teamwork ratings were recorded by nurses when they rated their teamwork with surgeons (“high” or “very high” 48% of the time). CRNA = certified registered nurse anesthetist.


---

**Table 12-3**

**Five-point operating room briefing**

What are the names and roles of the team members? Is the correct patient/procedure confirmed? (The Joint Commission Universal Protocol [TIME-OUT])

Have antibiotics been given? (if appropriate) What are the critical steps of the procedure? What are the potential problems for the case?

---

### Surgical Safety Checklist

<table>
<thead>
<tr>
<th>Before induction of anaesthesia</th>
<th>Before skin incision</th>
<th>Before patient leaves operating room</th>
</tr>
</thead>
<tbody>
<tr>
<td>(with at least nurse and anaesthetist)</td>
<td>(with nurse, anaesthetist and surgeon)</td>
<td>(with nurse, anaesthetist and surgeon)</td>
</tr>
</tbody>
</table>

- **Has the patient confirmed his/her identity, site, procedure, and consent?**
  - Yes
  - Not applicable
- **Is the site marked?**
  - Yes
  - Not applicable
- **Is the anaesthesia machine and medication check complete?**
  - Yes
- **Is the pulse oximeter on the patient and functioning?**
  - Yes
- **Does the patient have a: **
  - Known allergy?
    - No
    - Yes
  - Difficult airway or aspiration risk?
    - No
    - Yes, and equipment/assistance available
- **Risk of >500 ml blood loss (7ml/kg in children)?**
  - No
  - Yes, and two IVs/central access and fluids planned

- **Confirm team members have introduced themselves by name and role.**
- **Confirm the patient’s name, procedure, and where the incision will be made.**
- **Has antibiotic prophylaxis been given within the last 60 minutes?**
  - Yes
  - Not applicable

### Anticipated Critical Events

- **To Surgeon:**
  - What are the critical or non-routine steps?
  - How long will the case take?
  - What is the anticipated blood loss?
- **To Anaesthetist:**
  - Are there any patient-specific concerns?
- **To Nursing Team:**
  - Has sterile (including indicator results) been confirmed?
  - Are there equipment issues or any concerns?

- **Is essential imaging displayed?**
  - Yes
  - Not applicable

---

This checklist is not intended to be comprehensive. Additions and modifications to fit local practice are encouraged.

Revised 1/2009 © WHO, 2009

---


---

have incorporated them, such as intensive care and anesthesia, have made impressive strides in patient safety. Currently, communication breakdowns, information loss, hand off, multiple competing tasks, and high workload are considered “annoying but accepted features” of the perioperative environment.20 As physician attitudes toward errors, stress, and teamwork in medicine become more favorable toward the common goals of reducing error and improving teamwork and communication, medicine will likely achieve many of the milestones in safety that high-reliability industries such as aviation have already accomplished.

### COMPREHENSIVE UNIT-BASED SAFETY PROGRAM

As medical care and hospitals continue to expand, the care that is provided to patients is becoming more fragmented. This fragmentation makes communication more difficult and opportunities for medical errors more common. These problems require common sense solutions, often necessitating a change in the way that care is delivered on the local level. Unit-based meetings to discuss processes that are potentially dangerous for patients can quickly bring danger areas out into the open. These meetings

---

**Figure 12-5.** Incidence of identification errors observed per 1000 specimens (n = 21,351). *(Reproduced with permission from Makary MA, Epstein J, Pronovost PJ, et al. Surgical specimen identification errors: a new measure of quality in surgical care, Surgery. 2007 Apr;141(4):450-455.)*
Case 12-3 Inadequate sign out leading to medical error

Josie King was an 18-month-old child who was admitted to Johns Hopkins Hospital in January of 2001 for first- and second-degree burns. She spent 10 days in the pediatric intensive care unit and was well on her way to recovery. She was transferred to an intermediate care floor with the expectation that she would be sent home in a few days.

The following week, her central line was removed, but nurses would not allow Josie to drink anything by mouth. Around 1 pm the next day, a nurse came to Josie’s bedside with a syringe of methadone. Although Josie’s mother told the nurse that there was no order for narcotics, the nurse insisted that the orders had been changed and administered the drug. Josie’s heart stopped, and her eyes became fixed. She was moved to the pediatric intensive care unit and placed on life support. Two days later, on February 22, 2001, she died from severe dehydration.

After her death, Josie’s parents, Sorrel and Jay King, were motivated to work with leaders at Johns Hopkins to ensure that no other family would have to endure the death of a child due to medical error. They later funded the Josie King Patient Safety Program and an academic scholarship in the field of safety.

Implementing CUSP

The implementation of the Comprehensive Unit-based Safety Program (CUSP) involves measurement of a unit’s safety culture prior to starting the program and inclusion of hospital management from the start. Having management involved allows for more efficient allocation of resources and allows them to better understand the problems faced by front-line providers. Once CUSP is in place, changes can be made using local wisdom to advance patient care. The impact of changes made using CUSP can be measured using both patient outcomes and safety culture data.

Implementation of CUSP has been associated with improved patient outcomes, including decreased surgical site infections. In a 2-year study of colorectal patients, where the first year was pre-CUSP implementation and the second year was post-CUSP implementation, there was a 33% decrease in the surgical site infection rate after CUSP. In this study, the CUSP group met monthly and came up with a list of interventions based on their experience with these cases, including standardization of skin preparation and warming of patients in the preanesthesia area. This study showed that CUSP can be highly effective in ameliorating patient harm and improving patient care.

MEASURING QUALITY IN SURGERY

Despite the newfound focus on patient safety in surgery and the number of initiatives being undertaken by many organizations to improve their safety culture, there are few tools to actually measure whether these efforts are effective in reducing the number of errors. Several agencies and private groups have developed criteria to evaluate quality and safety within hospitals.

Practice Pattern Measures

New quality measures in healthcare focus on the appropriateness of medical care. These appropriateness indicators are doctor-defined and specialty-specific so they are smart and fair. One of the first of these new appropriateness metrics is the average number of tissue blocks a skin cancer (Mohs) surgeon will use to surgically remove a skin cancer. The American College of Mohs Surgeons formalized and endorsed the surgeon metric: average number of blocks a surgeon requires to remove a standardized skin cancer. In a report describing the national distribution of surgeons by their mean number of blocks per case, the national average was found to be 1.7 blocks per surgeon. Statistical outlier surgeons had an average four or more blocks per patient. Boundaries of normal variation was determined by expert physician leaders to define an acceptable range and an

![Figure 12-6](image-url)
unacceptable range (greater than two standard deviations from the national norm). The American College of Mohs Surgeons sent letters to outliers, letting them know where they stand, and offered coaching and retraining help. The new Mohs surgery metric demonstrates the opportunity to reduce unwarranted clinical variation and lower healthcare costs by simply using clinical wisdom and the power of peer-comparison.

Appropriateness measures approach quality differently than traditional quality measures and rely on expert physicians to define the metric and set boundaries of reasonable versus unsafe variation in an individual physician’s practice pattern relative to his or her peers nationally. This concept is being applied to utilization rates of minimally invasive surgery in candidate patients as well as rates of physical therapy utilization before elective spine surgery for chronic pain.

Agency for Healthcare Research and Quality Patient Safety Indicators
The Agency for Healthcare Research and Quality (AHRQ) was created in 1989 as a Public Health Service agency in the Department of Health and Human Services. Its mission is to improve the quality, safety, efficiency, and effectiveness of healthcare for all Americans. Nearly 80% of the AHRQ’s budget is awarded as grants and contracts to researchers at universities and other research institutions across the country. The AHRQ sponsors and conducts research that provides evidence-based information on healthcare outcomes, quality, cost, use, and access. It has advocated the use of readily available hospital inpatient administrative data to measure healthcare quality. The information helps healthcare decision makers make more informed decisions and improve the quality of healthcare services.29

One of the major contributions of the AHRQ is a set of Patient Safety Indicators (PSIs), initially released in 2003 and revised in 2010. PSIs are a tool to help health system leaders identify potential adverse events occurring during hospitalization. Developed after a comprehensive literature review, analysis of International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes, review by a clinician panel, implementation of risk adjustment, and empirical analyses, these 27 indicators provide information on potential in-hospital complications and adverse events following surgeries, procedures, and childbirth (Table 12-4).

Provider-level indicators provide a measure of the potentially preventable complications for patients who received their initial care and the complication of care within the same hospitalization. They include only those cases where a secondary diagnosis code flags a potentially preventable complication. Area-level indicators capture all cases of the potentially preventable complications that occur in a given area (e.g., metropolitan area or county), either during their initial hospitalization or resulting in subsequent hospitalization.30

Currently, PSIs are considered indicators, not definitive measures, of patient safety concerns. They can identify potential safety problems that merit further investigation. They also can be used to better prioritize and evaluate local and national initiatives, and even as benchmarks for tracking progress in patient safety. In the future, further growth in electronic health data will make administrative data-based tools like the PSIs more useful.31

The Surgical Care Improvement Project Measures
The Surgical Care Improvement Project (SCIP) was established in 2003 by a national partnership of organizations committed to improving surgical care by reducing surgical complications. The steering committee is comprised of groups such as the Centers for Medicare & Medicaid Services, the American Hospital Association, Centers for Disease Control and Prevention (CDC), Institute for Healthcare Improvement, The Joint Commission, and others.

The incidence of postoperative complications ranges from 6% for patients undergoing noncardiac surgery to more than 30% for patients undergoing high-risk surgery. Common postoperative complications include surgical site infections (SSIs), myocardial infarction, postoperative pneumonia, and thromboembolic complications. Patients who experience postoperative complications have increased hospital length of stay (3 to 11 days longer than those without complications), increased hospital costs (ranging from $1398 for an infectious complication to $18,310 for a thromboembolic event), and increased mortality (median patient survival decreases by up to 69%).32

Despite well-established evidence that many of these adverse events are preventable, failure to comply with standards

### Table 12-4

| Agency for Healthcare Research and Quality patient safety indicators |
|-------------------------|--------------------------|
| **Provider-level patient safety indicators** |
| • Complications of anesthesia |
| • Death in low mortality diagnosis-related groups |
| • Decubitus ulcer |
| • Failure to rescue |
| • Foreign body left in during procedure |
| • Iatrogenic pneumothorax |
| • Selected infections due to medical care |
| • Postoperative hip fracture |
| • Postoperative hemorrhage or hematoma |
| • Postoperative physiologic and metabolic derangements |
| • Postoperative respiratory failure |
| • Postoperative pulmonary embolism or deep vein thrombosis |
| • Postoperative sepsis |
| • Postoperative wound dehiscence in abdominopelvic surgical patients |
| • Accidental puncture and laceration |
| • Transfusion reaction |
| • Birth trauma—injury to neonate |
| • Obstetric trauma—vaginal delivery with instrument |
| • Obstetric trauma—vaginal delivery without instrument |
| • Obstetric trauma—cesarean delivery |
| **Area-level patient safety indicators** |
| • Foreign body left in during procedure |
| • Iatrogenic pneumothorax |
| • Selected infections due to medical care |
| • Postoperative wound dehiscence in abdominopelvic surgical patients |
| • Accidental puncture and laceration |
| • Transfusion reaction |
| • Postoperative hemorrhage or hematoma |

CHAPTER 12
QUALITY, PATIENT SAFETY, ASSESSMENTS OF CARE, AND COMPLICATIONS

The Surgical Care Improvement Project measures

<table>
<thead>
<tr>
<th>Process of care performance measures</th>
<th>Infection</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Prophylactic antibiotic received within 1 h before surgical incision</td>
<td></td>
</tr>
<tr>
<td>• Prophylactic antibiotic selection for surgical patients</td>
<td></td>
</tr>
<tr>
<td>• Prophylactic antibiotics discontinued within 24 h after surgery end time (48 h for cardiac patients)</td>
<td></td>
</tr>
<tr>
<td>• Cardiac surgery patients with controlled 6 A.M. postoperative serum glucose</td>
<td></td>
</tr>
<tr>
<td>• Surgery patients with appropriate hair removal</td>
<td></td>
</tr>
<tr>
<td>• Colorectal surgery patients with immediate postoperative normothermia</td>
<td></td>
</tr>
</tbody>
</table>

Venous thromboembolism

• Surgery patients with recommended venous thromboembolism prophylaxis ordered
• Surgery patients who received appropriate venous thromboembolism prophylaxis within 24 h before surgery to 24 h after surgery

Cardiac events

• Surgery patients on a β-blocker prior to arrival who received a β-blocker during the perioperative period

Proposed outcome measures

Infection
• Postoperative wound infection diagnosed during index hospitalization

Venous thromboembolism
• Intra- or postoperative pulmonary embolism diagnosed during index hospitalization and within 30 d of surgery
• Intra- or postoperative deep vein thrombosis diagnosed during index hospitalization and within 30 d of surgery

Cardiac events
• Intra- or postoperative acute myocardial infarction diagnosed during index hospitalization and within 30 d of surgery

Global measures
• Mortality within 30 d of surgery
• Readmission within 30 d of surgery

Data from The Joint Commission, 2012.

The SCIP effort provides an infrastructure and guidelines for data collection and quality improvement on a national scale. By achieving high levels of compliance with evidence-based practices to reduce SSIs, venous thromboembolism events, and perioperative cardiac complications, the potential number of lives saved in the Medicare patient population alone exceeds 13,000 annually.

National Surgical Quality Improvement Program

The National Surgical Quality Improvement Program (NSQIP) is a measurement program that allows hospitals to sample their rates of postoperative events and compare them to similar hospitals. Created by the Veterans Health Administration (VA) in 1991, NSQIP has been credited with measuring and improving morbidity and mortality outcomes at the VA, reducing 30-day mortality rate after major surgery by 31%, and 30-day postoperative morbidity by 45% in its first decade. Beta testing at 18 non-VA sites from 2001 to 2004 demonstrated the feasibility and utility of the program in the private sector. The program was subsequently expanded to the private sector in 2004 when the American College of Surgeons endorsed the program and encouraged hospital participation to measure and evaluate outcomes on a large scale. A study of 118 hospitals participating in NSQIP between 2005 to 2007 showed that 82% of hospitals decreased their complication rates and there was a decrease in morbidity of 11% and mortality of 7% annually per hospital. Currently, over 400 private-sector U.S. hospitals participate in the program.

NSQIP uses a risk-adjusted ratio of the observed to expected outcome (focusing primarily on 30-day morbidity and mortality) to compare the performance of participating hospitals with their peers. The data the program has compiled also can be used to conduct observational studies using prospectively collected information on more than 1.5 million patients and operations. The expansion of NSQIP to the private sector has helped shift the focus from merely preventing the provider errors and sentinel events highlighted by the IOM publication “To Err Is Human” to the larger goal of preventing all adverse postoperative outcomes.

Several insights about patient safety have arisen as a result of NSQIP. First, safety is indistinguishable from overall quality of surgical care and should not be addressed separately. Defining quality in terms of keeping a patient safe from adverse outcomes allows the NSQIP data to be used to assess and improve quality of care by making improvements in patient safety. In other words, prevention of errors is synonymous with the reduction of adverse outcomes and can be used as a reliable quality measure. Second, during an episode of surgical care, adverse
outcomes, and hence, patient safety, are primarily determined by the quality of the systems of care. Errors in hospitals with higher than expected observed to expected outcomes ratios are more likely to be from system errors than from provider incompetence. This underscores the importance of adequate communication, coordination, and teamwork in achieving quality surgical care. Finally, reliable comparative outcomes data are imperative for the identification of system problems. Risk-adjusted rates of adverse outcomes must be compared with those at peer institutions to appreciate more subtle system errors that lead to adverse outcomes to prompt changes in the quality of an institution’s processes and structures.

The Leapfrog Group

One of the largest efforts to standardize evidence-based medicine in the United States is led by The Leapfrog Group, an alliance of large public and private healthcare purchasers representing more than 37 million individuals across the United States. This healthcare consortium was founded in 2000 with the aim to exert their combined leverage toward improving nationwide standards of healthcare quality, optimizing patient outcomes, and ultimately lowering healthcare costs. The Leapfrog Group’s strategy to achieve these goals is through providing patient referral, financial incentives, and public recognition for hospitals that practice or implement evidence-based healthcare standards.

The healthcare quality and safety practices (leaps) that Leapfrog initially identified to measure healthcare standards were hospital use of computerized physician order entry systems, 24-hour ICU physician staffing, and evidence-based hospital referral (EBHR) standards for five high-risk operations. In 2010, after the National Quality Forum (NQF) released its updated Safe Practices for Better Healthcare, Leapfrog added a safe practices leap, which includes eight practices from the NQF report.

Leapfrog collects data on these practices through administration of an ongoing, voluntary, web-based hospital quality and safety survey. This survey is conducted in 41 regions that cover over half of the U.S. population and 62% of all hospital beds in the country. In 2011, more than 1200 urban, suburban, and rural hospitals participated in the survey. Leapfrog asks for information on eight high-risk conditions or procedures, including coronary artery bypass graft, percutaneous coronary intervention, abdominal aortic aneurysm (AAA) repair, pancreatic resection, and esophagectomy. These procedures were chosen because evidence exists that adherence to certain process measures can dramatically improve the outcomes of these procedures. In addition, more than 100 studies also have demonstrated that better results are obtained at high-volume hospitals when undergoing cardiovascular surgery, major cancer resections, and other high-risk procedures. Hospitals fulfilling the EBHR Safety Standard are expected to meet the hospital and surgeon volume criteria shown in Table 12-6. Hospitals that do not meet these criteria but adhere to the Leapfrog-endorsed process measures for coronary artery bypass graft surgery, percutaneous coronary intervention, AAA repair, and care for high-risk neonates, receive partial credit toward fulfilling the EBHR Safety Standard. Leapfrog purchasers work to recognize and reward hospitals that provide care for their enrollees who meet EBHR standards.

In a recent study, Brooke and associates analyzed whether achieving Leapfrog’s established evidence-based standards for AAA repair, including meeting targets for case volume and perioperative β-blocker usage, correlated with improved patient outcomes over time. After controlling for differences in hospital and patient characteristics, hospitals that implemented a policy for perioperative β-blocker usage had an estimated 51% reduction in mortality following open AAA repair cases. Among 111 California hospitals in which endovascular AAA repair was performed, in-hospital mortality was reduced by an estimated 61% over time among hospitals meeting Leapfrog case volume standards, although this result was not statistically significant. These results suggest that hospital compliance with Leapfrog standards for elective AAA repair is an effective means to help improve in-hospital mortality outcomes over time and support further efforts aimed at standardizing patient referral to hospitals that comply with evidence-based medicine standards for other surgical procedures.

The newest effort of the Leapfrog group is to promote transparency of hospital outcomes using a safety scorecard. This information can be viewed at www.hospitalsafetygrade.org.

World Health Organization “Safe Surgery Saves Lives” Initiative

In October 2004, the WHO launched a global initiative to strengthen healthcare safety and monitoring systems by creating the World Alliance for Patient Safety. As part of the group’s efforts to improve patient safety, the alliance implemented a series of safety campaigns that brought together experts in specific problem areas through individual Global Patient Safety Challenges. The second Global Patient Safety Challenge focuses on improving the safety of surgical care. The main goal of the campaign, called Safe Surgery Saves Lives, is to reduce surgical deaths and complications through the universal adaptation of a comprehensive perioperative surgical safety checklist in ORs worldwide (Fig. 12-4). In addition to the checklist, the WHO defined a set of uniform measures for national and international surveillance of surgical care to better assess the quantity and quality of surgical care being delivered worldwide. At the population level, metrics include the number of surgeon, anesthesia, and nurse providers per capita, the number of ORs per capita, and overall surgical case volumes and mortality rates. At the hospital level, metrics include safety improvement structures and a surgical “Apgar score,” a validated method of prognosticating patient outcomes based on intraoperative events (i.e., hypotension, tachycardia, blood loss).

National Quality Forum

The National Quality Forum (NQF) is a coalition of healthcare organizations that has worked to develop and implement a national strategy for healthcare quality measurement and
reporting. Their mission is to improve the quality of American healthcare by setting national priorities and goals for performance improvement, endorsing national consensus standards for measuring and publicly reporting on performance, and promoting the attainment of national goals through education and outreach programs.

One of the major contributions of the NQF is the development of a list of Serious Reportable Events, which are frequently referred to as “never events.” According to the NQF, “never events” are errors in medical care that are clearly identifiable, preventable, and serious in their consequences for patients and that indicate a real problem in the safety and credibility of a healthcare facility. Examples of never events include surgery performed on the wrong body part; a foreign body left in a healthcare facility. Despite widespread agreement that never events are preventable and many never events to help elucidate the root cause.

"NEVER EVENTS" IN SURGERY

Never events are errors in medical care that are clearly identifiable, preventable, and serious in their consequences for patients and that indicate a real problem in the safety and credibility of a healthcare facility. Despite widespread agreement that surgical never events are preventable and despite several national and local programs being launched to decrease them, never events are still a significant problem. A study from Mehtsun and colleagues showed that from October 1990 to October 2010, nationwide there were 9744 paid malpractice claims for never events. Of these, mortality was reported in 6.6%, permanent injury in 33.8%, and temporary injury in 49.7%. The cost of the never events totaled $1.3 billion. Also, of physicians who were named in a surgical never event claim, 12.4% were named in a future never events claim. Another study in 2010 by The Joint Commission found that wrong-site surgery occurs 40 times per week nationwide. Future directions for decreasing these problems include public reporting of never events by hospitals to increase hospital accountability, more formal training in teamwork, and CUSP programs in hospitals that have higher rates of never events to help elucidate the root cause.

Retained Surgical Items

A retained surgical item refers to any surgical item found to be inside a patient after he or she has left the OR, thus requiring a second operation to remove the item. Estimates of retained foreign bodies in surgical procedures range from one case per 8000 to 18,000 operations, corresponding to one case or more each year for a typical large hospital or approximately 1500 cases per year in the United States. This estimate is based on an analysis of malpractice claims and is likely to underestimate the true incidence. The risk of having a retained surgical item increases during emergency surgery, when there are unplanned changes in
sensation of a mass of fullness after a surgical procedure that leads to the discovery of a metallic object on a radiographic study. Commonly retained instruments include the malleable and "FISH" instrument that are used to protect the viscera when closing abdominal surgery.

A retained surgical foreign body should be included in the differential diagnosis of any postoperative patient who presents with pain, infection, a palpable mass, or a radiopaque structure on imaging. The diagnosis can usually be made using a computed tomographic (CT) scan, and this is often the only test needed. If a retained surgical item is identified in the setting of an acute clinical presentation, the treatment usually is removal of the item. However, if the attempt to remove the retained surgical item can potentially cause more harm than the item itself, as in the case of a needle or a small part of a surgical item, then removal is occasionally not recommended. Retained surgical sponges should always be removed.

The American College of Surgeons and the Association of Perioperative Registered Nurses, in addition to The Joint Commission, have issued guidelines to try to prevent the occurrence of retained surgical items. Current recommendations include the use of standard counting procedures, performing a thorough wound exploration before closing a surgical site, and using only X-ray–detectable items in the surgical wound. These organizations also strongly endorse the completion of a postoperative debriefing after every operation. An X-ray at the completion of an operation is encouraged if there is any concern for a foreign body based on confusion regarding the counts by even a single member of the OR team or in the presence of a risk factor.

### Surgical Counts

The benefit of performing surgical counts to prevent the occurrence of retained surgical items is controversial. The increased risk of a retained surgical item during emergency surgery in the study by Gawande and colleagues appeared to be related to bypassing the surgical count in many of these cases. However, in another study, the “falsely correct count,” in which a count is performed and declared correct when it is actually incorrect, occurred in 21% to 100% of cases in which a retained surgical item was found. This type of count was the most common circumstance encountered in all retained surgical item cases, which suggests that performing a surgical count in and of itself does not prevent this error from taking place. The counting protocol also imposes significant demands on the nursing staff and

---

**Table 12-9**

<table>
<thead>
<tr>
<th>Risk factors for retained surgical sponges</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Emergency surgery</td>
</tr>
<tr>
<td>• Unplanned changes in procedure</td>
</tr>
<tr>
<td>• Patient with higher body mass index</td>
</tr>
<tr>
<td>• Multiple surgeons involved in same operation</td>
</tr>
<tr>
<td>• Multiple procedures performed on same patient</td>
</tr>
<tr>
<td>• Involvement of multiple operating room nurses/staff members</td>
</tr>
<tr>
<td>• Case duration covers multiple nursing “shifts”</td>
</tr>
</tbody>
</table>
distracts them from focusing on other primarily patient-centered tasks, often during critical portions of the case.19

A retained surgical item can occur even in the presence of a known incorrect count. This event is usually a result of poor communication in which a surgeon will dismiss the incorrect count and/or fail to obtain a radiograph before the patient leaves the OR. Having stronger institutional policies in place in case of an incorrect count (such as requiring a mandatory radiograph while the patient is still in the OR) can avoid conflict among caregivers and mitigate the likelihood of a retained surgical item occurring as a result of a known incorrect count.

Although there is no single tool to prevent all errors, the development of multiple lines of defense to prevent retained surgical items and universally standardizing and adhering to OR safety protocols by all members of the surgical team will help reduce the incidence of this never event.46 Surgeons should take the lead in the prevention of retained surgical items by avoiding the use of small or nonradiologically detectable sponges in large cavities, performing a thorough wound inspection before closing any surgical incision, and having a vested interest in the counting procedure performed by nursing staff. The value of routine radiography to prevent a retained surgical item in emergency cases or when major procedures involving multiple surgical teams are being performed is becoming more apparent.

The widely accepted legal doctrine when a foreign body is erroneously left in a patient is that the mere presence of the item in the plaintiff’s body indicates that the patient did not receive proper surgical care. The characteristics of the surgeon and their style, bedside manner, honesty, and confidence demonstrated in the management of the case can go a long way in averting a lawsuit or mitigating damages.

Wrong-Site Surgery
Wrong-site surgery is any surgical procedure performed on the wrong patient, wrong body part, wrong side of the body, or wrong level of a correctly identified anatomic site. It is difficult to determine the true incidence of wrong-site surgery for several reasons. First, there is no standard definition for what constitutes wrong-site surgery among various healthcare organizations. Another factor is that wrong-site surgery is underreported by healthcare providers. Finally, the total number of potential opportunities for each type of wrong-site error is unknown. However, various studies show incidences ranging from 1 in 112,994 cases to 1 in 15,500 cases.46 The Washington University School of Medicine suggests a rate of 1 in 17,000 operations, which adds up to approximately 4000 wrong-site surgeries in the United States each year. If these numbers are correct, wrong-site surgery is the third most frequent life-threatening medical error in the United States.47

Several states now require mandatory reporting of all wrong-site surgery events, including near misses. These data provide some insight into the number of actual errors compared to the number of potential opportunities to perform wrong-site surgery. Of the 427 reports of wrong-site surgery submitted from June 2004 through December 2006 to the Pennsylvania Patient Safety Reporting System, more than 40% of the errors actually reached the patient, and nearly 20% involved completion of a wrong-site procedure.46

The risk of performing wrong-site surgery increases when there are multiple surgeons involved in the same operation or multiple procedures are performed on the same patient, especially if the procedures are scheduled or performed on different areas of the body.47 Time pressure, emergency surgery, abnormal patient anatomy, and morbid obesity are also thought to be risk factors. Communication errors are the root cause in more than 70% of the wrong-site surgeries reported to The Joint Commission.46 Other risk factors include receiving an incomplete preoperative assessment; having inadequate procedures in place to verify the correct surgical site; or having an organizational culture that lacks teamwork or reveres the surgeon as someone whose judgment should never be questioned.

There is a one in four chance that surgeons who work on symmetric anatomic structures will be involved in a wrong-site error sometime during their careers.47 The specialties most commonly involved in reporting wrong-site surgeries according to The Joint Commission are orthopedic/pediatric surgery (41%); general surgery (20%); neurosurgery (14%); urology (11%); and maxillofacial, cardiovascular, otolaryngology, and ophthalmology (14%).46 Most errors involved symmetric anatomic structures: lower extremities (30%), head/neck (24%), and genital/urinary/pelvic/groin (21%).42 Although orthopedic surgery is the most frequently involved, this may be due to the higher volume of cases performed as well as the increased opportunity for lateralization errors inherent in the specialty. In addition, because the American Academy of Orthopaedic Surgeons has historically tried as a professional organization to reduce wrong-site operations, orthopedic surgeons may be more likely to report these events when they do occur.47

The Joint Commission Universal Protocol to Ensure Correct Surgery
The movement to eliminate wrong-site surgery began among professional orthopedic societies in the mid-1990s, when both the Canadian Orthopaedic Association and the American Academy of Orthopaedic Surgeons issued position statements and embarked on educational campaigns to prevent the occurrence of wrong-site surgery within their specialty.47 Other organizations that issued position statements advocating for the elimination of wrong-site surgery include the North American Spine Society, the American Academy of Ophthalmology, the Association of Perioperative Registered Nurses, and the American College of Surgeons. After issuing a review of wrong-site surgery in their Sentinel Event Alert in 1998, The Joint Commission made the elimination of wrong-site surgery one of their first National Patient Safety Goals in 2003 and adopted a universal protocol for preventing wrong-site, wrong-procedure, and wrong-person surgery in 2004. The protocol has been endorsed by more than 50 professional associations and organizations.

A preoperative “time-out” or “pause for the cause” to confirm the patient, procedure, and site to be operated on before incision was recommended by The Joint Commission and is now mandatory for all ORs in the United States. Elements of the protocol include the following:

- Verifying the patient’s identity
- Marking the surgical site
- Using a preoperative site verification process such as a checklist
- Confirming the availability of appropriate documents and studies before the start of a procedure
- Taking a brief time-out immediately before skin incision, in which all members of the surgical team actively communicate and provide oral verification of the patient’s identity, surgical site, surgical procedure, administration of preoperative
medications, and presence of appropriate medical records, imaging studies, and equipment
• Monitoring compliance with protocol recommendations

Focusing on individual process components of the universal protocol, such as surgical site marking or the time-out, is not enough to prevent wrong-site surgery. Over a 30-month period in Pennsylvania, 21 wrong-site errors occurred despite the proper use of time-out procedures, with 12 of these errors resulting in complete wrong-site procedures. During the same period, correct site markings failed to prevent another 16 wrong-site surgeries, of which six were not recognized until after the procedure had been completed.47

Site verification begins with the initial patient encounter by the surgeon, continues throughout the preoperative verification process and during multiple critical points in the OR, and requires the active participation of the entire operating team, especially the surgeon and anesthesia provider. Based on a recent review of malpractice claims, two-thirds of wrong-site operations could have been prevented by a site-verification protocol.48

Despite the proliferation of wrong-site protocols in the last decade, their effectiveness is difficult to measure as the incidence of wrong-site surgery is too rare to measure as a rate. Interestingly, the number of sentinel events reported to The Joint Commission has not changed significantly since the widespread implementation of the Universal Protocol in 2004.47 This could be due to an increase in reporting rather than an actual increase in the incidence of wrong-site surgery.

The legal treatment of wrong-site surgery is similar to that of surgical items erroneously left in a patient: the mere fact that it occurred indicates that the patient did not receive proper surgical care. A malpractice claim may lead to a settlement or award on verdict in the six- or seven-figure range in 2011 U.S. dollars.41

Ultimately, the occurrence of retained surgical items or wrong-site surgery is a reflection of the quality of professional communication between caregivers and the degree of teamwork among the members of the operating team. In addition to standardizing procedures like the surgical count, instituting mandatory postoperative radiographs in the presence of a known miscount, and reforming the processes of patient identification and site verification, organizations should also strive to create a culture of safety, create independent and redundant checks for key processes, and create a system in which caregivers can learn from their mistakes (Table 12-10).49

Table 12-10

<table>
<thead>
<tr>
<th>Best practices for operating room safety</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Conduct The Joint Commission Universal Protocol (“time-out”) to prevent wrong-site surgery.</td>
</tr>
<tr>
<td>• Perform an operating room briefing (checklist) to identify and mitigate hazards early.</td>
</tr>
<tr>
<td>• Promote a culture of speaking up about safety concerns.</td>
</tr>
<tr>
<td>• Use a screening X-ray to detect foreign bodies in high-risk cases.</td>
</tr>
<tr>
<td>• Begin patient sign-outs with the most likely immediate safety hazard.</td>
</tr>
</tbody>
</table>


TRANSPARENCY IN HEALTHCARE

Despite a large increase in data being collected about patient safety and harm, much of it is not available to the public or other hospitals. This lack of transparency allows some hospitals to continue to practice outdated medicine and, in some cases, puts patients at a higher risk of serious complications. In a study by Mark Chassin, the health commissioner of New York State, having hospitals publicly disclose their mortality rates for coronary artery bypass graft (CABG) procedures resulted in a 41% decline in mortality from CABGs statewide.50 In this study, when CABG mortality data were initially made public, there was a wide range in cardiac surgery-related mortality from 1% to 18%, depending on the hospital; the standard of care is 2%. The reasons for higher mortality in the poorly performing hospitals ranged from poor communication between care teams to one rogue surgeon operating when the surgeon should not have been. The consequence of making this data transparent was that the hospitals held multidisciplinary, CUSP-like meetings, where as a team they decided on the measures to implement for improvement. Through this, over the next year, most hospitals decreased their mortality rate to below 2%. Even the hospital that had an 18% mortality rate decreased it to 7% within 3 years and 1.7% over the next several years.

Transparency in healthcare is becoming central to the healthcare quality discussion. A new SCIP core measure is publishing practitioner performance, and all Leapfrog survey results are published online where other hospitals and the public can see them. Additionally, different large medical societies, including the Society for Thoracic Surgery (STS), are encouraging and rewarding practitioners and hospitals that are transparent with their outcomes. Making hospital outcomes transparent makes hospitals accountable to the public for their outcomes and, in the case of New York, caused a radical improvement in the quality of care provided to patients. It also empowers patients by making them better informed about which hospital they choose for their care, which will further incentivize hospitals to improve.

Public Reporting and Patient Assessment of Care
The epiphany moment in contemporary healthcare created by the Institute of Medicine report2 generated far-reaching effects. One important aspect has been development of a variety of initiatives focused on the generation, endorsement, and reporting of numerous measures related to the safety and quality of healthcare—primarily process and outcomes measures. However, the science of measure development is slow paced and, unfortunately, has difficulty evolving at the same pace of change as clinical medicine or healthcare delivery systems.

Given the strong interest for improved knowledge and information by consumers of healthcare, the trend toward public reporting has rapidly gained momentum and outpaced reporting from the measurement science community. This has subsequently created occasional confusion and uncertainty in the marketplace—simply because the generation of public reports are not necessarily always based upon solid scientific data or evidence. The resulting net effect can be creation of a premature focus by organizations and providers on achieving success within influential public reporting venues (e.g., U.S. News Best Hospitals) and uncertainty by patients on what are optimal healthcare information resources.
Ideally centered on the public good, federal government sponsored healthcare payment plans are also focused upon measurement and reporting within the industry. One such initiative funded and overseen by the Agency for Healthcare Research and Quality (AHRQ) is the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) program. AHRQ works closely with a consortium of public and private research organizations to develop and maintain the HCAHPS surveys, but they do not administer any of the surveys to patients or require use of the surveys. The intent of the HCAHPS initiative is to provide a standardized survey instrument and data collection methodology for measuring patients’ perspectives on hospital care (Fig. 12-7).

While many hospitals have collected information on patient satisfaction, prior to HCAHPS there was no national standard for collecting or publicly reporting patients’ perspectives of care information that would enable valid comparisons to be made across all hospitals. Three broad goals have shaped the HCAHPS Survey. First, the survey is designed to produce comparable data on the patient’s perspective on care that allows objective and meaningful comparisons between hospitals on domains that are important to consumers. Second, public reporting of the survey results is designed to create incentives for hospitals to improve their quality of care. Hospitals frequently distribute the results of HCAHPS surveys of individual services or physicians to incentivize corrective steps and improve patients’ perceptions of their care. Third, public reporting will serve to enhance public accountability in healthcare by increasing the transparency of the quality of hospital care provided in return for the public investment (www.hcahpsonline.org).

HCAHPS scores are now directly tied to a hospital’s CMS reimbursement by federal law, and survey results account for 25% of the value-based purchasing score—directly impacting a hospital’s Medicare payments. In May 2005, the National Quality Forum (NQF), an organization established to standardize healthcare quality measurement and reporting, formally endorsed the HCAHPS Hospital Survey. The NQF endorsement represents the consensus of many healthcare providers, consumer groups, professional associations, purchasers, federal agencies, and research and quality organizations (www.qualityforum.org).

While the American College of Surgeons and other professional societies continue to develop and implement scientifically based healthcare measures, the escalating prevalence of less robust measures developed outside of scientific communities will also continue to expand. The patient population, provider organizations, payer organization, and the surgical community will necessarily need to find the balance of developing and utilizing valid patient-based information for decision-making.

RISK MANAGEMENT

Between one-half to two-thirds of hospital-wide adverse events are attributable to surgical care. Most surgical errors occur in the OR and are technical in nature. Surgical complications and adverse outcomes have previously been linked to lack of surgeon specialization, low hospital volume, communication breakdowns, fatigue, surgical residents and trainees, and numerous other factors.

However, poor surgical outcomes are not necessarily correlated with a surgeon’s level of experience in performing a certain procedure. In one study, three-fourths of the technical errors that occurred in a review of malpractice claims data involved fully trained and experienced surgeons operating within their area of expertise, and 84% occurred in routine operations that do not require advanced training. Rather than surgeon expertise, these errors likely occurred due to situations complicated by patient comorbidity, complex anatomy, repeat surgery, or equipment problems (Table 12-11). Because these errors occurred during routine operations, previous suggestions to limit the performance of high-complexity operations using selective referral, regionalization, or limitation of privileging may not actually be effective in reducing the incidence of technical error among surgical patients.

In any event, although there has been much emphasis on reducing the prevalence of surgical technical errors as a way of improving surgical care, a technical error in the OR may not be the most important indicator of whether a surgeon will be sued by a patient. Recent studies point to the importance of a surgeon’s communication skills in averting malpractice litigation. In the American College of Surgeons’ Closed Claims Study, although intraoperative organ injuries occurred in 40% of patients, a surgical technical misadventure was the most deficient component of care in only 12% of patients. In fact, communication and practice pattern violations were the most common deficiency in care for one third of patients in the Closed Claims Study who received the expected standard of surgical care.

The Importance of Communication in Managing Risk

The manner and tone in which a physician communicates is potentially more important to avoiding a malpractice claim than the actual content of the dialogue. For example, a physician relating to a patient in a “negative” manner may trigger litigious feelings when there is a bad result, whereas a physician relating in a “positive” manner may not. Expressions of dominance, in which the voice tone is deep, loud, moderately fast, unaccented, and clearly articulated, may communicate a lack of empathy and understanding for the patient, whereas concern or anxiety in the surgeon’s voice is often positively related to expressing concern and empathy. General and orthopedic surgeons whose tone of voice was judged to be more dominant were more likely to have been sued than those who sounded less dominant.

When significant medical errors do occur, physicians have an ethical and professional responsibility to immediately disclose them to patients. Failure to disclose errors to patients undermines public confidence in medicine and can create legal liability related to fraud. Physicians’ fear of litigation represents a major barrier to error disclosure. However, when handled appropriately, immediate disclosure of errors frequently leads to improved patient rapport, improved satisfaction, and fewer malpractice claims. In fact, rapport is the most important factor in determining whether a lawsuit is filed against a physician.

In 1987, the Department of Veterans Affairs Hospital in Lexington, Kentucky, implemented the nation’s first formal apology and medical error full disclosure program, which called for the hospital and its doctors to work with patients and their families to settle a case. As a result, the hospital improved from having one of the highest malpractice claims totals in the VA system to being ranked among the lowest quartile of a comparative group of similar hospitals for settlement and litigation costs over a 7-year period. Its average payout in 2005 was $16,000 per settlement vs the national VA average of $98,000 per settlement, and only two lawsuits went to trial during a
YOUR CARE FROM DOCTORS

5. During this hospital stay, how often did doctors treat you with courtesy and respect?
   1. Never
   2. Sometimes
   3. Usually
   4. Always

6. During this hospital stay, how often did doctors listen carefully to you?
   1. Never
   2. Sometimes
   3. Usually
   4. Always

7. During this hospital stay, how often did doctors explain things in a way you could understand?
   1. Never
   2. Sometimes
   3. Usually
   4. Always

THE HOSPITAL ENVIRONMENT

8. During this hospital stay, how often were your room and bathroom kept clean?
   1. Never
   2. Sometimes
   3. Usually
   4. Always

9. During this hospital stay, how often was the area around your room quiet at night?
   1. Never
   2. Sometimes
   3. Usually
   4. Always

YOUR EXPERIENCES IN THIS HOSPITAL

10. During this hospital stay, did you need help from nurses or other hospital staff in getting to the bathroom or in using a bedpan?
    1. Yes
    2. No ➔ If No, Go to Question 12

11. How often did you get help in getting to the bathroom or in using a bedpan as soon as you wanted?
    1. Never
    2. Sometimes
    3. Usually
    4. Always

12. During this hospital stay, did you have any pain?
    1. Yes
    2. No ➔ If No, Go to Question 15

13. During this hospital stay, did hospital staff talk with you about how much pain you had?
    1. Never
    2. Sometimes
    3. Usually
    4. Always

14. During this hospital stay, did hospital staff talk with you about how to treat your pain?
    1. Never
    2. Sometimes
    3. Usually
    4. Always

Table 12-11

Common causes of lawsuits in surgery

- Positional nerve injury
- Common bile duct injury
- Failure to diagnose or delayed diagnosis
- Failure to treat, delayed treatment, or wrong treatment
- Inadequate documentation
- Inappropriate surgical indication
- Failure to call a specialist
- Cases resulting in amputation/limb loss

10-year period. As a result of the success of this program, the Department of Veteran Affairs expanded the program to all VA hospitals nationwide in October 2005. This model also was replicated at the University of Michigan Health System with similar results. Its full-disclosure program cut the number of pending lawsuits by one half and reduced litigation costs per case from $65,000 to $35,000, saving the hospital approximately $2 million in defense litigation bills each year. In addition, University of Michigan’s doctors, patients, and lawyers are happier with this system. The cultural shift toward honesty and openness also has led to the improvement of systems and processes to reduce medical errors, especially repeat medical errors.
With regard to risk management, the importance of good communication by surgeons and other care providers cannot be overemphasized. Whether alerting other members of the care team about a patient’s needs, openly discussing concerns the patient and/or family might have, or disclosing the cause of a medical error, open communication with all parties involved can reduce anger and mistrust of the medical system; the frequency, morbidity, and mortality of preventable adverse events; and the likelihood of litigation.

**COMPLICATIONS**

Despite the increased focus on improving patient safety and minimizing medical errors, it is impossible to eliminate human error entirely. Individual errors in judgment or technique can cause minor or major complications during or after a surgical procedure. Although these types of errors may not be quantified as easily as wrong-site surgery or a retained surgical item, they can still lead to surgical complications that prolong the course of illness, lengthen hospital stay, and increase morbidity and mortality rates. In addition to technical and management errors, patient comorbidities also increase the risk of complications. The recognition and management of complications is a critical component of surgical care.

**Robotic Surgery**

Surgical advancements would not exist without intellectual curiosity, innovation, and technical developments; robotic surgery is a prime example of such an advance. With these advancements, however, errors and complications appear to be an inevitable and recognized risk by institutions and stakeholders due to unforeseen risks inherent in the new technology and the failure or delay of achieving expertise with a new device or technology. Although the reward for adopting new advances may be notoriety, increased patient referrals, improved patient satisfaction, decreased pain, and possibly decreased length of stay, the risks of adopting new technologies and methods become apparent only after widespread use.

Multiple surgical specialties have begun or continue to develop their experiences using robotic surgery from general surgery procedures such as inguinal hernia repairs to pancreatecduodenectomies to complex thoracic, urologic and ear, nose, and throat procedures. When robotic surgery goes awry, however, the complications can be serious. The MAUDE (Manufacturer and User Facility Device Experience) is an open access database where mandatory and voluntary adverse events are collected. As it relates to robotic surgery, some important information has been elucidated. Device failures (electrocautery, instrument malfunctions), make up roughly half of the complications. A retrospective study over the past 14 years in the United States documented over 10,000 robotic device-related complications that have occurred, of which 98% were reported by the manufacturers and distributors, and 2% were voluntary reports by hospitals and physicians. The data revealed that 1535 adverse events (14.4%) led to significant negative patient experiences (1391 injuries and 144 deaths). Additionally, the absolute number of reports increased 32 times since 2006, while in the same time period, the number of cases performed has only increased tenfold. Despite the large number of reports contained within this database, the extent to which it is a true representation of the complications associated with robotic surgery is uncertain due to the lack of comprehensive and mandatory reporting.

Despite the numbers and trends reported from this database, few prospective, controlled trials exist that examine the risks and benefits of robotic surgery with those of open and laparoscopic surgery. The data from 5 to 10 years ago may also be misleading as the approved use of robotic surgery continues to expand to additional specialties. The more recent adoption of robotic surgery by specialties such as gynecologic surgery, for example, appears to be accompanied by a disproportionately high rate of morbidity and mortality in robotically assisted procedures. Teaching institutions are producing a newer generation of robotic surgeons who will continue important advances in surgery and identify those patients who would benefit most from this type of approach. The challenge for the surgical community is to develop robust and effective training programs to allow trainees and practicing surgeons to acquire the skills necessary to perform robotic procedures with the highest degree of safety. This need replicates the development of skill acquisition processes that reversed the high number of bile duct injuries after the introduction of laparoscopic cholecystectomy and suggests that validated curricula and the use of robotic simulation applications will be crucial to achieve these goals (see Chapter 53, Skills and Simulation).

**Complications in Minor Procedures**

When performing procedures such as central line insertion or arterial line insertion, one should consider the necessity of the access, the use of less invasive or lower risk alternatives such as PICC line insertion instead of central line insertion, and non-invasive cardiac monitoring instead of arterial line insertion. While these alternatives may not be reliable substitutes in all patients, considering less invasive procedures can reduce the problem of avoidable harm.

**Central Venous Access Catheters.** Complications of central venous access catheters are common. Improvements in ultrasound technology and mass education surrounding the use and techniques in ultrasonography have led to increased employment and enthusiasm for its use in central venous catheter placement. Numerous institutions have mandated the use of ultrasound for placement of all central venous lines. In addition, many subclavian catheters have been alternatively placed at the internal jugular position due to a perceived benefit of decreasing the complication of pneumothorax. This theoretical benefit may be offset by an increase in line infections as the neck is a difficult site to keep clean and the dressing intact. Steps to decrease complications include:

- Ensure that central venous access is indicated.
- Experienced personnel should insert the catheter or should supervise the insertion.
- Use proper positioning and sterile technique.
- Ultrasound is recommended for internal jugular vein insertion.
- All central venous catheters should be assessed on a daily basis and should be exchanged only for specific indications (not as a matter of routine).
- All central catheters should be removed as soon as possible.

Common complications of central venous access include the following.

**Pneumothorax** Occurrence rates from both subclavian and internal jugular vein approaches are 1% to 6%. Prevention requires proper positioning of the patient and correct insertion technique. A postprocedure chest X-ray is recommended to confirm the presence or absence of a pneumothorax, regardless
of whether a pneumothorax is suspected. Recent reports have questioned whether a chest X-ray is required when the line is placed and confirmed under ultrasound guidance. Pneumothorax rates are higher among inexperienced providers and under-weight patients but occur with experienced operators as well. If the patient is stable, and the pneumothorax is small (<15%), close expectant observation may be adequate. If the patient is symptomatic, a thoracostomy tube should be placed. Occasionally, pneumothorax will occur as late as 48 to 72 hours after central venous access attempts. This usually creates sufficient compromise that a tube thoracostomy is required.

**Arrhythmias** Arrhythmias can result from myocardial irritability secondary to guidewire placement and usually resolve when the catheter or guidewire is withdrawn from the right heart. Prevention requires electrocardiogram (ECG) monitoring whenever possible during catheter insertion and rapid recognition when a new arrhythmia occurs.

**Arterial Puncture** Inadvertent puncture or laceration of an adjacent artery with bleeding can occur, but the majority will resolve with direct pressure on or near the arterial injury site. Rarely will angiography, stent placement, or surgery be required to repair the puncture site, but close observation and a chest X-ray are indicated. Ultrasound-guided insertion has not mitigated this complication, but it may decrease the incidence of arterial puncture. Ultrasound use has also been shown to decrease the number of attempts and the time it takes to complete insertion.

**Lost Guidewire** A guidewire or catheter that inadvertently migrates further into the vascular space away from the insertion site can be readily retrieved with interventional angiography techniques. A prompt chest X-ray and close monitoring of the patient until retrieval are indicated.

**Air Embolus** Although estimated to occur in only 0.2% to 1% of patients, an air embolism can be dramatic and fatal. If an embolus is suspected, the patient should immediately be placed into a left lateral decubitus Trendelenburg position so the entrapped air can be stabilized within the right ventricle. Auscultation over the precordium may reveal a “crunching” sound, but a portable chest X-ray will help confirm the diagnosis. Aspiration via a central venous line accessing the heart may decrease the volume of gas in the right side of the heart and minimize the amount traversing into the pulmonary circulation. Subsequent recovery of intracardiac and intrapulmonary air may require open surgical or angiographic techniques. Treatment may prove futile if the air bolus is larger than 50 mL, however.

**Pulmonary Artery Rupture** Flow-directed, pulmonary artery (Swan-Ganz) catheters can cause pulmonary artery rupture due to excessive advancement of the catheter into the pulmonary circulation. There usually is a sentinel bleed with coughing noted when a pulmonary artery catheter balloon is inflated, followed by uncontrolled hemoptysis. Reinflation of the catheter balloon is the initial step in management, followed by immediate airway intubation with mechanical ventilation, an urgent portable chest X-ray, and notification of the OR that an emergent thoracotomy may be required. If there is no further bleeding after the balloon is reinflated, the X-ray shows no significant consolidation of lung fields from ongoing bleeding, and the patient is easily ventilated, then a conservative nonoperative approach may be considered. However, more typically a pulmonary angiogram with angiembolization or vascular stenting is required. Hemodynamically unstable patients rarely survive because of the time needed to initiate and perform interventional procedures or a thoracotomy and to identify the ruptured branch of the pulmonary artery.

**Central Venous Line Infection** The CDC reports mortality rates of 12% to 25% when a central venous line infection becomes systemic, with a cost of approximately $25,000 per episode.60-62 The CDC does not recommend routine central line changes, but when the clinical suspicion of infection is high, the site of venous access must be changed. Nearly 15% of hospitalized patients will acquire central venous line sepsis. In many instances, once an infection is recognized as central line sepsis, removing the line is adequate. *Staphylococcus aureus* infections, however, present a unique problem because of the potential for metastatic seeding of bacterial emboli. The required treatment is 4 to 6 weeks of tailored antibiotic therapy. Using a checklist when inserting central venous catheters has been shown to significantly decrease rates of line infections.63-65 Following a checklist strategy and close monitoring of catheters has resulted in significant reductions in infection rates for numerous institutions, and many are now reporting zero annual infection rates.

**Arterial Lines.** Arterial lines are placed to facilitate arterial blood gas sampling and hemodynamic monitoring. The use of ultrasound to assist in placement of these catheters has become commonplace and markedly reduces the number of attempts and time for insertion completion.

Arterial access requires a sterile Seldinger technique, and a variety of arteries are used, including the radial, femoral, brachial, axillary, dorsalis pedis, or superficial temporal arteries. Although complications occur less than 1% of the time, they can be catastrophic. Complications include thrombosis, bleeding, hematoma, arterial spasm (nonthrombotic pulselessness), and infection. Thrombosis or embolization of an extremity arterial catheter can result in the loss of a digit, hand, or foot, and the risk is nearly the same for both femoral and radial cannulation. Thrombosis with distal tissue ischemia is treated with anticoagulation, but occasionally surgical intervention is required. Pseudoaneurysms and arteriovenous fistulae can also occur.

**Endoscopy and Bronchoscopy.** The principal risk of gastrointestinal (GI) endoscopy is perforation. Perforations occur in 1 in 10,000 patients with endoscopy alone but have a higher incidence rate when biopsies are performed (up to 10%). This increased risk is due to complications of intubating a GI diverticulum (either esophageal or colonic) or from the presence of weakened or inflamed tissue in the intestinal wall (e.g., diverticulitis, glucocorticoid use, or inflammatory bowel disease).

Patients will usually complain of diffuse abdominal pain shortly after the procedure and then progress with worsening abdominal discomfort and peritonitis on examination. In obtunded or elderly patients, a change in clinical status may be delayed for 24 to 48 hours. Radiologic studies to look for free intraperitoneal air, retroperitoneal air, or a pneumothorax are diagnostic. Open or laparoscopic exploration locates the perforation and allows repair and local decontamination of the surrounding tissues.

The occasional patient who may be a candidate for nonoperative management is one in whom perforation arises during an elective, bowel-prepped endoscopy and who does not have significant pain or clinical signs of infection. These patients must be closely observed in a monitored setting and must be on strict dietary restriction and broad-spectrum antibiotics.

Complications of bronchoscopy include bronchial plugging, hypoxemia, pneumothorax, lobar collapse, and bleeding.
When diagnosed in a timely fashion, they are rarely life-threatening. Bleeding usually resolves spontaneously and rarely requires surgery but may require repeat endoscopy for thermocoagulation or fibrin glue application. The presence of a pneumothorax necessitates placement of a thoracostomy tube when significant deoxygenation occurs or the pulmonary mechanics are compromised. Lobar collapse or mucous plugging usually responds to aggressive pulmonary toilet but occasionally requires repeat bronchoscopy. If biopsies have been performed, the risk for these complications increases.

**Tracheostomy.** Tracheostomy facilitates weaning from a ventilator, may decrease length of ICU or hospital stay, and improves pulmonary toilet. Tracheostomies are performed open, percutaneously, with or without bronchoscopy, and with or without Doppler guidance. The advantages of percutaneous tracheostomy include efficiency and cost containment over open tracheostomy. A recent literature review examining early (<3–7 days) vs late (>14 days) tracheostomy after endotracheal intubation demonstrates little difference in outcomes but does demonstrate greater patient comfort in those patients with tracheostomy than those with an endotracheal tube. Complications and outcomes between the two different methods remain largely equivalent.

Recent studies do not support obtaining a routine chest X-ray after percutaneous or open tracheostomy. However, significant lobar collapse can occur from copious tracheal secretions or mechanical obstruction. The most dramatic complication of tracheostomy is tracheoinnominate artery fistula (TIAF) (Fig. 12-8). This occurs rarely (<0.3%) but carries a 50% to 80% mortality rate. TIAF can occur as early as 2 days or as late as 2 months after tracheostomy. A sentinel bleed occurs in 50% of TIAF cases, followed by a large-volume bleed. Should a TIAF be suspected, the patient should be transported immediately to the OR for fiberoptic evaluation. If needed, remove the tracheostomy and place a finger through the tracheostomy site to apply direct pressure anteriorly for compression of the innominate artery while preparation for a more definitive approach is organized.

**Percutaneous Endogastronomy.** A misplaced percutaneous endogastronomy (PEG) tube may lead to intra-abdominal sepsis with peritonitis and/or an abdominal wall abscess with necrotizing fasciitis. As in other minor procedures, the initial placement technique must be fastidious to avoid complications.

Endoscopic transillumination of the abdomen from within the stomach has been proposed to decrease the risk for error, but this is without supporting evidence. Inadvertent colotomies, intraperitoneal placement of the tube and subsequent leakage of tube feeds with peritonitis, and abdominal wall abscesses require surgery to correct the complications and to replace the PEG with an alternate feeding tube, usually a jejunostomy.

A dislodged or prematurely removed PEG tube should be replaced as early as possible after dislodgment because the gastrostomy site closes rapidly. A contrast X-ray (sinogram) should be performed to confirm the tube’s intragastric position before feeding. If there is uncertainty of the tube location, conversion to an open tube placement procedure is required.

**Tube Thoracostomy.** Chest tube insertion is performed for pneumothorax, hemothorax, pleural effusions, or empyema. In most patients, a chest tube can be easily placed with a combination of local analgesia and light conscious sedation. Common complications include inadequate analgesia or sedation, incomplete penetration of the pleura with formation of a subcutaneous tube track, lacerations to the lung or diaphragm, intraperitoneal placement of the tube through the diaphragm, and bleeding. Additional problems include slippage of the tube out of position or mechanical problems related to the drainage system. In patients with bullous disease, there can be significant intrapleural scarring, and it can be easy to mistakenly place the chest tube into bullae. All of these complications can be avoided with proper initial insertion techniques, plus a daily review of the drainage system and follow-up radiographs. Tube removal can create a residual pneumothorax if the patient does not maintain positive intrapleural pressure by Valsalva maneuver during tube removal and dressing application.

**Complications of Angiography.** Intramural dissection of a cannulated artery can lead to complications such as ischemic stroke from a carotid artery dissection or occlusion, mesenteric ischemia from dissection of the superior mesenteric artery, or a more innocuous finding of “blue toe syndrome” from a dissected artery in a peripheral limb. Invasive or noninvasive imaging studies confirm the suspected problem. The severity of ischemia, extent of dissection determine if anticoagulation therapy or urgent surgical exploration is indicated.

Bleeding from a vascular access site usually is obvious, but may not be visible when the blood loss is tracking into the retroperitoneal tissue planes after femoral artery cannulation. These patients can present with hemorrhagic shock, an abdominopelvic CT scan delineates the extent of bleeding along the retroperitoneum. Initial management is direct compression at the access site and resuscitation as indicated. Urgent surgical exploration may be required to control the bleeding site and evacuate larger hematomas.

Renal complications of angiography occur in 1% to 2% of patients. Contrast nephropathy is a temporary and preventable complication of radiologic studies such as CT, angiography, and/or venography. Intravenous (IV) hydration before and after the procedure is the most efficient method for preventing contrast nephropathy. Nonionic contrast also may be of benefit in higher-risk patients. Close communication between providers is often required to resolve the priorities in care as well as to balance the risks versus benefits of renal protection when managing patients in need of angiographic procedures.

**Complications of Biopsies.** Lymph node biopsies have direct and indirect complications that include bleeding, infection,
lymph leakage, and seromas. Measures to prevent direct complications include proper surgical hemostasis, proper skin preparation, and a single preoperative dose of antibiotic to cover skin flora 30 to 60 minutes before incision. Bleeding at a biopsy site usually can be controlled with direct pressure. Infection at a biopsy site will appear 5 to 10 days postoperatively and may require opening of the wound to drain the infection. Seromas or lymphatic leaks resolve with aspiration of seromas and the application of pressure dressings but may require repeated treatments or even placement of a vacuum drain.

**Organ System Complications**

**Neurologic System.** Neurologic complications that occur after surgery include motor or sensory deficits and mental status changes. Peripheral motor and sensory deficits are often due to neurapraxia secondary to improper positioning and/or padding during operations. Treatment is largely clinical observation, and the majority of deficits resolve spontaneously within 1 to 3 months.

Direct injury to nerves during a surgical intervention is a well-known complication of several specific operations, including superficial parotidectomy (facial nerve), carotid endarterectomy (hypoglossal nerve), thyroidectomy (recurrent laryngeal nerve), prostatectomy (nervi erigentes), inguinal herniorrhaphy (ilioinguinal nerve), and mastectomy (long thoracic and thoracodorsal nerves). The nerve injury may be a stretch injury or an unintentionally severed nerve. In addition to loss of function, severed nerves can result in a painful neuroma that may require subsequent surgery.

Mental status changes in the postoperative patient can have numerous causes (Table 12-12). Mental status changes must be continually assessed. A noncontrast CT scan should be used early to detect new or evolving intracranial causes.

Atherosclerotic disease increases the risk for intraoperative and postoperative stroke (cerebrovascular accident). Postoperatively, hypotension and hypoxemia are the most likely causes of a cerebrovascular accident. Neurologic consultation should be obtained immediately to confirm the diagnosis. Management is largely supportive and includes adequate intravascular volume replacement plus optimal oxygen delivery. Advents in interventional radiology by radiologists and vascular and neurologic surgeons have proven successful alternatives in patients requiring diagnostic and therapeutic care in the immediate and acute postoperative period. Catheter-directed therapy with anticoagulants such as the kinases and tissue plasminogen activator (tPA) has potential benefit in postoperative thrombosis where reoperation carries significant risk. In addition, endoluminal stents with drug-eluting stents (DESs) or non-DESs have been used with some degree of success. DESs do require systemic antplatelet therapy due to the alternative coagulation pathway. Duration of antiplatelet therapy of 1 year is routine.

**Eyes, Ears, and Nose.** Corneal abrasions are unusual, but are due to inadequate protection of the eyes during anesthesia. Overlooked contact lenses in patients occasionally cause conjunctivitis.

Persistent epistaxis can occur after nasogastric tube placement or removal, and nasal packing is the best treatment option if prolonged persistent direct pressure on the external nares fails. Anterior and posterior nasal gauze packing with balloon tamponade, angioembolization, and fibrin glue placement may be required in refractory cases. The use of antibiotics for posterior packing is controversial.

External otitis and otitis media occasionally occur postoperatively. Patients complain of ear pain or decreased hearing, and treatment includes topical antibiotics and nasal decongestion for symptomatic improvement.

Otoxicity due to aminoglycoside administration occurs in up to 10% of patients and is often irreversible. Vancomycin-related otoxicity occurs about 3% of the time when used alone, and as often as 6% when used with other ototoxic agents.68

**Vascular Problems of the Neck.** Complications of carotid endarterectomy include central or regional neurologic deficits or bleeding with an expanding neck hematoma. An acute change in mental status or the presence of localized neurologic deficit requires an immediate return to the OR. An expanding hematoma may warrant emergent airway intubation and subsequent transfer to the OR for control of hemorrhage. Intraoperative anticoagulation with heparin during carotid surgery makes bleeding a postoperative risk. Other complications include arteriovenous fistulae, pseudoaneurysms, and infection, all of which are treated surgically.

---

**Table 12-12**

<table>
<thead>
<tr>
<th>ELECTROLYTE IMBALANCE</th>
<th>TOXINS</th>
<th>TRAUMA</th>
<th>METABOLIC</th>
<th>MEDICATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sodium</td>
<td>Ethanol</td>
<td>Closed head injury</td>
<td>Thyrotoxicosis</td>
<td>Aspirin</td>
</tr>
<tr>
<td>Magnesium</td>
<td>Methanol</td>
<td>Pain</td>
<td>Adrenal insufficiency</td>
<td>β-Blockers</td>
</tr>
<tr>
<td>Calcium</td>
<td>Venoms and poisons</td>
<td>Shock</td>
<td>Hypoxemia</td>
<td>Narcotics</td>
</tr>
<tr>
<td>Inflammation</td>
<td>Ethylene glycol</td>
<td>Psychiatric</td>
<td>Acidosis</td>
<td>Antiemetics</td>
</tr>
<tr>
<td>Sepsis</td>
<td>Carbon monoxide</td>
<td>Dementia</td>
<td>Severe anemia</td>
<td>MAOIs</td>
</tr>
<tr>
<td>AIDS</td>
<td>Depression</td>
<td>Hyperammononemia</td>
<td>TCAs</td>
<td></td>
</tr>
<tr>
<td>Cerebral abscess</td>
<td>ICU psychosis</td>
<td>Poor glycemic control</td>
<td>Amphetamines</td>
<td></td>
</tr>
<tr>
<td>Meningitis</td>
<td>Schizophrenia</td>
<td>Hypothermia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fever/hyperpyrexia</td>
<td>Hyperthermia</td>
<td>Corticosteroids, anabolic steroids</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AIDS = acquired immunodeficiency syndrome; ICU = intensive care unit; MAOI = monoamine oxidase inhibitor; TCA = tricyclic antidepressant.
Intraoperative hypotension during manipulation of the carotid bifurcation can occur and is related to increased tone from baroreceptors that reflexively cause bradycardia. Should hypotension occur when manipulating the carotid bifurcation, an injection of 1% lidocaine solution around this structure should attenuate this reflexive response.

The most common delayed complication following carotid endarterectomy remains myocardial infarction. The possibility of a postoperative myocardial infarction should be considered as a cause of labile blood pressure and arrhythmias in high-risk patients.

**Thyroid and Parathyroid Glands.** Surgery of the thyroid and parathyroid glands can result in hypocalcemia in the immediate postoperative period. Manifestations include ECG changes (shortened P-R interval), muscle spasm (tetany, Chvostek’s sign, and Trouseau’s sign), paresthesias, and laryngospasm. Treatment includes calcium gluconate infusion and, if tetany ensues, chemical paralysis with intubation. Maintenance treatment is thyroid hormone replacement (after thyroidectomy) in addition to calcium carbonate and vitamin D.

Recurrent laryngeal nerve (RLN) injury occurs in less than 5% of patients. Of those with injury, approximately 10% are permanent. Dissection near the inferior thyroid artery is a common area for RLN injury. At the conclusion of the operation, if there is suspicion of an RLN injury, direct laryngoscopy is diagnostic. The cord on the affected side will be in the paramedian position. With bilateral RLN injury, the chance of a successful extubation is poor. If paralysis of the cords is not permanent, function may return 1 to 2 months after injury. Permanent RLN injury can be treated by various techniques to stent the cords in a position of function.

Superior laryngeal nerve injury is less debilitating, as the common symptom is loss of projection of the voice. The glottic aperture is asymmetrical on direct laryngoscopy, and management is limited to clinical observation.

**Respiratory System.** Surgical complications that put the respiratory system in jeopardy are not confined to technical errors. Malnutrition, inadequate pain control, inadequate mechanical ventilation, inadequate pulmonary toilet, and aspiration can cause serious pulmonary problems.

Pneumothorax can occur from central line insertion during anesthesia or from a diaphragmatic injury during an abdominal procedure. Hypotension, hypoxemia, and tracheal deviation away from the affected side may be present. A tension pneumothorax can cause complete cardiovascular collapse. Treatment is by needle thoracostomy, followed by tube thoracostomy. The chest tube is inserted at the fifth intercostal space in the anterior axillary line. The anterior chest wall is up to 1 cm thicker than the lateral chest wall, so needle decompression is more effective in the lateral position. Attempted prehospital needle decompression in the traditional anterior position results in only 50% needle entry into the thoracic cavity.

Hemothoraces should be evacuated completely. Delay in evacuation of a hemothorax leaves the patient at risk for empyema and entrapped lung. If evacuation is incomplete with tube thoracostomy, video-assisted thoracoscopic or open evacuation and pleurodysis may be required.

Pulmonary atelectasis results in a loss of functional residual capacity (FRC) of the lung and can predispose to pneumonia. Poor pain control in the postoperative period contributes to poor inspiratory effort and collapse of the lower lobes in particular. The prevention of atelectasis is facilitated by sitting the patient up as much as possible, early ambulation, and adequate pain control. An increase in FRC by 700 mL or more can be accomplished by sitting patients up to greater than 45°. For mechanically ventilated patients, simply placing the head of the bed at 30° to 45° elevation and delivering adequate tidal volumes (8–10 mL/kg) improves pulmonary outcomes.

Patients with inadequate pulmonary toilet are at increased risk for bronchial plugging and lobar collapse. Patients with copious and tenacious secretions develop these plugs most often, but foreign bodies in the bronchus can be the cause of lobar collapse as well. The diagnosis of bronchial plugging is based on chest X-ray and clinical suspicion with acute pulmonary decompensation with increased work of breathing and hypoxemia. Fiberoptic bronchoscopy can be useful to clear mucous plugs and secretions.

Aspiration complications include pneumonitis and pneumonia. The treatment of pneumonitis is similar to that for acute respiratory distress syndrome (see later in this section) and includes oxygenation with general supportive care. Antibiotics are not indicated. Hospitalized patients who develop aspiration pneumonitis have a mortality rate as high as 70% to 80%. Early, aggressive, and repeated bronchoscopy for suctioning of aspirated material from the tracheobronchial tree will help minimize the inflammatory reaction of pneumonitis and facilitate improved pulmonary toilet. Forced diuresis to overcome anasarca and over-resuscitation remains controversial and unsubstantiated. Complications of forced diuresis include electrolyte disturbances, replacement of those electrolytes, metabolic alkalosis, hypotension, and acute kidney injury.

Pneumonia is the second most common nosocomial infection and is the most common infection in ventilated patients. Ventilator-associated pneumonia (VAP) occurs in 15% to 40% of ventilated ICU patients, with a probability rate of 5% per day, up to 70% at 30 days. The 30-day mortality rate of nosocomial pneumonia can be as high as 40% and depends on the microorganisms involved and the timeliness of initiating appropriate antimicrobials. Protocol-driven approaches for prevention and treatment of VAP are recognized as beneficial in managing these difficult infectious complications.

Once the diagnosis of pneumonia is suspected (an abnormal chest X-ray, fever, productive cough with purulent sputum, and no other obvious fever sources), it is invariably necessary to initially begin treatment with broad-spectrum antibiotics until proper identification, colony count (≥100,000 colony-forming units [CFU]), and sensitivity of the microorganisms are determined. The spectrum of antibiotic coverage should be narrowed as soon as the culture sensitivities are determined. Double-coverage antibiotic strategy for the two pathogens, *Pseudomonas* and *Acinetobacter* spp., may be appropriate if the local prevalence of these particularly virulent organisms is high. One of the most helpful tools in treating pneumonia and other infections is the tracking of a medical center’s antibiogram every 6 to 12 months.

Epidural analgesia decreases the risk of perioperative pneumonia. This method of pain control improves pulmonary toilet and the early return of bowel function; both have a significant impact on the potential for aspiration and for acquiring pneumonia. The routine use of epidural analgesia results in a lower incidence of pneumonia than patient-controlled analgesia.

Acute lung injury (ALI) was a diagnosis applied to patients with similar findings to those with acute respiratory distress...
syndrome (ARDS). The Berlin definition of ARDS developed by the American-European Consensus Conference of 2012 not only simplifies the definition of ARDS but also eliminates the term ALI from critical care vernacular. ARDS is now classified by partial pressure of oxygen in arterial blood (Pao₂)/fraction of inspired oxygen (Fio₂) ratios as mild (300–201 mmHg), moderate (200–101 mmHg), and severe (<100 mmHg). Elements of modification of the definition include the following: <7 days of onset; removal of pulmonary artery occlusion pressure; and clinical judgment for characterizing hydrostatic pulmonary edema is acceptable, unless risk factors for ARDS have been eliminated, in which case objective analysis is necessary.72-75

The definition of ARDS traditionally included five criteria (Table 12-13). The multicenter ARDS Research Network (ARDSnet) research trial demonstrated improved clinical outcomes for ARDS patients ventilated at tidal volumes of only 5 to 7 mL/kg.76 This strategy is no longer prescribed solely for patients with ARDS but is also recommended for patients with normal pulmonary physiology who are intubated for reasons other than acute respiratory failure. The beneficial effects of positive end-expiratory pressure (PEEP) for ARDS were confirmed in this study as well. The maintenance of PEEP during ventilatory support is determined based on blood gas analysis, pulmonary mechanics, and requirements for supplemental oxygen. As gas exchange improves with resolving ARDS, the initial step in decreasing ventilatory support should be to decrease the levels of supplemental oxygen first, and then to slowly bring the PEEP levels back down to minimal levels.77 This is done to minimize the potential for recurrent alveolar collapse and a worsening gas exchange.

Not all patients can be weaned easily from mechanical ventilation. When the respiratory muscle energy demands are not balanced or there is an ongoing active disease state external to the lungs, patients may require prolonged ventilatory support. Protocol-driven ventilator weaning strategies are successful and have become part of the standard of care. The use of a weaning protocol for patients on mechanical ventilation greater than 48 hours reduces the incidence of VAP and the overall length of ventilator time, and the number of ICU patient days.

The occurrence of PE is probably underdiagnosed. Its etiology is thought to stem from DVT. This concept, however, has recently been questioned by Spaniolas et al.81 The diagnosis of PE is made when a high degree of clinical suspicion for PE leads to imaging techniques such as ventilation–perfusion nuclear scans or CT pulmonary angiogram. Clinical findings include elevated central venous pressure, hypoxemia, shortness of breath, hypocarbia secondary to tachypnea, and right heart strain on ECG. Ventilation–perfusion nuclear scans are often indeterminate in patients who have an abnormal chest X-ray and are less sensitive than a CT angiogram or pulmonary angiogram for diagnosing PE. The pulmonary angiogram remains the gold standard for diagnosing PE, but spiral CT angiogram has become an alternative method because of its relative ease of use and reasonable rates of diagnostic accuracy. For cases without clinical contraindications to therapeutic anticoagulation, patients should be empirically started on heparin infusion until the imaging studies are completed if the suspicion of a PE is high.

Sequential compression devices on the lower extremities and low-dose subcutaneous heparin or low molecular weight heparinoid administration are routinely used to prevent DVT and, by inference, the risk of PE. Neurosurgical and orthopedic patients have higher rates of PE, as do obese patients and those at prolonged bed rest.

When anticoagulation is contraindicated, or when a known clot exists in the inferior vena cava (IVC), decreasing the risk for PE includes insertion of an IVC filter. The Greenfield filter has been most widely studied, and it has a failure rate of less than 4%. Newer devices include those with nitinol wire that expands with body temperature and retrievable filters. Retrievable filters, however, must be considered as permanent. In most studies, the actual retrievable rate only reached about 20%. Some studies recognize the benefit of automated reminders and diligence of outlying patient follow-up, where higher retrieval rates have been achieved.82 Patients with spinal cord injury and multiple long-bone or pelvic fractures frequently receive IVC filters, and

| Table 12-13 |
| Inclusion criteria for the acute respiratory distress syndrome |
| Acute onset |
| Predisposing condition |
| Pao₂:Fio₂ <200 (regardless of positive end-expiratory pressure) |
| Bilateral infiltrates |
| Pulmonary artery occlusion pressure <18 mmHg |
| No clinical evidence of right heart failure |

Fio₂ = fraction of inspired oxygen; Pao₂ = partial pressure of arterial oxygen.
Cardiac System. Arrhythmias are often seen preoperatively in elderly patients but may occur postoperatively in any age group. Atrial fibrillation is the most common arrhythmia and occurs between postoperative days 3 to 5 in high-risk patients. This is typically when patients begin to mobilize their intrastitial fluid into the vascular fluid space. Contemporary evidence suggests that rate control is more important than rhythm control for atrial fibrillation. The first-line treatment includes β-blockade and/or calcium channel blockade. β-Blockade must be used judiciously because hypotension, as well as withdrawal from β-blockade with rebound hypertension, is possible. Calcium channel blockers are an option if β-blockers are not tolerated by the patient, but caution must be exercised in those with a history of congestive heart failure. Although digoxin is still a standby medication, it has limitations due to the need for optimal dosing levels. Cardioversion may be required if patients become hemodynamically unstable and the rhythm cannot be controlled.

Ventricular arrhythmias and other tachyarrhythmias may occur in surgical patients as well. Similar to atrial rhythm problems, these are best controlled with β-blockade, but the use of other antiarrhythmics or cardioversion may be required if patients become hemodynamically unstable.

Cardiac ischemia is a cause of postoperative mortality. Acute myocardial infarction (AMI) can present insidiously, or it can be more dramatic with the classic presentation of shortness of breath, severe angina, and sudden cardiogenic shock. The workup to rule out an AMI includes an ECG and cardiac enzyme measurements. The patient should be transferred to a monitored (telemetry) floor. Morphine, supplemental oxygen, nitroglycerine, and aspirin (MONA) are the initial therapeutic maneuvers for those being investigated for AMI.

Gastrointestinal System. Surgery of the esophagus is potentially complicated because of its anatomic location and blood supply. Nutritional support strategies should be considered for esophageal resection patients to maximize the potential for survival. The two primary types of esophageal resection performed are the transthiatal resection and the transthoracic (Ivor-Lewis) resection. The transthiatal resection has the advantage that a formal thoracotomy incision is avoided. However, dissection of the esophagus is blind, and anastomotic leaks occur more than with other resections. However, when a leak does occur, simple opening of the cervical incision and draining the leak is all that is usually required.

The transthoracic Ivor-Lewis resection includes an esophageal anastomosis performed in the chest near the level of the azygos vein. These have lower leak rates, but the leaks that do occur result in mediastinitis and can be difficult to control. The reported mortality is about 50% with an anastomotic leak, and the overall mortality of the procedure is about 5%, which is similar to transthiatal resection.

Postoperative ileus is related to dysfunction of the neural reflex axis of the intestine. Excessive narcotic use may delay return of bowel function. Epidural anesthesia results in better pain control, and there is an earlier return of bowel function and a shorter length of hospital stay. The limited use of nasogastric tubes and the initiation of early postoperative feeding are associated with an earlier return of bowel function. The use of chewing gum and other oral stimulants to minimize ileus remains controversial.

Pharmacologic agents commonly used to stimulate bowel function include metoclopramide and erythromycin. Metoclopramide’s action is limited to the stomach and duodenum, and it may help primarily with gastroparesis. Erythromycin is a motilin agonist that works throughout the stomach and bowel. Several studies demonstrate significant benefit from the administration of erythromycin in those suffering from an ileus. Alvimopan, a newer agent and a µ-opioid receptor antagonist, has shown some promise in many studies for earlier return of gut function and subsequent reduction in length of stay. Neostigmine has been used in refractory pan-ileus patients (Ogilvie’s syndrome) with some degree of success. It is recommended for patients receiving this type of therapy to be in a monitored unit.

Small bowel obstruction occurs in less than 1% of early postoperative patients. When it does occur, adhesions are usually the cause. Internal and external hernias, technical errors, and infections or abscesses are also causative. Hyaluronidase is a mucolytic enzyme that degrades connective tissue, and the use of a methylcellulose form of hyaluronidase, Seprafilm®, has been shown to result in a 50% decrease in adhesion formation in some patients. This may translate into a lower occurrence of postoperative bowel obstruction, but has yet to be proven.

Fistulae are the abnormal communication of one structure to an adjacent structure or compartment and are associated with extensive morbidity and mortality. Common causes for fistula formation are summarized in the mnemonic FRIENDS (Foreign body, Radiation, Ischemia/Inflammation/Infection, Epithelialization of a tract, Neoplasia, Distal obstruction, and Steroid use). Postoperatively, they are most often caused by infection or obstruction leading to an anastomotic leak. The cause of the fistula must be recognized early, and treatment may include nonoperative management with observation and nutritional support, or a delayed operative management strategy that also includes nutritional support and wound care.

Gastrointestinal (GI) bleeding can occur perioperatively (Table 12-14). Technical errors such as a poorly tied suture, a nonhemostatic staple line, or a missed injury can all lead to

### Table 12-14

| Common causes of upper and lower gastrointestinal (GI) hemorrhage |
|---------------------------------|----------------|
| UPPER GI BLEED                  | LOWER GI BLEED |
| Erosive esophagitis             | Angiodysplasia |
| Gastric varices                 | Radiation proctitis |
| Esophageal varices              | Hemangioma     |
| Dieulafoy’s lesion              | Diverticulosis |
| Aortoduodenal fistula           | Neoplastic diseases |
| Mallory-Weiss tear              | Trauma         |
| Peptic ulcer disease            | Vasculitis     |
| Trauma                          | Hemorrhoids    |
| Neoplastic disease              | Aortoenteric fistula |
|                                 | Intussusception |
|                                 | Ischemic colitis |
|                                 | Inflammatory bowel disease |
|                                 | Postprocedure bleeding |
postoperative intestinal bleeding.\textsuperscript{94,95} The source of bleeding is in the upper GI tract about 85% of the time and is usually detected and treated endoscopically. Surgical control of intestinal bleeding is required in up to 40% of patients.\textsuperscript{96}

When patients in the ICU have a major bleed from stress gastritis, the mortality risk is as high as 50%. It is important to keep the gastric pH greater than 4 to decrease the overall risk for stress gastritis in patients mechanically ventilated for 48 hours or greater and patients who are coagulopathic.\textsuperscript{97} Proton pump inhibitors, $\text{H}_2$-receptor antagonists, and intragastric antacid installation are all effective measures. However, patients who are not mechanically ventilated or who do not have a history of gastritis or peptic ulcer disease should not be placed on gastritis prophylaxis postoperatively because it carries a higher risk of causing pneumonia.

**Hepatobiliary-Pancreatic System.** Complications involving the hepatobiliary system are usually due to technical errors. Laparoscopic cholecystectomy has become the standard of care for cholecystectomy, but common bile duct injury remains a nemesis of this approach. Intraoperative cholangiography has not been shown to decrease the incidence of common bile duct injuries because the injury to the bile duct usually occurs before the cholangiogram.\textsuperscript{98,99} Early recognition and immediate repair of an injury are important because delayed bile duct leaks often require a more complex repair.

Ischemic injury due to devascularization of the common bile duct has a delayed presentation days to weeks after an operation. Endoscopic retrograde cholangiopancreatography (ERCP) demonstrates a stenotic, smooth common bile duct, and liver function studies are elevated. The recommended treatment is a Roux-en-Y hepaticojejunostomy.

A bile leak due to an unrecognized injury to the ducts may present after cholecystectomy as a biloma. These patients may present with abdominal pain and hyperbilirubinemia. The diagnosis of a biliary leak can be confirmed by CT scan, ERCP, or radionuclide scan. Once a leak is confirmed, a retrograde biliary stent and external drainage are the treatment of choice.

Hyperbilirubinemia in the surgical patient can be a complex problem. Cholestasis makes up the majority of causes for hyperbilirubinemia, but other mechanisms of hyperbilirubinemia include reabsorption of blood (e.g., hematoma from trauma), decreased bile excretion (e.g., sepsis), increased unconjugated bilirubin due to hemolysis, hyperthyroidism, and impaired excretion due to congenital abnormalities or acquired disease. Errors in surgery that cause hyperbilirubinemia largely involve missed or iatrogenic injuries.

The presence of cirrhosis predisposes to postoperative complications. Abdominal or hepatobiliary surgery is problematic in the cirrhotic patient. Ascites leak in the postoperative period can be an issue when any abdominal operation has been performed. Maintaining proper intravascular oncotic pressure in the immediate postoperative period can be difficult, and resuscitation should be maintained with crystalloid solutions. Prevention of renal failure and the management of the hepatorenal syndrome can be difficult, as the demands of fluid resuscitation and altered glomerular filtration become competitive. Spironolactone with other diuretic agents may be helpful in the postoperative care. These patients often have a labile course, and bleeding complications due to coagulopathy are common. The operative mortality in cirrhotic patients is 10% for Child class A, 30% for Child class B, and 82% for Child class C patients.\textsuperscript{100}

Pyogenic liver abscess occurs in less than 0.5% of adult admissions, due to retained necrotic liver tissue, occult intestinal perforations, benign or malignant hepatobiliary obstruction, sepsis, and hepatic arterial occlusion. The treatment is long-term antibiotics with percutaneous drainage of large abscesses.

Pancreatitis can occur following injection of contrast during cholangiography and after endoscopic cholangiopancreatography (ERCP). These episodes range from a mild elevation in amylase and lipase with abdominal pain, to a fulminant course of pancreatitis with necrosis requiring surgical debridement. The incidence of post-ERCP pancreatitis has been shown to be reduced by the administration of rectal indomethacin.\textsuperscript{101} Studies are underway to determine whether the prophylactic use of pancreatic duct stenting in patients at high risk for post-ERCP pancreatitis can be avoided with the use of rectal indomethacin.

Traumatic injuries to the pancreas can occur during surgical procedures on the kidneys, GI tract, and spleen most commonly. Treatment involves serial CT scans and percutaneous drainage to manage infected fluid and abscess collections; sterile collections should not be drained because drain placement can introduce infection. A pancreatic fistula may respond to antisecretory therapy with a somatostatin analogue. Management of these fistulae initially includes ERCP with or without pancreatic stenting, percutaneous drainage of any fistula fluid collections, total parenteral nutrition (TPN) with bowel rest, and repeated CT scans. The majority of pancreatic fistulae will eventually heal spontaneously.

**Renal System.** Renal failure can be classified as prerenal failure, intrinsic renal failure, and postrenal failure. Postrenal failure, or obstructive renal failure, should always be considered when low urine output (oliguria) or anuria occurs. The most common cause is a misplaced or clogged urinary catheter. Other, less common causes to consider are unintentional ligation or transection of ureters during a difficult surgical dissection (e.g., colon resection for diverticular disease) or a large retroperitoneal hematoma (e.g., ruptured aortic aneurysm).

Oliguria is initially evaluated by flushing the urinary catheter using sterile technique. Urine electrolytes should also be measured (Table 12-15). A hemoglobin and hematocrit level should be checked immediately. Patients in compensated shock from acute blood loss may manifest anemia and end-organ malperfusion as oliguria.

Acute tubular necrosis (ATN) carries a mortality risk of 25% to 50% due to the many complications that can cause, or result from, this insult. When ATN is due to poor inflow (prerenal failure), the remedy begins with IV administration of crystalloid or colloid fluids as needed. If cardiac insufficiency is the problem, the optimization of vascular volume is achieved first, followed by inotropic agents, as needed. Intrinsic renal failure

<table>
<thead>
<tr>
<th>FE$<em>{\text{Na}}$, OSMOLARITY, UR$</em>{\text{Na}}$, ETIOLOGY</th>
<th>PRERENAL</th>
<th>INTRINSIC FAILURE</th>
</tr>
</thead>
<tbody>
<tr>
<td>FE$<em>{\text{Na}}$&lt;1, OSMOLARITY&gt;500, UR$</em>{\text{Na}}$&lt;20</td>
<td>CHF, cirrhosis</td>
<td></td>
</tr>
<tr>
<td>FE$<em>{\text{Na}}$&gt;1, OSMOLARITY&lt;350, UR$</em>{\text{Na}}$&gt;40</td>
<td>Sepsis, shock</td>
<td></td>
</tr>
</tbody>
</table>

CHF = congestive heart failure; FE$_{\text{Na}}$ = fractional excretion of sodium; UR$_{\text{Na}}$ = urinary excretion of sodium.
and subsequent ATN are often the result of direct renal toxins. Aminoglycosides, vancomycin, and furosemide, among other commonly used agents, contribute directly to nephrotoxicity. Contrast-induced nephropathy usually leads to a subtle or transient rise in creatinine. In patients who are volume depleted or have poor cardiac function, contrast nephropathy may permanently impair renal function.102-105

The treatment of renal failure due to myoglobinuria has shifted away from the use of sodium bicarbonate for alkalinizing the urine, to merely maintaining brisk urine output of 100 mL per hour with crystalloid fluid infusion. Mannitol and furosemide are not recommended. Patients who do not respond to resuscitation are at risk for needing renal replacement therapy. Fortunately, most of these patients eventually recover from their renal dysfunction.

Musculoskeletal System. A compartment syndrome can develop in any compartment of the body. Compartment syndrome of the extremities generally occurs after a closed fracture. The injury alone may predispose the patient to compartment syndrome, but aggressive fluid resuscitation can exacerbate the problem. Pain with passive motion is the hallmark of compartment syndrome, and the anterior compartment of the leg is usually the first compartment to be involved. Confirmation of the diagnosis is obtained by direct pressure measurement of the individual compartments. If the pressures are greater than 20 to 25 mmHg in any of the compartments, then a four-compartment fasciotomy is considered. Compartment syndrome can be due to ischemia-reperfusion injury, after an ischemic time of 4 to 6 hours. Renal failure (due to myoglobinuria), tissue loss, and a permanent loss of function are possible results of untreated compartment syndrome.

Decubitus ulcers are preventable complications of prolonged bed rest due to traumatic paralysis, dementia, chemical paralysis, or coma. Unfortunately, they are still occurring despite extensive research and clinical initiatives that demonstrate successful prevention strategies. Ischemic changes in the microcirculation of the skin can be significant after 2 hours of sustained pressure. Routine skin care and turning of the patient help ensure a reduction in skin ulceration. This can be labor intensive, and special mattresses and beds are available to help. The treatment of a decubitus ulcer in the noncoagulopathic patient is surgical debridement. Once the wound bed has a viable granulation base without an excess of fibrinous debris, a vacuum-assisted closure dressing can be applied. Wet to moist dressings with frequent dressing changes is the alternative and is labor intensive. Expensive topical enzyme preparations are also available. If the wounds fail to respond to these measures, soft tissue coverage by flap is considered.

Contractures are the result of muscle disuse. Whether from trauma, amputation, or vascular insufficiency, contractures can be prevented by physical therapy and splinting. If not attended to early, contractures will prolong rehabilitation and may lead to further wounds and wound healing issues. Depending on the functional status of the patient, contracture releases may be required for long-term care.

Hematologic System. The traditional transfusion guideline of maintaining the hematocrit level in all patients at greater than 30% is no longer valid. Only patients with symptomatic anemia, who have significant cardiac disease, or who are critically ill and require increased oxygen-carrying capacity to adequately perfuse end organs require higher levels of hemoglobin. Other than these select patients, the decision to transfuse should generally not occur until the hemoglobin level falls to 7 mg/dL or the hematocrit reaches 21%.

Transfusion reactions are common complications of blood transfusion. These can be attenuated with a leukocyte filter, but not completely prevented. The manifestations of a transfusion reaction include simple fever, pruritus, chills, muscle rigidity, and renal failure due to myoglobinuria secondary to hemolysis. Discontinuing the transfusion and returning the blood products to the blood bank is an important first step, but administration of antihistamine and possibly steroids may be required to control the reaction symptoms. Severe transfusion reactions are rare but can be fatal.

Infectious complications in blood transfusion range from cytomegalovirus transmission, which is benign in the nontransplant patient, to human immunodeficiency virus (HIV) infection, to passage of the hepatitis viruses (Table 12-16).

Patients on warfarin (Coumadin) who require surgery can have anticoagulation reversal by administration of fresh frozen plasma. Each unit of fresh frozen plasma contains 200 to 250 mL of plasma and includes one unit of coagulation factor per milliliter of plasma.

Thrombocytopenia may require platelet transfusion for a platelet count less than 20,000/mL when invasive procedures are performed, or when platelet counts are low and ongoing bleeding from raw surface areas persists. One unit of platelets will increase the platelet count by 5000 to 7500 per mL in adults. It is important to delineate the cause of the low platelet count. Usually there is a self-limiting or reversible condition such as sepsis. Rarely, it is due to heparin-induced thrombocytopenia I and II. Complications of heparin-induced thrombocytopenia II can be serious because of the diffuse thrombogenic nature of the disorder. Simple precautions to limit this hypercoagulable state include saline solution flushes instead of heparin solutions and limiting the use of heparin-coated catheters. The treatment is anticoagulation with synthetic agents such as argatroban.

For patients with uncontrollable bleeding due to disseminated intravascular coagulopathy (DIC), a potentially useful drug is factor VIIa, but its use should be judicious.106-109 Originally used in hepatic trauma and obstetric emergencies, this agent was lifesaving in some circumstances. The CONTROL Trial,109 however, has largely decreased overuse of this agent because investigators demonstrated no benefit over simple factor replacement in severely coagulopathic patients. Factor VIIa use may also be limited due to its potential thrombotic complications. For some situations, the combination of ongoing.

### Table 12-16

<table>
<thead>
<tr>
<th></th>
<th>Rate of viral transmission in blood product transfusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV</td>
<td>1:1.9 million</td>
</tr>
<tr>
<td>HBV</td>
<td>1:137,000</td>
</tr>
<tr>
<td>HCV</td>
<td>1:1 million</td>
</tr>
</tbody>
</table>

*Post-nucleic acid amplification technology (1999). Earlier rates were erroneously reported higher due to lack of contemporary technology.

*HBV is reported with prenucleic acid amplification technology. Statistical information is unavailable with postnucleic acid amplification technology at this writing.

Note that bacterial transmission is 50 to 250 times higher than viral transmission per transfusion.

HBV = hepatitis B virus; HCV = hepatitis C virus.
nonsurgical bleeding and renal failure can occasionally be successfully treated with desmopressin.

In addition to classic hemophilia, other inherited coagulation factor deficiencies can be difficult to manage in surgery. When required, transfusion of appropriate replacement products is coordinated with the regional blood bank center before surgery. Other blood dyscrasias seen by surgeons include hypercoagulopathic patients. Those who carry congenital anomalies such as the most common factor V Leiden deficiency, as well as protein C and S deficiencies, are likely to form thromboses if inadequately anticoagulated, and these patients should be managed in consultation with a hematologist.

**Abdominal Compartment Syndrome.** Multisystem trauma, thermal burns, retroperitoneal injuries, and surgery related to the retroperitoneum are the major initial causative factors that may lead to abdominal compartment syndrome (ACS). Ruptured AAA, major pancreatic injury and resection, or multiple intestinal injuries are also examples of clinical situations in which a large volume of IV fluid resuscitation puts these patients at risk for intra-abdominal hypertension. Manifestations of ACS typically include progressive abdominal distention followed by increased peak airway ventilator pressures, oliguria followed by anuria, and an insidious development of intracranial hypertension. These findings are related to elevation of the diaphragm and inadequate venous return from the vena cava or renal veins secondary to the transmitted pressure on the venous system.

Measurement of abdominal pressures is easily accomplished by transducing bladder pressures from the urinary catheter after instilling 100 mL of sterile saline into the urinary bladder. A pressure greater than 20 mmHg constitutes intra-abdominal hypertension, but the diagnosis of ACS requires intra-abdominal pressure greater than 25 to 30 mmHg, with at least one of the following: compromised respiratory mechanics and ventilation, oliguria or anuria, or increasing intracranial pressures.

The treatment of ACS is to open any recent abdominal incision to release the abdominal fascia or to open the fascia directly if no abdominal incision is present. Immediate improvement in mechanical ventilation pressures, intracranial pressures, and urine output is usually noted. When expectant management for ACS is considered in the OR, the abdominal fascia should be left open and covered under sterile conditions (e.g., a vacuum-assisted open abdominal wound closure system) with plans made for a second-look operation and delayed fascial closure. Patients with intra-abdominal hypertension should be monitored closely with repeated examinations and measurements of bladder pressure, so that any further deterioration is detected and operative management can be initiated. Left untreated, ACS may lead to multiple system end-organ dysfunction or failure and has a high mortality.

Abdominal wall closure should be attempted every 48 to 72 hours until the fascia can be reapproximated. If the abdomen cannot be closed within 5 to 7 days following release of the abdominal fascia, a large incisional hernia is the net result. A variety of surgical options have evolved for prevention and closure of the resultant hernias, but no standard approach has yet evolved.

**Wounds, Drains, and Infection**

**Wound (Surgical Site) Infection.** No prospective, randomized, double-blind, controlled studies exist that demonstrate antibiotics used beyond 24 hours in the perioperative period prevent infections. Prophylactic use of antibiotics should simply not be continued beyond this time. Irrigation of the operative field and the surgical wound with saline solution has shown benefit in controlling wound inoculum. Irrigation with an antibiotic-based solution has not demonstrated significant benefit in controlling postoperative infection.

Antibacterial-impregnated polyvinyl placed over the operative wound area for the duration of the surgical procedure has not been shown to decrease the rate of wound infection.

Although skin preparation with 70% isopropyl alcohol has the best bactericidal effect, it is flammable and could be hazardous when electrocautery is used. The contemporary formulas of chlorhexidine gluconate with isopropyl alcohol remain more advantageous.

There is a difference between wound colonization and infection. Over-treating colonization is just as injurious as undertreating infection. The strict definition of wound infection is more than 10^5 CFU per gram of tissue. This warrants expeditious and proper antibiotic/antifungal treatment. Often, however, clinical signs raise enough suspicion that the patient is treated before a confirmatory culture is undertaken. The clinical signs of wound infection include rubor, tumor, calor, and dolor (redness, swelling, heat, and pain). Once the diagnosis of wound infection has been established, the most definitive treatment remains open drainage of the wound. The use of antibiotics for wound infection treatment should be limited.

One type of wound dressing/drainage system that has gained popularity is the vacuum-assisted closure dressing. The principle of the system is to decrease local wound edema and to promote healing through the application of a sterile dressing that is then covered and placed under controlled suction for a period of 2 to 4 days at a time. Although costly, the benefits are frequently dramatic and may offset the costs of nursing care, frequent dressing changes, and operative wound debridement.

**Drain Management.** The four indications for applying a surgical drain are:

- To collapse surgical dead space in areas of redundant tissue (e.g., neck and axilla)
- To provide focused drainage of an abscess or grossly infected surgical site
- To provide early warning notice of a surgical leak (either bowel contents, secretions, urine, air, or blood)—the so-called sentinel drain
- To control an established fistula leak

Open drains are often used for large contaminated wounds such as perirectal or perianal fistulas and subcutaneous abscess cavities. They prevent premature closure of an abscess cavity in a contaminated wound. More commonly, surgical sites are drained by closed suction drainage systems, but data do not support closed suction drainage to “protect an anastomosis” or to “control a leak” when placed at the time of surgery. Closed suction devices can exert a negative pressure of 70 to 170 mmHg at the level of the drain; therefore, the presence of this excess suction may call into question whether an anastomosis breaks down on its own or whether the drain creates a suction injury that promotes leakage (Fig. 12-9). On the other hand, CT- or ultrasound-guided placement of percutaneous drains is now the standard of care for abscesses, loculated infections, and other isolated fluid collections such as protein C and S deficiencies, are likely to form thromboses if inadequately anticoagulated, and these patients should be managed in consultation with a hematologist.

**Abdominal Compartment Syndrome.** Multisystem trauma, thermal burns, retroperitoneal injuries, and surgery related to the retroperitoneum are the major initial causative factors that may lead to abdominal compartment syndrome (ACS). Ruptured AAA, major pancreatic injury and resection, or multiple intestinal injuries are also examples of clinical situations in which a large volume of IV fluid resuscitation puts these patients at risk for intra-abdominal hypertension. Manifestations of ACS typically include progressive abdominal distention followed by increased peak airway ventilator pressures, oliguria followed by anuria, and an insidious development of intracranial hypertension. These findings are related to elevation of the diaphragm and inadequate venous return from the vena cava or renal veins secondary to the transmitted pressure on the venous system.

Measurement of abdominal pressures is easily accomplished by transducing bladder pressures from the urinary catheter after instilling 100 mL of sterile saline into the urinary bladder. A pressure greater than 20 mmHg constitutes intra-abdominal hypertension, but the diagnosis of ACS requires intra-abdominal pressure greater than 25 to 30 mmHg, with at least one of the following: compromised respiratory mechanics and ventilation, oliguria or anuria, or increasing intracranial pressures.

The treatment of ACS is to open any recent abdominal incision to release the abdominal fascia or to open the fascia directly if no abdominal incision is present. Immediate improvement in mechanical ventilation pressures, intracranial pressures, and urine output is usually noted. When expectant management for ACS is considered in the OR, the abdominal fascia should be left open and covered under sterile conditions (e.g., a vacuum-assisted open abdominal wound closure system) with plans made for a second-look operation and delayed fascial closure. Patients with intra-abdominal hypertension should be monitored closely with repeated examinations and measurements of bladder pressure, so that any further deterioration is detected and operative management can be initiated. Left untreated, ACS may lead to multiple system end-organ dysfunction or failure and has a high mortality.

Abdominal wall closure should be attempted every 48 to 72 hours until the fascia can be reapproximated. If the abdomen cannot be closed within 5 to 7 days following release of the abdominal fascia, a large incisional hernia is the net result. A variety of surgical options have evolved for prevention and closure of the resultant hernias, but no standard approach has yet evolved.

**Wounds, Drains, and Infection**

**Wound (Surgical Site) Infection.** No prospective, randomized, double-blind, controlled studies exist that demonstrate antibiotics used beyond 24 hours in the perioperative period prevent infections. Prophylactic use of antibiotics should simply not be continued beyond this time. Irrigation of the operative field and the surgical wound with saline solution has shown benefit in controlling wound inoculum. Irrigation with an antibiotic-based solution has not demonstrated significant benefit in controlling postoperative infection.

Antibacterial-impregnated polyvinyl placed over the operative wound area for the duration of the surgical procedure has not been shown to decrease the rate of wound infection.

Although skin preparation with 70% isopropyl alcohol has the best bactericidal effect, it is flammable and could be hazardous when electrocautery is used. The contemporary formulas of chlorhexidine gluconate with isopropyl alcohol remain more advantageous.

There is a difference between wound colonization and infection. Over-treating colonization is just as injurious as undertreating infection. The strict definition of wound infection is more than 10^5 CFU per gram of tissue. This warrants expeditious and proper antibiotic/antifungal treatment. Often, however, clinical signs raise enough suspicion that the patient is treated before a confirmatory culture is undertaken. The clinical signs of wound infection include rubor, tumor, calor, and dolor (redness, swelling, heat, and pain). Once the diagnosis of wound infection has been established, the most definitive treatment remains open drainage of the wound. The use of antibiotics for wound infection treatment should be limited.

One type of wound dressing/drainage system that has gained popularity is the vacuum-assisted closure dressing. The principle of the system is to decrease local wound edema and to promote healing through the application of a sterile dressing that is then covered and placed under controlled suction for a period of 2 to 4 days at a time. Although costly, the benefits are frequently dramatic and may offset the costs of nursing care, frequent dressing changes, and operative wound debridement.

**Drain Management.** The four indications for applying a surgical drain are:

- To collapse surgical dead space in areas of redundant tissue (e.g., neck and axilla)
- To provide focused drainage of an abscess or grossly infected surgical site
- To provide early warning notice of a surgical leak (either bowel contents, secretions, urine, air, or blood)—the so-called sentinel drain
- To control an established fistula leak

Open drains are often used for large contaminated wounds such as perirectal or perianal fistulas and subcutaneous abscess cavities. They prevent premature closure of an abscess cavity in a contaminated wound. More commonly, surgical sites are drained by closed suction drainage systems, but data do not support closed suction drainage to “protect an anastomosis” or to “control a leak” when placed at the time of surgery. Closed suction devices can exert a negative pressure of 70 to 170 mmHg at the level of the drain; therefore, the presence of this excess suction may call into question whether an anastomosis breaks down on its own or whether the drain creates a suction injury that promotes leakage (Fig. 12-9). On the other hand, CT- or ultrasound-guided placement of percutaneous drains is now the standard of care for abscesses, loculated infections, and other isolated fluid collections such as
QUALITY, PATIENT SAFETY, ASSESSMENTS OF CARE, AND COMPLICATIONS

CHAPTER 12

Figure 12-9 This illustration demonstrates typical intraoperative placement of closed suction devices in pancreatic or small bowel surgery, where there may be an anastomosis. At negative pressures of 70 to 170 mmHg, these devices may actually encourage anastomotic leaks and not prevent them or become clogged by them.

as pancreatic leaks. The risk of surgery is far greater than the placement of an image-guided drain.

The use of antibiotics when drains are in place is often unnecessary as the drain provides direct source control. Twenty-four to 48 hours of antibiotic use after drain placement is prophylactic, and after this period, only specific treatment of positive cultures should be performed to avoid increased drug resistance and superinfection.

**Urinary Catheters.** Several complications of urinary catheters can occur that lead to an increased length of hospital stay and morbidity. In general, use of urinary catheters should be minimized and every opportunity to expeditiously remove them should be encouraged. If needed, it is recommended that the catheter be inserted its full length up to the hub and that urine flow is established before the balloon is inflated because misplacement of the catheter in the urethra with premature inflation of the balloon can lead to tears and disruption of the urethra.

Enlarged prostatic tissue can make catheter insertion difficult, and a catheter could be required. If this attempt is also unsuccessful, then a urologic consultation for endoscopic placement of the catheter may be required to prevent harm to the urethra. For patients with urethral strictures, filiform-tipped catheters and followers may be used, but these can potentially cause bladder injury. If endoscopic attempts fail, the patient may require a percutaneously placed suprapubic catheter to obtain decompression of the bladder. Follow-up investigations of these patients are recommended so definitive care of the urethral abnormalities can be pursued.

The most frequent nosocomial infection is urinary tract infection (UTI). These infections are classified into complicated and uncomplicated forms. The uncomplicated type is a UTI that can be treated with outpatient antibiotic therapy. The complicated UTI usually involves a hospitalized patient with an indwelling catheter whose UTI is diagnosed as part of a fever workup. The interpretation of urine culture results of less than 100,000 CFU/mL is controversial. Before treating such a patient, one should change the catheter and then repeat the culture to see if the catheter was simply colonized with organisms. Cultures with more than 100,000 CFU/mL should be treated with the appropriate antibiotics and the catheter changed or removed as soon as possible. Undertreatment or misdiagnosis of a UTI can lead to urosepsis and septic shock.

Recommendations are mixed on the proper way to treat *Candida albicans* fungal bladder infections. Continuous bladder washings with fungicidal solution for 72 hours have been recommended, but this is not always effective. Replacement of the urinary catheter and a course of fluconazole are appropriate treatments, but some infectious disease specialists claim that *C. albicans* in the urine may serve as an indication of fungal infection elsewhere in the body. If this is the case, then screening cultures for other sources of fungal infection should be performed whenever a fungal UTI is found.

**Empyema.** One of the most debilitating infections is an empyema, or infection of the pleural space. Frequently, an overwhelming pneumonia is the source of an empyema, but a retained hemothorax, systemic sepsis, esophageal perforation from any cause, and infections with a predilection for the lung (e.g., tuberculosis) are potential etiologies as well. The diagnosis is confirmed by chest X-ray or CT scan, followed by aspiration of pleural fluid for bacteriologic analysis. Gram’s stain, lactate dehydrogenase, protein, pH, and cell count are obtained, and broad-spectrum antibiotics are initiated while the laboratory studies are performed. Once the specific organisms are confirmed, anti-infective agents are tailored appropriately. Placement of a thoracostomy tube is needed to evacuate and drain the infected pleural fluid, but depending on the specific nidus of infection, video-assisted thoracoscopy may also be
helpful for irrigation and drainage of the infection. Refractory empyemas require specialized surgical approaches.

**Abdominal Abscesses.** Postoperative intra-abdominal abscesses can present with vague complaints of intermittent abdominal pain, fever, leukocytosis, and a change in bowel habits. Depending on the type and timing of the original procedure, the clinical assessment of these complaints is sometimes difficult, and a CT scan is usually required. When a fluid collection within the peritoneal cavity is found on CT scan, antibiotics and percutaneous drainage of the collection is the treatment of choice. Initial antibiotic treatment is usually with broad-spectrum antibiotics such as piperacillin-tazobactam or imipenem. Should the patient exhibit signs of peritonitis and/or have free air on X-ray or CT scan, then re-exploration should be considered.

For patients who present primarily (i.e., not postoperatively) with the clinical and radiologic findings of an abscess but are clinically stable, the etiology of the abscess must be determined. A plan for drainage of the abscess and decisions about further diagnostic studies with consideration of the timing of any definitive surgery all need to be balanced. This can be a complex set of decisions, depending on the etiology (e.g., appendicitis or diverticulitis), but if the patient exhibits signs of peritonitis, urgent surgical exploration should be performed.

**Necrotizing Fasciitis.** Postoperative infections that progress to the fulminant soft tissue infection known as necrotizing fasciitis are uncommon. Group A streptococcal (M types 1, 3, 12, and 28) soft tissue infections, as well as infections with *Clostridium perfringens* and *C. septicum*, carry a mortality of 30% to 70%. Septic shock can be present, and patients can become hypotensive less than 6 hours following inoculation. Manifestations of a group A *Streptococcus pyogenes* infection in its most severe form include hypotension, renal insufficiency, coagulopathy, hepatic insufficiency, ARDS, tissue necrosis, and erythematous rash.

These findings constitute a surgical emergency, and the mainstay of treatment remains wide debridement of the necrotic tissue to the level of bleeding, viable tissue. A gray serous fluid at the level of the necrotic tissue is usually noted, and as the infection spreads, thrombosed blood vessels are noted along the tissue planes involved with the infection. Typically, the patient requires serial trips to the OR for wide debridement until the infection is under control. Antibiotics are an important adjunct to surgical debridement, and broad-spectrum coverage should be used because these infections may be polymicrobial (i.e., so-called mixed-synergistic infections). *Streptococcus pyogenes* is eradicated with penicillin, and it should still be used as the initial drug of choice.

**Systemic Inflammatory Response Syndrome, Sepsis, and Multiple-Organ Dysfunction Syndrome.** The systemic inflammatory response syndrome (SIRS) and the multiple-organ dysfunction syndrome (MODS) carry significant mortality risks (Table 12-17). Specific criteria have been established for the diagnosis of SIRS (Table 12-18), but two criteria are not required for the diagnosis of SIRS: lowered blood pressure and blood cultures positive for infection. SIRS is the result of proinflammatory cytokines related to tissue malperfusion or injury. The dominant cytokines implicated in this process include interleukin (IL)-1, IL-6, and tissue necrosis factor (TNF). Other mediators include nitric oxide, inducible macrophage-type nitric oxide synthase, and prostaglandin $\alpha$.

### Table 12-17

<table>
<thead>
<tr>
<th>PROGNOSIS</th>
<th>MORTALITY (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 SIRS criteria</td>
<td>5</td>
</tr>
<tr>
<td>3 SIRS criteria</td>
<td>10</td>
</tr>
<tr>
<td>4 SIRS criteria</td>
<td>15–20</td>
</tr>
</tbody>
</table>

Sepsis is categorized as sepsis, severe sepsis, and septic shock. Sepsis is SIRS plus infection. Severe sepsis is sepsis plus signs of cellular hypoperfusion or end-organ dysfunction. Septic shock is sepsis plus hypotension after adequate fluid resuscitation.

MODS is the culmination of septic shock and multiple end-organ failure. Usually there is an inciting event (e.g., perforated sigmoid diverticulitis), and as the patient undergoes resuscitation, he or she develops cardiac hypokinesis and oliguric or anuric renal failure, followed by the development of ARDS and eventually septic shock with death.

The international Surviving Sepsis Campaign (www.sccm.org/Documents/SSC-Guidelines.pdf) continues to demonstrate the importance of early recognition and initiation of specific treatment guidelines for optimal management of sepsis. Management of SIRS/MODS includes aggressive global resuscitation and support of end-organ perfusion, correction of the inciting etiology, control of infectious complications, and management of iatrogenic complications. Drotrecogin-$\alpha$, or recombinant activated protein C, appears to specifically counteract the cytokine cascade of SIRS/MODS, but its use is still limited. Other adjuncts for supportive therapy include tight glucose control, low tidal volumes in ARDS, vasopressin in septic shock, and steroid replacement therapy.

**Nutritional and Metabolic Support Complications**

**Nutrition-Related Complications.** A basic principle is to use enteral feeding whenever possible, but complications can intervene such as aspiration, ileus, and to a lesser extent, sinusitis. There is no difference in aspiration rates when a small-caliber feeding tube is placed postpyloric or if it remains in the stomach. Patients who are fed via nasogastric tubes are at risk for aspiration pneumonia because these large-bore tubes stent open the gastroesophageal junction, creating the possibility of gastric reflux. The use of enteric and gastric feeding tubes obviates

### Table 12-18

<table>
<thead>
<tr>
<th>Inclusion criteria for the systemic inflammatory response syndrome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature $&gt;$38°C or $&lt;$36°C ($&gt;$100.4°F or $&lt;$96.8°F)</td>
</tr>
<tr>
<td>Heart rate $&gt;$90 beats/min</td>
</tr>
<tr>
<td>Respiratory rate $&gt;$20 breaths/min or $Paco_2$ $&lt;$32 mmHg</td>
</tr>
<tr>
<td>White blood cell count $&lt;$4000 or $&gt;$12,000 cells/mm$^3$ or $&gt;$10% immature forms</td>
</tr>
</tbody>
</table>

$Paco_2$ = partial pressure of arterial carbon dioxide.
complications of TPN, such as pneumothorax, line sepsis, upper extremity DVT, and the related expense. There is growing evidence to support the initiation of enteral feeding in the early postoperative period, before the return of bowel function, where it is usually well tolerated.

In patients who have had any type of nasal intubation who are having high, unexplained fevers, sinusitis must be entertained as a diagnosis. CT scan of the sinuses is warranted, followed by aspiration of sinus contents so the organism(s) are appropriately treated.

Patients who have not been enterally fed for prolonged periods secondary to multiple operations, those who have had enteral feeds interrupted for any other reason, or those with poor enteral access are at risk for the refeeding syndrome, which is characterized by severe hypophosphatemia and respiratory failure. Slow progression of the enteral feeding administration rate can avoid this complication.

Common TPN problems are mostly related to electrolyte abnormalities that may develop. These electrolyte errors include deficits or excesses in sodium, potassium, calcium, magnesium, and phosphate. Acid-base abnormalities can also occur with the improper administration of acetate or bicarbonate solutions.

The most common cause for hypernatremia in hospitalized patients is under-resuscitation, and, conversely, hyponatremia is most often caused by fluid overload. Treatment for hyponatremia is fluid restriction in mild or moderate cases and the administration of hypertonic saline for severe cases. An overly rapid correction of the sodium abnormality may result in central pontine myelinolysis, which results in a severe neurologic deficit. Treatment for hyponatremic patients includes fluid restriction to correct the free water deficit by 50% in the first 24 hours. An overcorrection of hyponatremia can result in severe cerebral edema, a neurologic deficit, or seizures.

**Glycemic Control.** In 2001, Van den Berghe and colleagues demonstrated that tight glycemic control by insulin infusion is associated with a 50% reduction in mortality in the critical care setting. This prospective, randomized, controlled trial of 1500 patients had two study arms: the intensive-control arm, where the serum glucose was maintained between 80 and 110 mg/dL with insulin infusion; and the control arm, where patients received an insulin infusion only if blood glucose was greater than 215 mg/dL, but serum glucose was then maintained at 180 to 200 mg/dL.

The tight glycemic control group had an average serum glucose level of 103 mg/dL, and the average glucose level in the control group was 153 mg/dL. Hypoglycemic episodes (glucose <40 mg/dL) occurred in 39 patients in the tightly controlled group, while the control group had episodes in six patients. The overall mortality was reduced from 8% to 4.6%, but the mortality of those patients whose ICU stay lasted longer than 5 days was reduced from 20% to 10%. Secondary findings included an improvement in overall morbidity, a decreased percentage of ventilator days, less renal impairment, and a lower incidence of bloodstream infections. These finding have been corroborated by subsequent similar studies, and the principal benefit appears to be a greatly reduced incidence of nosocomial infections and sepsis. It is not known whether the benefits are due to strict euglycemia, to the anabolic properties of insulin, or both, but the maintenance of strict euglycemia between 140 and 180 mg/dL appears to be a powerful therapeutic strategy. A number of studies followed this sentinel publication of tight glycemic control. The NICE-SUGAR trials revisited the Van den Berghe study and found that the glycemic goals found initially to improve outcomes in critically ill patients were now found to be associated with a higher mortality when glucose was kept below 180 mg/dL, due to an increase in incidents of hypoglycemia. When targeted goals of 180 mg/dL are achieved, fewer occurrences of hypoglycemia have been documented, and improved survivorship has been achieved. In addition, some studies find no relationship between glycemic control and improved outcomes. Thus, glycemic control in the critically ill still remains unclear and elusive at best. Part of the difficulty in achieving “tight glycemic control” is the necessity for frequent (every 1–2 hours) blood glucose determinations. When this is performed, glycemic control is enhanced and hypoglycemia is avoided.

**Metabolism-Related Complications.** “Stress dose steroids” have been advocated for the perioperative treatment of patients on corticosteroid therapy, but recent studies strongly discourage the use of supraphysiologic doses of steroids when patients are on low or maintenance doses (e.g., 5–15 mg) of prednisone daily. Parenteral glucocorticoid treatment need only replicate physiologic replacement steroids in the perioperative period. When patients are on steroid replacement doses equal to or greater than 20 mg per day of prednisone, it may be appropriate to administer additional glucocorticoid doses for no more than 2 perioperative days.

Adrenal insufficiency may be present in patients with a baseline serum cortisol less than 20 μg/dL. A rapid provocative test with synthetic adrenocorticotropic hormone may confirm the diagnosis. After a baseline serum cortisol level is drawn, 250 μg of cosynotropin is administered. At exactly 30 and 60 minutes following the dose of cosynotropin, serum cortisol levels are obtained. There should be an incremental increase in the cortisol level of between 7 and 10 μg/dL for each half hour. If the patient is below these levels, a diagnosis of adrenal insufficiency is made, and glucocorticoid and mineralocorticoid administration is then warranted. Mixed results are common, but the complication of performing major surgery on an adrenally insufficient patient is sudden or profound hypotension that is not responsive to fluid resuscitation.

Thyroid hormone abnormalities usually consist of previously undiagnosed thyroid abnormalities. Hypothyroidism and the so-called sick-euthyroid syndrome are more commonly recognized in the critical care setting. When surgical patients are not progressing satisfactorily in the perioperative period, screening for thyroid abnormalities should be performed. If the results show mild to moderate hypothyroidism, then thyroid replacement should begin immediately, and thyroid function studies should be monitored closely. All patients should be reassessed after the acute illness has subsided regarding the need for chronic thyroid replacement therapy.

**Problems with Thermoregulation**

**Hypothermia.** Hypothermia is defined as a core temperature less than 35°C (95°F) and is divided into subsets of mild (35°C–32°C [95.5°F–89.6°F]), moderate (32°C–28°C [89.6°F–82.4°F]), and severe (<28°C [<82.4°F]) hypothermia. Shivering, the body’s attempt to reverse the effects of hypothermia, occurs between 37°C and 31°C (98.6°F and 87.8°F), but ceases at temperatures below 31°C (87.8°F). Patients who are moderately hypothermic are at higher risk for complications than are those who are more profoundly hypothermic.
Hypothermia creates a coagulopathy that is related to platelet and clotting cascade enzyme dysfunction. This triad of metabolic acidosis, coagulopathy, and hypothermia is commonly found in long operative cases and in patients with blood dyscrasias. The enzymes that contribute to the clotting cascade and platelet activity are most efficient at normal body temperatures; therefore, all measures must be used to reduce heat loss intraoperatively.\(^\text{146}\)

The most common cardiac abnormality is the development of arrhythmias when body temperature drops below 35°C (95°F). Bradycardia occurs with temperatures below 30°C (86°F). It is well known that hypothermia may induce CO\(_2\) retention, resulting in respiratory acidosis. Renal dysfunction of hypothermia manifests itself as a paradoxical polyuria and is related to an increased glomerular filtration rate, as peripheral vascular constriction creates central shunting of blood. This is potentially perplexing in patients who are undergoing resuscitation for hemodynamic instability because the brisk urine output provides a false sense of an adequate intravascular fluid volume.

Induced peripheral hypothermia for hyperpyrexia due to infection (not to include neurologic or cardiac disease) is likely deleterious and does not appear to be beneficial. Placing cooling blankets on or under the patient or ice packs in the axillae or groin may be effective in cooling the skin, and when this occurs, a subsequent feedback loop triggers the hypothalamus to raise the internally regulated set point, thus raising core temperature even higher. This paradoxical reaction may be why those who feel the need to treat a fever in the ICU by cooling the skin and arguably the core have worse outcomes. Cooling core temperatures can be achieved reliably with catheter-directed therapy with commercially available devices. Whether this is a worthwhile practice or not may be controversial. Poor data exist in support of treating fevers lower than 42°C in any fashion.\(^\text{147-149}\)

Adult trauma patients who underwent induced hypothermia had poor outcomes in a recent investigation, and thus, this remains a procedure to be avoided. In a similar vein, pediatric patients who were induced did not show any improvement, and therefore, induced hypothermia is not recommended. Complications with induced hypothermia include, but are not limited to, hypokalemia, diuresis, DVT (due to catheter-related vein injury), arrhythmias, shivering, undiagnosed catheter-related bloodstream infection, and bacteremia.\(^\text{150-152}\)

Neurologic dysfunction is inconsistent in hypothermia, but a deterioration in reasoning and decision-making skills progresses as body temperature falls, and profound coma (and profound unresponsiveness) occurs as the temperature drops below 30°C (86°F). The diagnosis of hypothermia is important, so accurate measurement techniques are required to get a true core temperature.

Methods used to warm patients include warm air circulation over the patient and heated IV fluids, as well as more aggressive measures such as bilateral chest tubes with warm solution lavage, intraperitoneal rewarming lavage, and extracorporeal membrane oxygenation. A rate of temperature rise of 2°C to 4°C (3.6°F–7.2°F) per hour is considered adequate, but the most common complication for nonbypass rewarming is arrhythmia with ventricular arrest.

**Hyperthermia.** Hyperthermia is defined as a core temperature greater than 38.6°C (101.5°F) and has a host of etiologies (Table 12-19).\(^\text{145}\) Hyperthermia can be environmentally induced (e.g., summer heat with inability to dissipate heat or control exposure), iatrogenically induced (e.g., heat lamps and medications), endocrine in origin (e.g., thyrotoxicosis), or neurologically induced (i.e., hypothalamic dysfunction).

Malignant hyperthermia occurs intraoperatively after exposure to agents such as succinylcholine and some halothane-based inhalational anesthetics. The presentation is dramatic, with rapid onset of increased temperature, rigors, and myoglobinuria related to myonecrosis. Medications must be discontinued immediately and dantrolene administered (2.5 mg/kg every 5 minutes) until symptoms subside. Aggressive cooling methods are also implemented, such as an alcohol bath, or packing in ice. In cases of severe malignant hyperthermia, the mortality rate is nearly 30%.\(^\text{153,154}\)

Thyrotoxicosis can occur after surgery due to undiagnosed Graves’ disease. Hallmarks of the syndrome include hyperthermia (>40°C [104°F]), anxiety, copious diaphoresis, congestive heart failure (present in about one fourth of episodes), tachycardia (most commonly atrial fibrillation), and hypokalemia (in up to 50% of patients). The treatment of thyrotoxicosis includes glucocorticoids, propylthiouracil, β-blockade, and iodide (Lugol’s solution) delivered in an emergent fashion. As the name suggests, these patients are usually toxic and require supportive measures as well. Acetaminophen, the cooling modalities noted in the previous paragraph, and vasoactive agents often are indicated.

### Table 12-19

<table>
<thead>
<tr>
<th>Common causes of elevated temperature in surgical patients</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HYPERTHERMIA</strong></td>
</tr>
<tr>
<td>Environmental</td>
</tr>
<tr>
<td>Malignant hyperthermia</td>
</tr>
<tr>
<td>Neuroleptic malignant syndrome</td>
</tr>
<tr>
<td>Thyrotoxicosis</td>
</tr>
<tr>
<td>Pheochromocytoma</td>
</tr>
<tr>
<td>Carcinoid syndrome</td>
</tr>
<tr>
<td>Iatrogenic</td>
</tr>
<tr>
<td>Central/hypothalamic responses</td>
</tr>
<tr>
<td>Pulmonary embolism</td>
</tr>
<tr>
<td>Adrenal insufficiency</td>
</tr>
<tr>
<td><strong>HYPERPYREXIA</strong></td>
</tr>
<tr>
<td>Sepsis</td>
</tr>
<tr>
<td>Infection</td>
</tr>
<tr>
<td>Drug reaction</td>
</tr>
<tr>
<td>Transfusion reaction</td>
</tr>
<tr>
<td>Collagen disorders</td>
</tr>
<tr>
<td>Factitious syndrome</td>
</tr>
<tr>
<td>Neoplastic disorders</td>
</tr>
</tbody>
</table>

**REFERENCES**

Entries highlighted in bright blue are key references.

1. Makary MA, Daniel M. Medical error—the third leading cause of death in the US. *Brit Med J.* 2016;353:i2139. doi:10.1136/bmj.i2139


QUALITY, PATIENT SAFETY, ASSESSMENTS OF CARE, AND COMPLICATIONS

Chapter 12


INTRODUCTION

The Latin verb *monere*, which means “to warn, or advise” is the origin for the English word *monitor*. In modern medical practice, patients undergo monitoring to detect pathologic variations in physiologic parameters, providing advanced warning of impending deterioration in the status of one or more organ systems. The intended goal of this endeavor is to allow the clinician to take appropriate actions in a timely fashion to prevent or ameliorate the physiologic derangement. Furthermore, physiologic monitoring is used not only to warn, but also to titrate therapeutic interventions, such as fluid resuscitation or the infusion of vasoactive or inotropic drugs. The intensive care unit (ICU) and operating room are the two locations where the most advanced monitoring capabilities are routinely employed in the care of critically ill patients.

In the broadest sense, physiologic monitoring encompasses a spectrum of endeavors, ranging in complexity from the routine and intermittent measurement of the classic vital signs (i.e., temperature, heart rate, arterial blood pressure, and respiratory rate) to the continuous recording of the oxidation state of cytochrome oxidase, the terminal element in the mitochondrial electron transport chain. The ability to assess clinically relevant parameters of tissue and organ status and employ this knowledge to improve patient outcomes represents the “holy grail” of critical care medicine. Unfortunately, consensus is often lacking regarding the most appropriate parameters to monitor in order to achieve this goal. Furthermore, making an inappropriate therapeutic decision due to inaccurate physiologic data or misinterpretation of good data can lead to a worse outcome than having no data at all. Of the highest importance is the integration of physiologic data obtained from monitoring into a coherent and evidenced-based treatment plan. Current technologies available to assist the clinician in this endeavor are summarized in this chapter. Also presented is a brief look at emerging techniques that may soon enter into clinical practice.

In essence, the goal of hemodynamic monitoring is to ensure that the flow of oxygenated blood through the microcirculation is sufficient to support aerobic metabolism at the cellular level. In general, mammalian cells cannot store oxygen for subsequent use in oxidative metabolism, although a relatively tiny amount is stored in muscle tissue as oxidized myoglobin. Thus, aerobic synthesis of adenosine triphosphate (ATP), the energy “currency” of cells, requires the continuous delivery of oxygen by diffusion from hemoglobin in red blood cells to the oxidative machinery within mitochondria. Delivery of oxygen to mitochondria may be insufficient for several reasons. For example, cardiac output, hemoglobin concentration of blood, or the oxygen content of arterial blood each can be inadequate...
Key Points

1. The delivery of modern critical care is predicated on the ability to monitor a large number of physiologic variables and formulate evidenced-based therapeutic strategies to manage these variables. Technological advances in monitoring have at least a theoretical risk of exceeding our ability to understand the clinical implications of the derived information. This could result in the use of monitoring data to make inappropriate clinical decisions. Therefore, the implementation of any new monitoring technology must take into account the relevance and accuracy of the data obtained, the risks for independent reasons. Alternatively, despite adequate cardiac output, perfusion of capillary networks can be impaired as a consequence of dysregulation of arteriolar tone, microvascular thrombosis, or obstruction of nutritive vessels by sequestered leukocytes or platelets. Hemodynamic monitoring that does not take into account all of these factors will portray an incomplete and perhaps misleading picture of cellular physiology.

Under normal conditions when the supply of oxygen is plentiful, aerobic metabolism is determined by factors other than the availability of oxygen. These factors include the hormonal milieu and mechanical workload of contractile tissues. However, in pathologic circumstances when oxygen availability is inadequate, oxygen utilization (VO₂) becomes dependent upon oxygen delivery (DO₂). The relationship of VO₂ to DO₂ over a broad range of DO₂ values is commonly represented as two intersecting straight lines (Fig. 13-1). In the region of higher DO₂ values, the slope of the line is approximately equal to zero, indicating that VO₂ is largely independent of DO₂. In contrast, in the region of low DO₂ values, the slope of the line is nonzero and positive, indicating that VO₂ is supply-dependent. The region where the two lines intersect is called the point of critical oxygen delivery (DO₂crit), and represents the transition from supply-independent to supply-dependent oxygen uptake. Below a critical threshold of oxygen delivery, increased oxygen extraction cannot compensate for the delivery deficit; hence, oxygen consumption begins to decrease. The slope of the supply-dependent region of the plot reflects the maximal oxygen extraction capability of the vascular bed being evaluated.

The subsequent sections will describe the techniques and utility of monitoring various physiologic parameters.

**ARTERIAL BLOOD PRESSURE**

The pressure exerted by blood in the systemic arterial system, commonly referred to simply as “blood pressure,” is a cardinal parameter measured as part of the hemodynamic monitoring of patients. Extremes in blood pressure are either intrinsically deleterious or are indicative of a serious perturbation in normal physiology. Arterial blood pressure is a complex function of both cardiac output and vascular input impedance. Thus, inexperienced clinicians may assume that the presence of a normal blood pressure is evidence that cardiac output and tissue perfusion are adequate. This assumption is frequently incorrect and is the reason why some critically ill patients may benefit from forms of hemodynamic monitoring in addition to measurement of arterial pressure.

Blood pressure can be determined directly by measuring the pressure within the arterial lumen or indirectly using a cuff around an extremity. When the equipment is properly set up and calibrated, direct intra-arterial monitoring of blood pressure provides accurate and continuous data. Additionally, intra-arterial catheters provide a convenient way to obtain samples of blood for measurements of arterial blood gases and other laboratory studies. Despite these advantages, intra-arterial catheters are invasive devices and occasionally are associated with serious complications.

**Noninvasive Measurement of Arterial Blood Pressure**

Both manual and automated means for the noninvasive determination of blood pressure use an inflatable sphygmomanometer cuff to increase pressure around an extremity and to detect the presence or absence of arterial pulsations. Several methods exist for this purpose. The time-honored approach is the auscultation of the Korotkoff sounds, which are heard over an artery distal to the cuff as the cuff is deflated from a pressure higher than systolic pressure to one less than diastolic pressure. **Systolic pressure** is defined as the pressure in the cuff when tapping sounds are first audible. **Diastolic pressure** is the pressure in the cuff when audible pulsations first disappear.

---

Figure 13-1. Graphical representation of the relationship between oxygen utilization (VO₂) and oxygen delivery (DO₂). Under most normal physiologic conditions oxygen utilization does not depend on oxygen delivery, but below the critical value DO₂crit oxygen utilization decreases linearly as a function of oxygen delivery, rendering tissues susceptible to ischemic injury.
Another means for pulse detection when measuring blood pressure noninvasively depends upon the detection of oscillations in the pressure within the bladder of the cuff. This approach is simple, and unlike auscultation, can be performed even in a noisy environment (e.g., a busy emergency department or medical helicopter). Unfortunately, this approach is neither accurate nor reliable. Other methods, however, can be used to reliably detect the reappearance of a pulse distal to the cuff and thereby estimate systolic blood pressure. Two excellent and widely available approaches for pulse detection are use of a Doppler stethoscope (reappearance of the pulse produces an audible amplified signal) or a pulse oximeter (reappearance of the pulse is indicated by flashing of a light-emitting diode).

A number of automated devices are capable of repetitively measuring blood pressure noninvasively. Some of these devices measure pressure oscillations in the inflatable bladder encircling the extremity to detect arterial pulsations as pressure in the cuff is gradually lowered from greater than systolic to less than diastolic pressure. Other automated noninvasive devices use a piezoelectric crystal positioned over the brachial artery as a pulse detector. The accuracy of these devices is variable, and often dependent on the size mismatch between the arm circumference and the cuff size. If the cuff is too narrow (relative to the extremity), the measured pressure will be artifactualy elevated. Therefore, the width of the cuff should be approximately 40% of its circumference.

Another noninvasive approach for measuring blood pressure relies on a technique called photoplethysmography. This method is capable of providing continuous information, since systolic and diastolic blood pressures are recorded on a beat-to-beat basis. Photoplethysmography uses the transmission of infrared light to estimate the amount of hemoglobin (directly related to the volume of blood) in a finger placed under a servocommanded inflatable cuff. A feedback loop controlled by a microprocessor continually adjusts the pressure in the cuff to maintain the blood volume of the finger constant. Under these conditions, the pressure in the cuff reflects the pressure in the digital artery. The measurements obtained using photoplethysmography generally agree closely with those obtained by invasive monitoring of blood pressure. However, these readings may be less accurate in patients with hypotension or hyperthermia.

**Invasive Monitoring of Arterial Blood Pressure**

Direct and continuous monitoring of arterial pressure in critically ill patients may be performed by using fluid-filled tubing to connect an intra-arterial catheter to an external strain-gauge transducer. The signal generated by the transducer is electronically amplified and displayed as a continuous waveform by an oscilloscope or computerized display. Digital values for systolic and diastolic pressure also are displayed. Mean pressure, calculated by electronically averaging the amplitude of the pressure waveform, can also be displayed. The fidelity of the catheter-tubing-transducer system is determined by numerous factors, including the compliance of the tubing, the surface area of the transducer diaphragm, and the compliance of the diaphragm. If the system is underdamped, then the inertia of the system, which is a function of the mass of the fluid in the tubing and the mass of the diaphragm, causes overshoot of the points of maximum positive and negative displacement of the diaphragm during systole and diastole, respectively. Thus, in an underdamped system, systolic pressure will be overestimated and diastolic pressure will be underestimated. In an overdamped system, displacement of the diaphragm fails to track the rapidly changing pressure waveform, and systolic pressure will be underestimated and diastolic pressure will be overestimated. It is important to note that even in an underdamped or overdamped system, mean pressure will be accurately recorded, provided the system has been properly calibrated. For these reasons, when using direct measurement of intra-arterial pressure to monitor patients, clinicians should make clinical decisions based primarily on the measured mean arterial blood pressure.

The radial artery at the wrist is the site most commonly used for intra-arterial pressure monitoring. Other sites include the femoral and axillary artery. It is important to recognize, however, that measured arterial pressure is determined in part by the site where the pressure is monitored. Central (i.e., aortic) and peripheral (e.g., radial artery) pressures typically are different as a result of the impedance and inductance of the arterial tree. Systolic pressures typically are higher and diastolic pressures are lower in the periphery, whereas mean pressure is approximately the same in the aorta and more distal sites.

Distal ischemia is an uncommon complication of intra-arterial catheterization. The incidence of thrombosis is increased when larger-caliber catheters are employed and when catheters are left in place for an extended period of time. The incidence of thrombosis can be minimized by using a 20-gauge (or smaller) catheter in the radial artery and removing the catheter as soon as feasible. The risk of distal ischemic injury can be reduced by ensuring that adequate collateral flow is present prior to catheter insertion. At the wrist, adequate collateral flow can be documented by performing a modified version of the Allen test, wherein the artery to be cannulated is digitally compressed while using a Doppler stethoscope to listen for perfusion in the palmar arch vessels.

Another potential complication of intra-arterial monitoring is retrograde embolization of air bubbles or thrombi into the intracranial circulation. In order to minimize this risk care should be taken to avoid flushing arterial lines when air is present in the system, and only small volumes of fluid (less than 5 mL) should be employed for this purpose. Catheter-related infections can occur with any intravascular monitoring device. However, catheter-related bloodstream infection is a relatively uncommon complication of intra-arterial lines used for monitoring, occurring in 0.4% to 0.7% of catheterizations. The incidence increases with longer duration of arterial catheterization.

**ELECTROCARDIOGRAPHIC MONITORING**

The electrocardiogram (ECG) records the electrical activity associated with cardiac contraction by detecting voltages on the body surface. A standard 3-lead ECG is obtained by placing electrodes that correspond to the left arm (LA), right arm (RA), and left leg (LL). The limb leads are defined as lead I (LA-RA), lead II (LL-RA), and lead III (LL-LA). The ECG waveforms can be continuously displayed on a monitor, and the devices can be set to sound an alarm if an abnormality of rate or rhythm is detected. Continuous ECG monitoring is widely available and applied to critically ill and perioperative patients. Monitoring of the ECG waveform is essential in patients with acute coronary syndromes or blunt myocardial injury because dysrhythmias are the most common lethal complication. In patients with shock or sepsis, dysrhythmias can occur as a consequence of inadequate myocardial oxygen delivery or as a complication of vasoactive or inotropic drugs used to support blood pressure and cardiac...
output. Dysrhythmias can be detected by continuously monitoring the ECG tracing, and timely intervention may prevent serious complications. With appropriate computing hardware and software, continuous ST-segment analysis also can be performed to detect ischemia or infarction.

Additional information can be obtained from a 12-lead ECG, which is essential for patients with potential myocardial ischemia or to rule out cardiac complications in other acutely ill patients. Continuous monitoring of the 12-lead ECG may be beneficial in certain patient populations. In a study of 185 vascular surgical patients, continuous 12-lead ECG monitoring was able to detect transient myocardial ischemic episodes in 20.5% of the patients. This study demonstrated that the precordial lead V₄₅, which is not routinely monitored on a standard 3-lead ECG, is the most sensitive for detecting perioperative ischemia and infarction. To detect 95% of the ischemic episodes, two or more precordial leads were necessary. Furthermore, in a prospective observational study, 51 peripheral artery vascular surgery patients underwent ambulatory continuous 12-lead ECG monitoring in the postoperative setting. Ischemic load, defined as the area under the curve defined by ischemic ST-segment deviation and ischemic time, was shown to predict perioperative myocardial infarction with an area under the receiver operating characteristics curve of 0.87. Notably, ischemia was asymptomatic in 14 of the 17 identified patients, demonstrating value of this modality as a warning tool. Thus, continuous 12-lead ECG monitoring may provide greater sensitivity than 3-lead ECG for the detection of perioperative myocardial ischemia, and may become standard for monitoring high-risk surgical patients.

Currently, there is considerable interest in using computerized approaches to analyze ECG waveforms and patterns to uncover hidden information that can be used to predict sudden cardiac death or the development of serious dysrhythmias. ECG patterns of interest include repetitive changes in the morphology of the T-wave (T-wave alternans; TWA) and heart rate variability.

**ALGORITHMIC INTEGRATIVE MONITORING**

Integrated monitoring systems employ software that integrates vital signs to produce a single-parameter index that allows early detection of physiologic perturbations. The input variables include noninvasive measurements of heart rate, respiratory rate, blood pressure, SpO₂, and temperature. The software uses neural networking to develop a probabilistic model of normality, previously developed from a representative sample patient training set. Variance from this data set is used to evaluate the probability that the patient-derived vital signs are within the normal range. An abnormal index can occur while no single vital sign parameter is outside the range of normal if their combined patterns are consistent with known instability patterns. Employing such an integrated monitoring system in step-down unit patients has been shown to be a sensitive method to detect early physiologic abnormalities that may precede hemodynamic instability. This subsequently was demonstrated to reduce the amount of overall patient instability by facilitating earlier identification and appropriate intervention by the medical team.

The large expansion of the electronic medical record (EMR) is also driving the development of new algorithmic assessment tools for inpatient monitoring. The Rothman Index (RI) is a proprietary data analysis toolkit encompassing a total of 26 variables including vital signs, nursing assessments, laboratory test values, and cardiac rhythms and was developed to make use of the vast amount of data input into the EMR on a real-time basis to help provide a global assessment of patient status. In the initial derivation, Rothman and colleagues demonstrated concordance of the RI with the Modified Early Warning Score (MEWS) system, which is designed to alert medical teams to clinical deterioration that precedes cardiac or pulmonary arrest events. Subsequent publications evaluated performance of the RI in predicting both readmission to surgical ICUs in the postoperative setting as well as for rapid response team activations. Although more work is required to evaluate the broad applicability of the RI and similar measures, the evidence to date is compelling. Furthermore, as EMR interfaces become more sophisticated, other real-time data analysis software packages will likely be developed that provide further insight into the care of postsurgical patients.

**CARDIAC OUTPUT AND RELATED PARAMETERS**

Bedside catheterization of the pulmonary artery was introduced into clinical practice in the 1970s. Although the pulmonary artery catheter initially was used primarily to manage patients with cardiogenic shock and other acute cardiac diseases, indications for this form of invasive hemodynamic monitoring gradually expanded to encompass a wide variety of clinical conditions. Clearly, many clinicians believe that information valuable for the management of critically ill patients is afforded by having a pulmonary artery catheter (PAC) in place. However, unambiguous data in support of this view are scarce, and several studies suggest that bedside pulmonary artery catheterization may not benefit most critically ill patients and in fact may lead to some serious complications (see “Effect of Pulmonary Artery Catheterization on Outcome”).

**Determinants of Cardiac Performance**

Cardiac performance requires the integration of multiple mechanical and physiologic parameters of both the heart itself and of the circulatory system through which blood flows. The following sections discuss some of these factors, including preload, contractility, and afterload. A brief review of some of the graphical tools for evaluating cardiac physiology is demonstrated in Fig. 13-2.

**Preload.** Starling’s law of the heart states that the force of muscle contraction depends on the initial length of the cardiac fibers. Using terminology that derives from early experiments using isolated cardiac muscle preparations, preload is the stretch of ventricular myocardial tissue just prior to the next contraction. Strictly speaking, preload is determined by end-diastolic volume (EDV). In practice, EDV is challenging to measure precisely during the cardiac cycle, and so clinicians utilize the end-diastolic pressure (EDP) as a reasonable surrogate. For the right ventricle, central venous pressure (CVP) approximates right ventricular EDP. For the left ventricle, pulmonary artery occlusion pressure (PAOP), which is measured by transiently inflating a balloon at the end of a pressure monitoring catheter positioned in a small branch of the pulmonary artery, approximates left ventricular EDP. The presence of atrioventricular valvular stenosis may alter this relationship.

There are limits to the utilization of EDP as a surrogate for EDV when evaluating preload. For example, EDP is determined not only by volume but also by the diastolic compliance of the ventricular chamber. Ventricular compliance is altered by
Figure 13-2 A-D. Left ventricular pressure-volume loops constructed for various clinically relevant scenarios. For further information refer to the text. A. Standard left ventricular pressure-volume loop, with stroke volume, end systolic volume, and end diastolic volume highlighted for reference. Note the directionality of the pressure-volume loop, which is not annotated in the other figures for clarity. B-D. Demonstration of the effect of changing preload (B), contractility (C), or afterload (D) on the pressure-volume relationships in the left ventricle. Note the differences in stroke volume for various conditions, as well as the end-systolic volume and pressures, as these represent clinically significant parameters that govern patient care.

various pathologic conditions and pharmacologic agents. Furthermore, the relationship between EDP and true preload is not linear, but rather is exponential (Fig. 13-2A,B). This fact limits the utility of EDP as a surrogate marker at extremes of EDV.

**Contractility.** Contractility is defined as the inotropic state of the myocardium. Contractility is said to increase when the force of ventricular contraction increases at constant preload and afterload. Clinically, contractility is difficult to quantify because virtually all of the available measures are dependent to a certain degree on preload and afterload. If pressure-volume loops are constructed for each cardiac cycle, small changes in preload and/or afterload will result in shifts of the point defining the end of systole. These end-systolic points on the pressure-versus-volume diagram describe a straight line, known as the end-systolic pressure-volume line. A steeper slope of this line indicates greater contractility, as illustrated in Fig. 13-2C.

**Afterload.** Afterload is another term derived from in vitro experiments using isolated strips of cardiac muscle and is defined as the force resisting fiber shortening once systole begins. Defined specifically for the in vivo system, afterload is the resistance to the expulsion of blood from the heart chamber of interest, usually the left ventricle. Several factors comprise the in vivo correlate of ventricular afterload, including ventricular chamber geometry, intracavitary pressure generation, and the arterial impedance in the systemic circulation. Since these factors are difficult to assess clinically, afterload is commonly approximated by calculating systemic vascular resistance (SVR), defined as mean arterial pressure (MAP) divided by cardiac output (Fig. 13-2D).

**PLACEMENT OF THE PULMONARY ARTERY CATHETER**

In its simplest form, the PAC has four channels. One channel terminates in a balloon at the tip of the catheter. The proximal end of this channel is connected to a syringe to permit inflation of the balloon with air. Prior to insertion of the PAC, the integrity of the balloon should be verified by inflating it. In order to minimize the risk of vascular or ventricular perforation by the relatively inflexible catheter, it also is important to verify that the inflated balloon extends just beyond the tip of the device. A second channel in the catheter contains wires that are connected
to a thermistor located near the tip of the catheter. At the proximal end of the PAC, the wires terminate in a fitting that permits connection to appropriate hardware for the calculation of cardiac output using the thermodilution technique. The final two channels are used for pressure monitoring and the injection of the thermal indicator for determinations of cardiac output. One of these channels terminates at the tip of the catheter; the other terminates 20 cm proximal to the tip.

Placement of a PAC requires access to the central venous circulation. Such access can be obtained at a variety of sites, including the antecubital, femoral, jugular, and subclavian veins. Percutaneous placement through either the jugular or subclavian vein generally is preferred. Right internal jugular vein cannulation carries the lowest risk of complications, and the path of the catheter from this site into the right atrium is straight. In the event of inadvertent arterial puncture, local pressure is significantly more effective in controlling bleeding from the carotid artery as compared to the subclavian artery. Nevertheless, it is more difficult to keep occlusive dressings in place on the neck than in the subclavian fossa. Furthermore, the anatomic landmarks in the subclavian position are quite constant, even in patients with anasarca or massive obesity; the subclavian vein is always attached to the deep (concave) surface of the clavicle. In contrast, the appropriate landmarks to guide jugular venous cannulation are sometimes difficult to discern in obese or very edematous patients. However, ultrasonic guidance, which should be used routinely, has been shown to facilitate bedside jugular venipuncture.14

Cannulation of the vein is normally performed percutaneously, using the Seldinger technique. A small-bore needle is inserted through the skin and subcutaneous tissue into the vein. After documenting return of venous blood, a guidewire with a flexible tip is inserted through the needle into the vein, and the needle is withdrawn. A dilator/introducer sheath is passed over the wire, and the wire and the dilator are removed. The proximal terminus of the distal port of the PAC is connected through low-compliance tubing to a strain-gauge transducer, and the tubing-catheter system is flushed with fluid. While constantly observing the pressure tracing on a monitor screen, the PAC is advanced with the balloon deflated until respiratory excursions are observed. The balloon is then inflated, and the catheter advanced further, while monitoring pressures sequentially in the right atrium and right ventricle en route to the pulmonary artery. The pressure waveforms for the right atrium, right ventricle, and pulmonary artery are each characteristic (Fig. 13-3). The catheter is advanced out the pulmonary artery until a damped tracing indicative of the “wedged” position is obtained. The balloon is then deflated, taking care to ensure that a normal pulmonary arterial tracing is again observed on the monitor; leaving the balloon inflated can increase the risk of pulmonary infarction or perforation of the pulmonary artery. Unnecessary measurements of the pulmonary artery occlusion pressure are discouraged as rupture of the pulmonary artery may occur.

HEMODYNAMIC MEASUREMENTS

Even in its simplest embodiment, the PAC is capable of providing clinicians with a remarkable amount of information about the hemodynamic status of patients. Additional information may be obtained if various modifications of the standard PAC are employed. By combining data obtained through use of the PAC with results obtained by other means (i.e., blood hemoglobin concentration and oxyhemoglobin saturation), derived estimates of systemic oxygen transport and utilization can be calculated. Direct and derived parameters obtainable by bedside pulmonary arterial catheterization, along with several associated approximate normal ranges, are summarized in Table 13-1.
**Table 13-1**

Directly measured and derived hemodynamic data obtainable by bedside pulmonary artery catheterization, with normal associated ranges

<table>
<thead>
<tr>
<th>PARAMETER</th>
<th>NORMAL RANGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVP</td>
<td>0–6 mmHg</td>
</tr>
<tr>
<td>PAP</td>
<td>Varies</td>
</tr>
<tr>
<td>PAOP</td>
<td>6–12 mmHg</td>
</tr>
<tr>
<td>SVO₂ (intermittent or continuous)</td>
<td>65%–70%</td>
</tr>
<tr>
<td>QT (intermittent or continuous)</td>
<td>4–6 L/min</td>
</tr>
<tr>
<td>QT* (intermittent or continuous)</td>
<td>2.5–3.5 L·min⁻¹·m⁻²</td>
</tr>
<tr>
<td>RVEF</td>
<td>&gt;55%</td>
</tr>
<tr>
<td>SV</td>
<td>40–80 mL</td>
</tr>
<tr>
<td>SVR</td>
<td>800–1400 dyne·sec·cm⁻⁵</td>
</tr>
<tr>
<td>SVRI</td>
<td>1500–2400 dyne·sec·cm⁻³·m⁻²</td>
</tr>
<tr>
<td>PVR</td>
<td>100–150 dyne·sec·cm⁻⁵</td>
</tr>
<tr>
<td>PVRI</td>
<td>200–400 dyne·sec·cm⁻³·m⁻²</td>
</tr>
<tr>
<td>RVEDV</td>
<td>Variable</td>
</tr>
<tr>
<td>DO₂</td>
<td>400–660 mL·min⁻¹·m⁻²</td>
</tr>
<tr>
<td>VO₂</td>
<td>115–165 mL·min⁻¹·m⁻²</td>
</tr>
<tr>
<td>ER</td>
<td>Variable</td>
</tr>
<tr>
<td>Qₜ/QT</td>
<td>Variable</td>
</tr>
</tbody>
</table>

CVP = mean central venous pressure; DO₂ = systemic oxygen delivery; ER = systemic oxygen extraction ratio; PAOP = pulmonary artery occlusion (wedge) pressure; PAP = pulmonary artery pressure; PVR = pulmonary vascular resistance; PVRI = pulmonary vascular resistance index; QT = cardiac output; QT* = cardiac output indexed to body surface area (cardiac index); RVEF = right ventricular end-diastolic volume; RVEDV = right ventricular ejection fraction; SV = stroke volume; SVI = stroke volume index; SVO₂ = fractional mixed venous (pulmonary artery) hemoglobin saturation; SVR = systemic vascular resistance; SVRI = systemic vascular resistance index; VO₂ = systemic oxygen utilization.

Measurement of Cardiac Output by Thermodilution

Before the development of the PAC, determining cardiac output (QT) at the bedside required careful measurements of oxygen consumption (Fick method) or spectrophotometric determination of indocyanine green dye dilution curves. Measurements of QT using the thermodilution technique are simple and reasonably accurate. The measurements can be performed repetitively, and the principle is straightforward. If a bolus of an indicator is rapidly and thoroughly mixed with a moving fluid upstream from a detector, then the concentration of the indicator at the detector will increase sharply and then exponentially diminish back to zero. The area under the resulting time-concentration curve is a function of the volume of indicator injected and the flow rate of the moving stream of fluid. Larger volumes of indicator result in greater areas under the curve, and faster flow rates of the mixing fluid result in smaller areas under the curve. When QT is measured by thermodilution, the indicator is heat and the detector is a temperature-sensing thermistor at the distal end of the PAC. The relationship used for calculating QT is called the Stewart-Hamilton equation:

\[
Q_T = \frac{V K_1 K_2 (T_B - T_I)}{\int T_B(t) dt}
\]

where V is the volume of the indicator injected, TB is the temperature of blood (i.e., core body temperature), TI is the temperature of the indicator, K1 is a constant that is the function of the specific heats of blood and the indicator, K2 is an empirically derived constant that accounts for several factors (the dead space volume of the catheter, heat lost from the indicator as it traverses the catheter, and the injection rate of the indicator), and \(\int T_B(t) dt\) is the area under the time-temperature curve. In clinical practice, the Stewart-Hamilton equation is solved by a microprocessor.

Determination of cardiac output by the thermodilution method is generally quite accurate, although it tends to systematically overestimate QT at low values. Changes in blood temperature and QT during the respiratory cycle can influence the measurement. Therefore, results generally should be recorded as the mean of two or three determinations obtained at random points in the respiratory cycle. Using cold injectate widens the difference between TB and TI and thereby increases signal-to-noise ratio. Nevertheless, most authorities recommend using room temperature injectate (normal saline or 5% dextrose in water) to minimize errors resulting from warming of the fluid as it transferred from its reservoir to a syringe for injection.

Technologic innovations have been introduced that permit continuous measurement of QT by thermodilution. In this approach, thermal transients are not generated by injecting a bolus of a cold indicator, but rather by heating the blood with a tiny filament located on the PAC upstream from the thermistor. By correlating the amount of current supplied to the heating element with the downstream temperature of the blood, it is possible to estimate the average blood flow across the filament and thereby calculate QT. Based upon the results of several studies, continuous determinations of QT using this approach agree well with data generated by conventional measurements using bolus injections of a cold indicator. Information is lacking regarding the clinical value of being able to monitor QT continuously.

**Mixed Venous Oximetry**

The Fick equation can be written as

\[
Q_T = \frac{\text{VO}_2}{(\text{CaO}_2 - \text{CvO}_2)}
\]

where \(\text{C}_a\text{O}_2\) is the content of oxygen in arterial blood and \(\text{C}_v\text{O}_2\) is the content of oxygen in mixed venous blood. The oxygen content in both arterial and venous blood is a function of the hemoglobin concentration in the blood, the hemoglobin saturation, and the partial pressure of oxygen:

\[
\text{C}_a\text{O}_2 = (1.36 \times \text{Hgb} \times \frac{\text{S}_a\text{O}_2}{100}) + 0.0031 \times P_a\text{O}_2
\]

\[
\text{C}_v\text{O}_2 = (1.36 \times \text{Hgb} \times \frac{\text{S}_v\text{O}_2}{100})
\]

where \(\text{S}_a\text{O}_2\) is the fractional saturation of hemoglobin in either arterial or venous blood, Hgb is the concentration of hemoglobin...
in blood, and \( P_{a}O_2 \) is the partial pressure of oxygen in the arterial or venous blood. Under most circumstances the contribution of dissolved oxygen to both \( C_{a}O_2 \) and \( C_{v}O_2 \) is negligible, allowing the second portion of equation to be functionally eliminated (see previous equation). Given that, if the Fick equation is rearranged to the following:

\[
C_{a}O_2 = C_{v}O_2 - \frac{VO_2}{Q_T}
\]

Oxygen saturation can replace oxygen content, yielding the final clinically valuable equation:

\[
S_{O_2} = S_{a}O_2 - \frac{VO_2}{(Q_T \times Hgb \times 1.36)}
\]

where \( S_{O_2} \) is the fractional saturation of hemoglobin in mixed venous blood, \( S_{a}O_2 \) is the fractional saturation of hemoglobin in arterial blood, and \( Hgb \) is the concentration of hemoglobin in blood. Thus, it can be seen that \( S_{O_2} \) is a function of \( VO_2 \) (i.e., metabolic rate), \( Q_T \), \( S_{a}O_2 \), and \( Hgb \). Accordingly, subnormal values of \( S_{O_2} \) can be caused by a decrease in \( Q_T \) (due, for example, to heart failure or hypovolemia), a decrease in \( S_{a}O_2 \) (due, for example, to intrinsic pulmonary disease), a decrease in \( Hgb \) (i.e., anemia), or an increase in metabolic rate (due, for example, to seizures or fever). With a conventional PAC, measurements of \( S_{O_2} \) require aspirating a sample of blood from the distal (i.e., pulmonary arterial) port of the catheter and injecting the sample into a blood gas analyzer. Therefore, for practical purposes, measurements of \( S_{O_2} \) can be performed only intermittently.

By adding a fifth channel to the PAC, it is possible to monitor \( S_{O_2} \) continuously. The fifth channel contains two fiber-optic bundles, which are used to transmit and receive light of the appropriate wavelengths to permit measurements of hemoglobin saturation by reflectance spectrophotometry. Continuous \( S_{O_2} \) devices provide measurements of \( S_{O_2} \) that agree quite closely with those obtained by conventional analyses of blood aspirated from the pulmonary artery. Despite the theoretical value of being able to monitor \( S_{O_2} \) continuously, data are lacking to show that this capability favorably improves outcomes. In a prospective, observational study of 3265 patients undergoing cardiac surgery with either a standard PAC or a PAC with continuous \( S_{O_2} \) monitoring, the oximetric catheter was associated with fewer arterial blood gases and thermodilution cardiac output determinations but no difference in patient outcome. Since pulmonary artery catheters that permit continuous monitoring of \( S_{O_2} \) are more expensive than conventional PACs, the routine use of these devices cannot be recommended.

The saturation of oxygen in the right atrium or superior vena cava (\( S_{a}O_2 \)) correlates closely with \( S_{O_2} \) over a wide range of conditions, although the correlation between \( S_{a}O_2 \) and \( S_{O_2} \) has been questioned. Since measurement of \( S_{a}O_2 \) requires placement of a central venous catheter rather than a PAC, it is somewhat less invasive and easier to carry out. By using a central venous catheter equipped to permit fiber-optic monitoring of \( S_{a}O_2 \), it may be possible to titrate the resuscitation of patients with shock using a less invasive device than the PAC. The Surviving Sepsis Campaign international guidelines for the management of severe sepsis and septic shock recommends that during the first 6 hours of resuscitation, the goals of initial resuscitation of sepsis-induced hypoperfusion should include all of the following: CVP 8 to 12 mm Hg, MAP \( \geq 65 \) mm Hg, urine output \( \geq 0.5 \) mL/kg per hour, and \( S_{a}O_2 \) of 70% or \( S_{v}O_2 \) 65%.20

**EFFECT OF PULMONARY ARTERY CATHETERIZATION ON OUTCOME**

Despite initial enthusiasm for using the PAC in the management of critically ill patients, several studies have failed to show improved outcomes with their use. Connors and colleagues reported results of a major observational study evaluating the value of pulmonary artery catheterization in critically ill patients. These researchers compared two groups of patients: those who did undergo placement of a PAC during their first 24 hours of stay and those who did not. The investigators recognized that the value of their intended analysis was completely dependent on the robustness of their methodology for case-matching because sicker patients (i.e., those at greater risk of mortality based upon the severity of their illness) were presumably more likely to undergo pulmonary artery catheterization. Accordingly, the authors used sophisticated statistical methods for generating a cohort of study (i.e., PAC) patients, each one having a paired control matched carefully for severity of illness. Connors and associates concluded that placement of a pulmonary artery catheter during the first 24 hours of stay in an ICU is associated with a significant increase in the risk of mortality, even when statistical methods are used to account for severity of illness.

A number of prospective, randomized controlled trials of pulmonary artery catheterization are summarized in Table 13-2. The study by Pearson and associates was underpowered with only 226 patients enrolled. In addition, the attending anesthesiologists were permitted to exclude patients from the CVP group at their discretion; thus randomization was compromised. The study by Tuman and coworkers was large (1094 patients were enrolled), but different anesthesiologists were assigned to the different groups. Furthermore, 39 patients in the CVP group underwent placement of a PAC because of hemodynamic complications. All of the individual single-institution studies of vascular surgery patients were relatively underpowered, and all excluded at least certain categories of patients (e.g., those with a history of recent myocardial infarction).24,25

In the largest randomized controlled trial of the PAC, Sandham and associates randomized nearly 2000 American Society of Anesthesiologists (ASA) classes III and IV patients undergoing major thoracic, abdominal, or orthopedic surgery to placement of a PAC or CVP catheter. In the patients assigned to receive a PAC, physiologic goal-directed therapy was implemented by protocol. There were no differences in mortality at 30 days, 6 months, or 12 months between the two groups, and ICU length of stay was similar. There was a significantly higher rate of pulmonary emboli in the PAC group (0.9% vs. 0%). This study has been criticized because most of the patients enrolled were not in the highest risk category.

In the “PAC-Man” trial, a multicenter, randomized trial in 65 UK hospitals, over 1000 ICU patients were managed with or without a PAC. The specifics of the clinical management were then left up to the treating clinicians. There was no difference in hospital mortality between the 2 groups (with PAC 68% vs. without PAC 66%, \( P = 0.39 \)). However, a 9.5% complication rate was associated with the insertion or use of the PAC, although none of these complications were fatal. Clearly, these were critically ill patients, as noted by the high hospital mortality rates. Supporters of the PAC may cite methodology problems with this study, such as loose inclusion criteria and the lack of a defined treatment protocol.
Table 13-2
Summary of randomized, prospective clinical trials comparing pulmonary artery catheter (PAC) with central venous pressure (CVP) monitoring

<table>
<thead>
<tr>
<th>AUTHOR</th>
<th>STUDY POPULATION</th>
<th>GROUPS</th>
<th>OUTCOMES</th>
<th>STRENGTHS/WEAKNESSES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson et al(^{22})</td>
<td>“Low risk” patients undergoing cardiac or vascular surgery</td>
<td>CVP catheter (group 1); PAC (group 2); PAC with continuous Sv–O2 readout (group 3)</td>
<td>No differences among groups for mortality or length of ICU stay; significant differences in costs (group 1 &lt; group 2 &lt; group 3)</td>
<td>Underpowered (266 total patients enrolled); compromised randomization protocols</td>
</tr>
<tr>
<td>Tuman et al(^{23})</td>
<td>Cardiac surgical patients</td>
<td>PAC; CVP</td>
<td>No differences between groups for mortality, length of ICU stay, or significant noncardiac complications</td>
<td>Large trial (1094 patients); different anesthesiologists for different groups</td>
</tr>
<tr>
<td>Bender et al(^{24})</td>
<td>Vascular surgery patients</td>
<td>PAC; CVP</td>
<td>No differences between groups for mortality, length of ICU stay, or length of hospital stay</td>
<td>Relatively underpowered</td>
</tr>
<tr>
<td>Valentine et al(^{25})</td>
<td>Aortic surgery patients</td>
<td>PAC + hemodynamic optimization in ICU night before surgery; CVP</td>
<td>No difference between groups for mortality or length of ICU stay; significantly higher incidence of postoperative complications in PAC group</td>
<td>Relatively underpowered</td>
</tr>
<tr>
<td>Sandham et al(^{26})</td>
<td>“High risk” major surgery</td>
<td>PAC; CVP</td>
<td>No differences between groups for mortality, length of ICU stay; increased incidence of pulmonary embolism in PAC group</td>
<td>Largest RCT of PAC utilization; commonly criticized for smaller number of highest risk category patients</td>
</tr>
<tr>
<td>Harvey et al(^{27})</td>
<td>Medical and surgical ICU patients</td>
<td>PAC vs no PAC, with option for alternative CO measuring device in non-PAC group</td>
<td>No difference in hospital mortality between the 2 groups, increased incidence of complications in the PAC group</td>
<td>Loose inclusion criteria with lack of a defined treatment protocol for use of PAC data</td>
</tr>
<tr>
<td>Binanay et al(^{29})</td>
<td>Patients with CHF</td>
<td>PAC vs no PAC</td>
<td>No difference in hospital mortality between the groups, increased incidence of adverse events in the PAC group</td>
<td>No formal treatment protocol for PAC-driven therapy</td>
</tr>
<tr>
<td>Wheeler et al(^{30})</td>
<td>Patients with ALI</td>
<td>PAC vs CVC with a fluid and inotropic management protocol</td>
<td>No difference in ICU or hospital mortality, or incidence of organ failure between the groups; increased incidence of adverse events in the PAC group</td>
<td></td>
</tr>
</tbody>
</table>

ALI = acute lung injury; CHF = congestive heart failure; CO = cardiac output; CVC = central venous catheter; ICU = intensive care unit; PAC = pulmonary artery catheter; Sv–O2 = fractional mixed venous (pulmonary artery) hemoglobin saturation.

A meta-analysis of 13 randomized studies of the PAC that included over 5000 patients was published in 2005.\(^{28}\) A broad spectrum of critically ill patients was included in these heterogeneous trials, and the hemodynamic goals and treatment strategies varied. While the use of the PAC was associated with an increased use of inotropes and vasodilators, there were no differences in mortality or hospital length of stay between the patients managed with a PAC and those managed without a PAC. The ESCAPE trial (which was one of the studies included in the previous meta-analysis)\(^{29}\) evaluated 433 patients with severe or recurrent congestive heart failure (CHF) admitted to the ICU. Patients were randomized to management by clinical assessment and a PAC or clinical assessment without a PAC. The goal in both groups was resolution of CHF, with additional PAC targets of a pulmonary capillary occlusion pressure of 15 mmHg and a right atrial pressure of 8 mmHg. There was no formal treatment protocol, but inotropic support was discouraged. Substantial reduction in symptoms, jugular venous pressure, and edema was noted in both groups. There was no significant difference in the primary end point of days alive and out of the hospital during the first 6 months, or hospital mortality (PAC 10% vs without PAC 9%). Adverse events
were more common among patients in the PAC group (21.9% vs 11.5%; *P* = 0.04).

Finally, the Fluids and Catheters Treatment Trial (FACTTT) conducted by the Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network was published in 2006. The risks and benefits of PAC compared with central venous catheters (CVC) were evaluated in 1000 patients with acute lung injury. Patients were randomly assigned to receive either a PAC or a CVC to guide management for 7 days via an explicit protocol. Patients also were randomly assigned to a conservative or liberal fluid strategy in a 2 × 2 factorial design (outcomes based on the fluid management strategy were published separately). Mortality during the first 60 days was similar in the PAC and CVC groups (27% and 26%, respectively; *P* = .69). The duration of mechanical ventilation and ICU length of stay also were not influenced by the type of catheter used. The type of catheter employed did not affect the incidence of shock, respiratory or renal failure, ventilator settings, or requirement for hemodialysis or vasopressors. There was a 1% rate of crossover from CVC-guided therapy to PAC-guided therapy. The catheter used did not affect the administration of fluids or diuretics, and the fluid balance was similar in the two groups. The PAC group had approximately twice as many catheter-related adverse events (mainly arrhythmias).

Few subjects in critical care medicine have historically generated more emotional responses among experts in the field than the use of the PAC. As these studies demonstrate, it is not possible to show that therapy directed by use of the PAC saves lives when it is evaluated in a large population of patients. Certainly, given the available evidence, routine use of the PAC cannot be justified.

Whether selective use of the device in a few relatively uncommon clinical situations is warranted or valuable remains a controversial issue. Consequently, a marked decline in the use of the PAC from 5.66 per 1000 medical admissions in 1993 to 1.99 per 1000 medical admissions in 2004 has been seen. Based upon the results and exclusion criteria in these prospective randomized trials, reasonable criteria for perioperative monitoring without use of a PAC are presented in Table 13-3.

One of the reasons for using a PAC to monitor critically ill patients is to optimize cardiac output and systemic oxygen delivery. Defining what constitutes the optimum cardiac output, however, has proven to be difficult. A number of randomized trials evaluating the effect on outcome of goal-directed as compared to conventional hemodynamic resuscitation have been published. Some studies provide support for the notion that interventions designed to achieve supraphysiologic goals for DO₂, VO₂, and Qₜ improve outcome. However, other published studies do not support this view, and a meta-analysis concluded that interventions designed to achieve supraphysiologic goals for oxygen transport do not significantly reduce mortality rates in critically ill patients. At this time, supraphysiologic resuscitation of patients in shock cannot be endorsed.

There is no simple explanation for the apparent lack of effectiveness of pulmonary artery catheterization, although several concurrent possibilities exist. First, even though bedside pulmonary artery catheterization is quite safe, the procedure is associated with a finite incidence of serious complications, including ventricular arrhythmias, catheter-related sepsis, central venous thrombosis, pulmonary arterial perforation, and pulmonary embolism. The adverse effects of these complications on outcome may equal or even outweigh any benefits associated with using a PAC to guide therapy. Second, the data generated by the PAC may be inaccurate, leading to inappropriate therapeutic interventions. Third, the measurements, even if accurate, are often misinterpreted. Furthermore, the current state of understanding is primitive when it comes to deciding what is the best management for certain hemodynamic disturbances, particularly those associated with sepsis or septic shock. Taking all of this into consideration, it may be that interventions prompted by measurements obtained with a PAC are actually harmful to patients. As a result, the marginal benefit now available by placing a PAC may be quite small. Less invasive modalities are available that may provide clinically useful hemodynamic information.

It may be true that aggressive hemodynamic resuscitation of patients, guided by various forms of monitoring, is valuable only during certain critical periods, such as the first few hours after presentation with septic shock or during surgery. For example, Rivers and colleagues reported that survival of patients with septic shock is significantly improved when resuscitation in the emergency department is guided by a protocol that seeks to keep ScvO₂ greater than 70%. Similarly, a study using an ultrasound-based device (see “Doppler Ultrasonography”) to assess cardiac filling and SV showed that maximizing SV intraoperatively results in fewer postoperative complications and shorter hospital length of stay.

### Table 13-3

**Suggested criteria for perioperative monitoring without use of a pulmonary artery catheter in patients undergoing cardiac or major vascular surgical procedures**

<table>
<thead>
<tr>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>No anticipated need for suprarenal or supraceliac aortic cross-clamping</td>
</tr>
<tr>
<td>No history of myocardial infarction during 3 months prior to operation</td>
</tr>
<tr>
<td>No history of poorly compensated congestive heart failure</td>
</tr>
<tr>
<td>No history of coronary artery bypass graft surgery during 6 weeks prior to operation</td>
</tr>
<tr>
<td>No history of ongoing symptomatic mitral or aortic valvular heart disease</td>
</tr>
<tr>
<td>No history of ongoing unstable angina pectoris</td>
</tr>
</tbody>
</table>

**MINIMALLY INVASIVE ALTERNATIVES TO THE PULMONARY ARTERY CATHETER**

Because of the cost, risks, and questionable benefit associated with bedside pulmonary artery catheterization, there has been interest in the development of practical means for less invasive monitoring of hemodynamic parameters. Several approaches have been developed that have achieved variable degrees of success. None of these methods render the standard thermodilution technique of the pulmonary artery catheter obsolete. However, these strategies may contribute to improvements in the hemodynamic monitoring of critically ill patients.

**Transpulmonary Thermodilution**

In the standard PAC thermodilution technique, measurements rely on the detection of temperature changes in a relatively small area from the injection port to the thermistor on the same catheter. In contrast, the transpulmonary thermodilution (TPTD) technique measures temperature changes from cold
bolus solution injected centrally, then measured using an arterial thermistor on a special arterial line, generally placed in the femoral artery. Both standard PAC thermodilution and TPTD make use of the Stewart-Hamilton equation to subsequently calculate cardiac output. Studies have demonstrated that this technique provides comparable estimates of cardiac output when compared to routine PAC thermodilution and can accurately detect changes in cardiac output as small as 12%. However, due to the large blood circuit between the central injection point and the thermistor, data can be challenging to interpret in certain pathophysiologic conditions (e.g., in pulmonary edema, as excess lung water serves as a temperature sink). On the other hand, thoughtful application of TPTD data allows clinicians access to several additional variables that the traditional PAC does not provide, such as estimation of the global end-diastolic volume (GEDV) and the extravascular lung water volume (EVLW). While these variables are of scientific interest, they are not yet in widespread use, and further studies are required to determine their utility. However, TPTD does currently play a prominent role in the real-time calibration of pulse contour analysis, described in greater detail later in this chapter.

**Doppler Ultrasoundography**

When ultrasonic sound waves are reflected by moving erythrocytes in the bloodstream, the frequency of the reflected signal is increased or decreased, depending on whether the cells are moving toward or away from the ultrasonic source. This change in frequency is called the Doppler shift, and its magnitude is determined by the velocity of the moving red blood cells. Therefore, measurements of the Doppler shift can be used to calculate red blood cell velocity. With knowledge of both the cross-sectional area of a vessel and the mean red blood cell velocity of the blood flowing through it, one can calculate blood flow rate. If the vessel in question is the aorta, then Q_T can be calculated as:

\[
Q_T = HR \times A \times \int V(t) \, dt
\]

where A is the cross-sectional area of the aorta and \( \int V(t) \, dt \) is the red blood cell velocity integrated over the cardiac cycle.

Two approaches have been developed for using Doppler ultrasonography to estimate Q_T. The first approach uses an ultrasonic transducer, which is manually positioned in the suprasternal notch and focused on the root of the aorta. Aortic cross-sectional area can be estimated using a nomogram, which factors in age, height, and weight, back-calculated if an independent measure of Q_T is available, or by using two-dimensional transthoracic or transesophageal ultrasonography. While this approach is completely noninvasive, it requires a highly-skilled operator in order to obtain meaningful results and is labor-intensive. Moreover, unless Q_T measured using thermodilution is used to back-calculate aortic diameter, accuracy using the suprasternal notch approach is not acceptable. Accordingly, the method is useful only for obtaining very intermittent estimates of Q_T, and it has not been widely adopted by clinicians.

A more promising, albeit more invasive, approach has been introduced. In this method blood flow velocity is continuously monitored in the descending thoracic aorta using a continuous-wave Doppler transducer introduced into the esophagus. The probe is connected to a monitor which continuously displays the blood flow velocity profile in the descending aorta as well as the calculated Q_T. In order to maximize the accuracy of the device, the probe position must be adjusted to obtain the peak velocity in the aorta. In order to transform blood flow in the descending aorta into Q_T, a correction factor is applied that is based on the assumption that only 70% of the flow at the root of the aorta is still present in the descending thoracic aorta. Aortic cross-sectional area is estimated using a nomogram based on the patient’s age, weight, and height. Results using these methods appear to be reasonably accurate across a broad spectrum of patients. A meta-analysis of the available data shows a good correlation between cardiac output estimates obtained by transesophageal Doppler and PAC in critically ill patients. The ultrasonic device also calculates a derived parameter termed flow time corrected (FTc), which is the systolic flow time in the descending aorta corrected for heart rate. FTc is a function of preload, contractility, and vascular input impedance. Although it is not a pure measure of preload, Doppler-based estimates of SV and FTc have been used successfully to guide volume resuscitation in high-risk surgical patients undergoing major operations.

**Impedance Cardiography**

The impedance to flow of alternating electrical current in regions of the body is commonly called bioimpedance. In the thorax, changes in the volume and velocity of blood in the thoracic aorta lead to detectable changes in bioimpedance. The first derivative of the oscillating component of thoracic bioimpedance (dZ/dt) is linearly related to aortic blood flow. On the basis of this relationship, empirically derived formulas have been developed to estimate SV, and subsequently Q_T, noninvasively. This methodology is called impedance cardiography. The approach is attractive because it is noninvasive, provides a continuous readout of Q_T, and does not require extensive training. Despite these advantages, measurements of Q_T obtained by impedance cardiography are not sufficiently reliable to be used for clinical decision making and have poor correlation with thermodilution.

Because of the limitations of bioimpedance devices, a newer approach for processing the impedance signal was developed and commercialized. This approach is based on the recognition that the impedance signal has two components: amplitude and phase. Whereas the amplitude of the thoracic impedance signal is determined by all of the components of the thoracic cavity (bone, blood, muscle, and other soft tissues), phase shifts are determined entirely by pulsatile flow. The vast majority of pulsatile flow is related to blood moving within the aorta. Therefore, the “bioreactance” signal correlates closely with aortic flow, and cardiac output determined using this approach agrees closely with cardiac output measured using conventional indicator dilution techniques.

**Pulse Contour Analysis**

Another method for determining cardiac output is an approach called pulse contour analysis for estimating SV on a beat-to-beat basis. The mechanical properties of the arterial tree and SV determine the shape of the arterial pulse waveform. The pulse contour method of estimating Q_T uses the arterial pressure waveform as an input for a model of the systemic circulation in order to determine beat-to-beat flow through the circulatory system. The parameters of resistance, compliance, and impedance are initially estimated based on the patient’s age and sex and can be subsequently refined by using a reference standard measurement of Q_T. The reference standard estimation of Q_T is obtained periodically using the indicator dilution approach by injecting the indicator into a central venous catheter and
detecting the transient increase in indicator concentration in the blood using an arterial catheter. In one commercially available embodiment of this approach, the lithium ion (Li+) is the indicator used for the periodic calibrations of the device. The lithium carbonate indicator can be injected into a peripheral vein, and the doses do not exert pharmacologically relevant effects in adult patients. The Li+ indicator dilution method has shown to be at least as reliable as other thermodilution methods over a broad range of CO in a variety of patients. An improved method uses the indicator for calibration, via TPTD approaches as described previously. When the pulse contour analysis is combined with intermittent TPTD in this fashion, the continuous data provided by contour analysis is more precise than TPTD alone.

Measurements of QT based on pulse contour monitoring using these two approaches are comparable in accuracy to standard pulmonary artery catheter (PAC)-thermodilution methods, but they are less invasive because transcardiac catheterization is not needed. Using online pressure waveform analysis, the computerized algorithms can calculate SV, QT, SVR, and an estimate of myocardial contractility, the rate of rise of the arterial systolic pressure (dP/dT). The use of pulse contour analysis has been applied using noninvasive photoplethysmographic measurements of arterial pressure. However, the accuracy of this technique has been questioned, and its clinical utility remains to be determined.

One commercially available device that can be used for estimating cardiac output does not require external calibration. Instead, the relationship between pulse pressure and stroke volume is determined using a proprietary algorithm that uses biometric data, such as age, gender and height, as inputs. Although this methodology is gaining fairly wide acceptance in critical care medicine, reported accuracy (in comparison to “gold standard” approaches) is not very good.

Partial Carbon Dioxide Rebreathing

Partial carbon dioxide (CO₂) rebreathing uses the Fick principle to estimate Qₜ noninvasively. By intermittently altering the dead space within the ventilator circuit via a rebreathing valve, changes in CO₂ production (VCO₂) and end-tidal CO₂ (ETCO₂) are used to determine cardiac output using a modified Fick equation:

\[
Q_T = \frac{\Delta V_{CO_2}}{\Delta ET_{CO_2}}
\]

Commercially available devices use this Fick principle to calculate Qₜ using intermittent partial CO₂ rebreathing through a disposable rebreathing loop. These devices consist of a CO₂ sensor based on infrared light absorption, an airflow sensor, and a pulse oximeter. Changes in intrapulmonary shunt and hemodynamic instability impair the accuracy of Qₜ estimated by partial CO₂ rebreathing. Continuous inline pulse oximetry and inspiration of inspired O₂ (FiO₂) are used to estimate shunt fraction to correct Qₜ.

Some studies of the partial CO₂ rebreathing approach suggest that this technique is not as accurate as thermodilution, the gold standard for measuring Qₜ. However, other studies suggest that the partial CO₂ rebreathing method for determination of Qₜ compares favorably to measurements made using a PAC in critically ill patients.

Transesophageal Echocardiography

Transesophageal echocardiography (TEE) has made the transition from operating room to intensive care unit. TEE requires that the patient be sedated and usually intubated for airway protection. Using this powerful technology, global assessments of LV and RV function can be made, including determinations of ventricular volume, EF, and QT. Segmental wall motion abnormalities, pericardial effusions, and tamponade can be readily identified with TEE. Doppler techniques allow estimation of atrial filling pressures. The technique is somewhat cumbersome and requires considerable training and skill in order to obtain reliable results. Recently, a TEE probe has been introduced into practice that is small enough in diameter that it can be left in place for as long as 72 hours. While only limited data are currently available with this probe, it seems like it will be a useful cardiac monitoring tool for use in selected, complex patients.

Assessing Preload Responsiveness

Although pulse contour analysis or partial CO₂ rebreathing may be able to provide estimates of SV and QT, these approaches alone can offer little or no information about the adequacy of preload. Thus, if QT is low, some other means must be employed to estimate preload. Many clinicians assess the adequacy of cardiac preload by determining CVP or PAOP. However, neither CVP nor PAOP correlate well with the true parameter of interest, left ventricular end-diastolic volume (LVEDV). Extremely high or low CVP or PAOP results are informative, but readings in a large middle zone (i.e., 5 to 20 mmHg) are less useful. Furthermore, changes in CVP or PAOP fail to correlate well with changes in stroke volume. Echocardiography can be used to estimate LVEDV, but this approach is dependent on the skill and training of the individual using it, and isolated measurements of LVEDV fail to predict the hemodynamic response to alterations in preload.

When intrathoracic pressure increases during the application of positive airway pressure in mechanically ventilated patients, venous return decreases, and as a consequence, left ventricular stroke volume (LVSV) also decreases. Therefore, pulse pressure variation (PPV) during a positive pressure episode can be used to predict the responsiveness of cardiac output to changes in preload. PPV is defined as the difference between the maximal pulse pressure and the minimum pulse pressure divided by the average of these two pressures (Fig. 13-4). This approach has validated this by comparing PPV, CVP, PAOP, and systolic pressure variation as predictors of preload responsiveness in a cohort of critically ill patients. Patients were classified as being “preload responsive” if their cardiac index increased by at least 15% after rapid infusion of a standard volume of intravenous fluid. Receiver-operating characteristic (ROC) curves demonstrated that PPV was the best predictor of preload responsiveness. Although atrial arrhythmias can interfere with the usefulness of this technique, PPV remains a useful approach for assessing preload responsiveness in most patients because of its simplicity and reliability.

Near-Infrared Spectroscopic Measurement of Tissue Hemoglobin Oxygen Saturation

Near-infrared spectroscopy (NIRS) allows continuous, noninvasive measurement of tissue hemoglobin oxygen saturation (StO₂) using near-infrared wave lengths of light (700–1000 nm). This technology is based on Beer’s law, which states that the transmission of light through a solution with a dissolved
solute decreases exponentially as the concentration of the solute increases. In mammalian tissue, three compounds change their absorption pattern when oxygenated: cytochrome aa3, myoglobin, and hemoglobin. Because of the distinct absorption spectra of oxyhemoglobin and deoxyhemoglobin, Beer’s law can be used to detect their relative concentrations within tissue. Thus, the relative concentrations of the types of hemoglobin can be determined by measuring the change in light intensity as it passes through the tissue. Since about 20% of blood volume is intra-arterial and the StO2 measurements are taken without regard to systole or diastole, spectroscopic measurements are primarily indicative of the venous oxyhemoglobin concentration.

NIRS has been evaluated to assess the severity of traumatic shock in animal models and in trauma patients. Studies have shown that peripheral muscle StO2, as determined by NIRS, is as accurate as other endpoints of resuscitation (i.e., base deficit, mixed venous oxygen saturation) in a porcine model of hemorrhagic shock. Continuously measured StO2 has been evaluated in blunt trauma patients as a predictor of the development of multiple organ dysfunction syndrome (MODS) and mortality. 383 patients were prospectively studied at seven level I trauma centers. StO2 was monitored for 24 hours after admission along with vital signs and other endpoints of resuscitation such as base deficit (BD). Minimum StO2 (using a minimum StO2 ≤75% as a cutoff) had a similar sensitivity and specificity in predicting the development of MODS as BD ≥6 mEq/L. StO2 and BD were also comparable in predicting mortality. Thus, NIRS-derived muscle StO2 measurements perform similarly to BD in identifying poor perfusion and predicting the development of MODS or death after severe torso trauma, yet have the additional advantages of being continuous and noninvasive. Ongoing prospective studies will help determine the clinical utility of continuous monitoring of StO2 in clinical scenarios such as trauma, hemorrhagic shock, sepsis, etc.

**RESPIRATORY MONITORING**

The ability to monitor various parameters of respiratory function is of utmost importance in critically ill patients. Many of these patients require mechanical ventilation. Monitoring of their respiratory physiology is necessary to assess the adequacy of oxygenation and ventilation, guide weaning and liberation from mechanical ventilation, and detect adverse events associated with respiratory failure and mechanical ventilation. These parameters include gas exchange, neuromuscular activity, respiratory mechanics, and patient effort.

**Arterial Blood Gases**

Blood gas analysis may provide useful information when caring for patients with respiratory failure. However, even in the absence of respiratory failure or the need for mechanical ventilation, blood gas determinations also can be valuable to detect alterations in acid-base balance due to low Qs, sepsis, renal failure, severe trauma, medication or drug overdose, or altered mental status. Arterial blood can be analyzed for pH, Po2, Pco2, HCO3– concentration and calculated base deficit. When indicated, carboxyhemoglobin and methemoglobin levels also can be measured. In recent years, efforts have been made to decrease the unnecessary use of arterial blood gas analysis. Serial arterial blood gas determinations are not necessary for routine weaning from mechanical ventilation in the majority of postoperative patients.

Most bedside blood gas analyses still involve removal of an aliquot of blood from the patient, although continuous bedside arterial blood gas determinations are now possible without sampling via an indwelling arterial catheter that contains a biosensor. In studies comparing the accuracy of continuous arterial blood gas and pH monitoring with a conventional laboratory blood gas analyzer, excellent agreement between the two methods has been demonstrated. Continuous monitoring can reduce the volume of blood loss due to phlebotomy and dramatically decrease the time necessary to obtain blood gas results. Continuous monitoring, however, is expensive and is not widely employed.

**Determinants of Oxygen Delivery**

The primary goal of the cardiovascular and respiratory systems is to deliver oxygenated blood to the tissues. DO2 is dependent to a greater degree on the oxygen saturation of hemoglobin (Hgb) in arterial blood (Sao2) than on the partial pressure of oxygen in arterial blood (Pao2). DO2 also is dependent on Qs and Hgb. As discussed earlier and illustrated mathematically by previous equations, the dissolved oxygen in blood makes only a negligible contribution to DO2. Sao2 in mechanically ventilated patients depends on the mean airway pressure, the fraction of inspired oxygen (Fio2), and S O2. Thus, when Sao2 is low, the clinician has only a limited number of ways to improve this parameter. The clinician can increase mean airway pressure by increasing positive-end expiratory pressure (PEEP) or inspiratory time. Fio2 can be increased to a maximum of 1.0 by decreasing the amount of room air mixed with the oxygen supplied to the ventilator. S O2 can be increased by increasing Hgb
Peak and Plateau Airway Pressure

Airway pressures are routinely monitored in mechanically ventilated patients. The peak airway pressure measured at the end of inspiration (P\text{peak}) is a function of the tidal volume, the resistance of the airways, lung/chest wall compliance, and peak inspiratory flow. The airway pressure measured at the end of inspiration when the inhaled volume is held in the lungs by briefly closing the expiratory valve is termed the plateau airway pressure (P\text{plateau}). As a static parameter, plateau airway pressure is independent of the airway resistance and peak airway flow and is related to the lung/chest wall compliance and delivered tidal volume. Mechanical ventilators monitor P\text{peak} with each breath and can be set to trigger an alarm if the P\text{peak} exceeds a predetermined threshold. P\text{plateau} is not measured routinely with each delivered tidal volume but rather is measured intermittently by setting the ventilator to close the exhalation circuit briefly at the end of inspiration and record the airway pressure when airflow is zero.

If both P\text{peak} and P\text{plateau} are increased (and tidal volume is not excessive), then the problem is a decrease in the compliance in the lung/chest wall unit. Common causes of this problem include pneumothorax, hemotorax, lobar atelectasis, pulmonary edema, pneumonia, acute respiratory distress syndrome (ARDS), active contraction of the chest wall or diaphragmatic muscles, abdominal distention, and intrinsic PEEP, such as occurs in patients with bronchospasm and insufficient expiratory times. When P\text{peak} is increased but P\text{plateau} is relatively normal, the primary problem is an increase in airway resistance, such as occurs with bronchospasm, use of a small-caliber endotracheal tube, or kinking or obstruction of the endotracheal tube. A low P\text{peak} also should trigger an alarm, as it suggests a discontinuity in the airway circuit involving the patient and the ventilator. These scenarios are outlined in Table 13-4.

Ventilator-induced lung injury (VILI) is now an established clinical entity of great relevance to the care of critically ill patients. Excessive airway pressure and tidal volume adversely affect pulmonary and possibly systemic responses to critical illness. Subjecting the lung parenchyma to excessive pressure, known as barotrauma, can result in parenchymal lung injury, diffuse alveolar damage similar to ARDS, and pneumothorax, and can impair venous return and therefore limit cardiac output. Lung-protective ventilation strategies have been developed to prevent the development of VILI and improve patient outcomes.

### Table 13-4

<table>
<thead>
<tr>
<th>Condition</th>
<th>P\text{peak}</th>
<th>P\text{plateau}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decreased compliance of the system (ARDS, abdominal distention, intrinsic PEEP)</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>Increase in airway resistance (bronchospasm, endotracheal tube obstruction/kinking, or small-caliber endotracheal tube)</td>
<td>↑</td>
<td>normal</td>
</tr>
<tr>
<td>Disconnected circuit</td>
<td>↓</td>
<td>↓</td>
</tr>
</tbody>
</table>

In a large, multicenter, randomized trial of patients with ARDS from a variety of etiologies, limiting plateau airway pressure to less than 30 cm H\text{2}O and tidal volume to less than 6 mL/kg of ideal body weight reduced 28-day mortality by 22% relative to a ventilator strategy that used a tidal volume of 12 mL/kg. For this reason, monitoring of plateau pressure and using a low tidal volume strategy in patients with ARDS is now the standard of care. Recent data also suggest that a lung-protective ventilation strategy is associated with improved clinical outcomes in ventilated patients without ARDS. Importantly, this strategy also has been shown to have benefit for high-risk patients undergoing general anesthesia for surgical procedures, leading to a reduced overall rate of pulmonary complications in the perioperative period as well as a reduced length of stay following surgery.

Pulse Oximetry

The pulse oximeter is a microprocessor-based device that integrates oximetry and plethysmography to provide continuous noninvasive monitoring of the oxygen saturation of arterial blood (Sao\text{2}). It is considered one of the most important and useful technologic advances in patient monitoring. Continuous, noninvasive monitoring of arterial oxygen saturation is possible using light-emitting diodes and sensors placed on the skin. Pulse oximetry employs two wavelengths of light (i.e., 660 nm and 940 nm) to analyze the pulsatile component of blood flow between the light source and sensor. Because oxyhemoglobin and deoxyhemoglobin have different absorption spectra, differential absorption of light at these two wavelengths can be used to calculate the fraction of oxygen saturation of hemoglobin. Under normal circumstances, the contributions of carboxyhemoglobin and methemoglobin are minimal. However, if carboxyhemoglobin levels are elevated, the pulse oximeter will incorrectly interpret carboxyhemoglobin as oxyhemoglobin and the arterial saturation displayed will be falsely elevated. When the concentration of methemoglobin is markedly increased, the Sao\text{2} will be displayed as 85%, regardless of the true arterial saturation. The accuracy of pulse oximetry begins to decline at Sao\text{2} values less than 92% and tends to be unreliable for values less than 85%.

Several studies have assessed the frequency of arterial oxygen desaturation in hospitalized patients and its effect on outcome. Monitoring pulse oximetry in surgical patients is associated with a reduction in unrecognized deterioration, rescue events, and transfers to the ICU. Because of its clinical relevance, ease of use, noninvasive nature, and cost-effectiveness, pulse oximetry has become a routine monitoring strategy in patients with respiratory disease, intubated patients, and those undergoing surgical intervention under sedation or general anesthesia. Pulse oximetry is especially useful in the titration of Fio\text{2} and PEEP for patients receiving mechanical ventilation, and during weaning from mechanical ventilation. The widespread use of pulse oximetry has decreased the need for arterial blood gas determinations in critically ill patients.

Pulse CO-Oximetry

While simple pulse oximeters such as those described previously are helpful for determination of the Sao\text{2}, extensions of the technology may prove valuable for determination of total hemoglobin concentration as well. Through the use of multiple additional wavelengths of light, clinicians can leverage the different spectrophotometric properties of the multiple different
oxidative states of hemoglobin to get a complete readout of the total hemoglobin present in a given volume, leading to a noninvasive measurement of Hgb. These devices are referred to as pulse CO-Oximeters, as opposed to pulse oximeters, to differentiate that they are capable of measuring other hemoglobin moieties. Currently, there are two such devices that are commercially available for clinical use.

Theoretically, the capacity to noninvasively measure Hgb concentration in real time would offer significant clinical benefit. These include obviating the need for serial blood draws, the early detection of potential postsurgical hemorrhage, and more judicious usage of blood transfusions. In practice, there are multiple factors that currently affect the accuracy of the technique. Multiple studies have demonstrated that biases with noninvasive Hgb monitoring are inversely correlated with hemoglobin concentration in a variety of monitoring scenarios; with decreasing hemoglobin values the noninvasive approaches tend to overestimate the true Hgb.62-64 This poses a significant challenge for monitoring the critically ill patient, as frequently anemia is a common comorbid condition. On the other hand, if the continuous monitoring capacity afforded by these monitors can provide usable trend data, that may still provide clinical utility despite less accuracy at low hemoglobin levels. To date, there have been relatively few studies validating the trending capacity of noninvasive Hgb monitoring compared to serial blood draws, with limited agreement due to differences in analysis and study design.65 Further studies are required to evaluate the clinical utility of this potentially useful technology.

**Capnometry**

Capnometry is the measurement of carbon dioxide in the airway throughout the respiratory cycle. Capnometry is most commonly measured by infrared light absorption. CO₂ absorbs infrared light at a peak wavelength of approximately 4.27 µm. Capnometry works by passing infrared light through a sample chamber to a detector on the opposite side. More infrared light passing through the sample chamber (i.e., less CO₂) causes a larger signal in the detector relative to the infrared light passing through a reference cell. Capnometric determination of the partial pressure of CO₂ in end-tidal exhaled gas (PetCO₂) is used as a surrogate reference cell. Capnometric determination of the partial pressure of CO₂ in end-tidal exhaled gas (PetCO₂) is used as a surrogate reference cell. Capnometry is most commonly measured by infrared light absorption. CO₂ absorbs infrared light at a peak wavelength of approximately 4.27 µm. Capnometry works by passing infrared light through a sample chamber to a detector on the opposite side. More infrared light passing through the sample chamber (i.e., less CO₂) causes a larger signal in the detector relative to the infrared light passing through a reference cell. Capnometric determination of the partial pressure of CO₂ in end-tidal exhaled gas (PetCO₂) is used as a surrogate reference cell. Capnometry is most commonly measured by infrared light absorption. CO₂ absorbs infrared light at a peak wavelength of approximately 4.27 µm. Capnometry works by passing infrared light through a sample chamber to a detector on the opposite side. More infrared light passing through the sample chamber (i.e., less CO₂) causes a larger signal in the detector relative to the infrared light passing through a reference cell. Capnometric determination of the partial pressure of CO₂ in end-tidal exhaled gas (PetCO₂) is used as a surrogate reference cell.

**Urine Output**

Bladder catheterization with an indwelling catheter allows the monitoring of urine output, usually recorded hourly by the nursing staff. With a patent Foley catheter, urine output is a gross indicator of renal perfusion. The generally accepted normal urine output is 0.5 mL/kg per hour for adults and 1 to 2 mL/kg per hour for neonates and infants. Oliguria may reflect inadequate renal artery perfusion due to hypotension, hypovolemia, or low Qs. Low urine flow also can be a sign of intrinsic renal dysfunction. It is important to recognize that normal urine output does not exclude the possibility of impending renal failure.

**Bladder Pressure**

The triad of oliguria, elevated peak airway pressures, and elevated intra-abdominal pressure is known as abdominal compartment syndrome (ACS). This syndrome, first described in patients after repair of ruptured abdominal aortic aneurysm, is associated with interstitial edema of the abdominal organs, resulting in elevated intra-abdominal pressure (IAP). When IAP exceeds venous or capillary pressures, perfusion of the kidneys and other intra-abdominal viscera is impaired. Oliguria is a cardinal sign. While the diagnosis of ACS is a clinical one, measuring IAP is useful to confirm the diagnosis. Ideally, a catheter inserted into the peritoneal cavity could measure IAP to substantiate the diagnosis. In practice, transurethral bladder pressure measurement reflects IAP and is most often used to confirm the presence of ACS. After instilling 50 to 100 mL of sterile saline into the bladder via a Foley catheter, the tubing is connected to a transducing system to measure bladder pressure in the supine position at end-expiration.

Intra-abdominal hypertension is defined as an IAP ≥12 mmHg recorded on three standard measurements conducted 4 to 6 hours apart and is separated into several grades. The diagnosis of ACS is the presence of an IAP ≥20 mmHg recorded by three measurements 1 to 6 hours apart, along with new onset of organ dysfunction (Table 13-5).67-69 Less commonly, gastric or inferior vena cava pressures can be monitored with appropriate catheters to detect elevated intra-abdominal pressures.

**Intracranial Pressure**

Because the brain is rigidly confined within the bony skull, cerebro spinal fluid or mass lesions increase intracranial pressure (ICP). Monitoring of ICP is currently recommended in patients with severe traumatic brain injury (TBI), defined as a Glasgow Coma Scale (GCS) score less than or equal to 8 with an abnormal computed tomography (CT) scan, and in patients with severe TBI and a normal CT scan if two or more of the following are present: age >40 years, unilateral or bilateral motor posturing,
or systolic blood pressure <90 mmHg. ICP monitoring also is indicated in patients with acute subarachnoid hemorrhage with coma or neurologic deterioration, intracranial hemorrhage with intraventricular blood, ischemic middle cerebral artery stroke, fulminant hepatic failure with coma and cerebral edema on CT scan, and global cerebral ischemia or anoxia with cerebral edema on CT scan. The goal of ICP monitoring is to ensure that cerebral perfusion pressure (CPP) is adequate to support perfusion of the brain. CPP is equal to the difference between MAP and ICP: CPP = MAP – ICP.

One type of ICP measuring device, the ventriculostomy catheter, consists of a fluid-filled catheter inserted into a cerebral ventricle and connected to an external pressure transducer. This device permits measurement of ICP but also allows drainage of cerebrospinal fluid (CSF) as a means to lower ICP and sample CSF for laboratory studies. Other devices locate the pressure transducer within the central nervous system and are used only to monitor ICP. These devices can be placed in the intraventricular, parenchymal, subdural, or epidural spaces. Ventriculostomy catheters are the accepted standard for monitoring ICP in patients with TBI due to their accuracy, ability to drain CSF, and low complication rate. The associated complications include infection (5%), hemorrhage (1.1%), catheter malfunction or obstruction (6.3–10.5%), and malposition with injury to cerebral tissue.

The purpose of ICP monitoring is to detect and treat abnormal elevations of ICP that may be detrimental to cerebral perfusion and function. In TBI patients, ICP greater than 20 mmHg is associated with unfavorable outcomes. However, few studies have shown that treatment of elevated ICP improves clinical outcomes in human trauma patients. In a randomized, controlled, double-blind trial, Eisenberg and colleagues demonstrated that maintaining ICP less than 25 mmHg in patients without craniectomy and less than 15 mmHg in patients with craniectomy is associated with improved outcome. In patients with low CPP, therapeutic strategies to correct CPP can be directed at increasing MAP or decreasing ICP. While it has been recommended that CPP be maintained between 50 and 70 mmHg, the evidence to support this recommendation are not overly compelling. Furthermore, a retrospective cohort study of patients with severe TBI found that ICP/CPP-targeted neurointensive care was associated with prolonged mechanical ventilation and increased therapeutic interventions, without evidence for improved outcome in patients who survive beyond 24 hours.

### Electroencephalogram and Evoked Potentials

Electroencephalography offers the capacity to monitor global neurologic electrical activity, while evoked potential monitoring can assess pathways not detected by the conventional EEG. Continuous EEG (CEEG) monitoring in the intensive care unit permits ongoing evaluation of cerebral cortical activity. It is especially useful in obtunded and comatose patients. CEEG also is useful for monitoring of therapy for status epilepticus and detecting early changes associated with cerebral ischemia. CEEG can be used to adjust the level of sedation, especially if high-dose barbiturate therapy is being used to manage elevated ICP. Somatosensory and brain stem evoked potentials are less affected by the administration of sedatives than is the EEG. Evoked potentials are useful for localizing brain stem lesions or proving the absence of such structural lesions in cases of metabolic or toxic coma. They also can provide prognostic data in posttraumatic coma.

An advance in EEG monitoring is the use of the bispectral index (BIS) to titrate the level of sedative medications. While sedative drugs are usually titrated to the clinical neurologic examination, the BIS device has been used in the operating room to continuously monitor the depth of anesthesia. The BIS is an empiric measurement statistically derived from a database of over 5000 EEGs. The BIS is derived from bifrontal EEG recordings and analyzed for burst suppression ratio, relative alpha to beta ratio, and bicoherence. Using a multivariate regression model, a linear numeric index (BIS) is calculated, ranging from 0 (isoelectric EEG) to 100 (fully awake). Its use has been associated with lower consumption of anesthetics during surgery and earlier awakening and faster recovery from anesthesia. The BIS also has been validated as a useful approach for monitoring the level of sedation for ICU patients, using the revised Sedation-Agitation Scale as a gold standard.

### Transcranial Doppler Ultrasonography

This modality provides a noninvasive method for evaluating cerebral hemodynamics. Transcranial Doppler (TCD) measurements of middle and anterior cerebral artery blood flow velocity are useful for the diagnosis of cerebral vasospasm after subarachnoid hemorrhage. Qureshi and associates demonstrated that an increase in the middle cerebral artery mean flow velocity as assessed by TCD is an independent predictor of symptomatic vasospasm in a prospective study of patients with aneurysmal subarachnoid hemorrhage. In addition, while some have proposed using TCD to estimate ICP, studies have shown that TCD is not a reliable method for estimating ICP and CPP and currently cannot be endorsed for this purpose. TCD also is useful to confirm the clinical examination for determining brain death in patients with confounding factors such as the presence of CNS depressants or metabolic encephalopathy.

### Jugular Venous Oximetry

When the arterial oxygen content, hemoglobin concentration, and the oxyhemoglobin dissociation curve are constant, changes in jugular venous oxygen saturation (S_jo2) reflect changes in the difference between cerebral oxygen delivery and demand.
Generally, a decrease in Sj o2 reflects cerebral hypoperfusion, whereas an increase in Sj o2 indicates the presence of hyperemia. Sj o2 monitoring cannot detect decreases in regional cerebral blood flow if overall perfusion is normal or above normal. This technique requires the placement of a catheter in the jugular bulb, usually via the internal jugular vein. Catheters that permit intermittent aspiration of jugular venous blood for analysis or continuous oximetry catheters are available.

Low Sj o2 is associated with poor outcomes after TBI. Nevertheless, the value of monitoring Sj o2 remains unproven. If it is employed, it should not be the sole monitoring technique, but rather should be used in conjunction with ICP and CPP monitoring. By monitoring ICP, CPP, and Sj o2, early intervention with volume, vasopressors, and hyperventilation has been shown to prevent ischemic events in patients with TBI.

**Transcranial Near-Infrared Spectroscopy**

Transcranial near-infrared spectroscopy (NIRS) is a noninvasive continuous monitoring method to determine cerebral oxygenation. It employs technology similar to that of pulse oximetry to determine the concentrations of oxy- and deoxyhemoglobin with near-infrared light and sensors and takes advantage of the relative transparency of the skull to light in the near-infrared region of the spectrum. Continuous monitoring of cerebral perfusion via transcranial NIRS may provide a method to detect early cerebral ischemia in patients with traumatic brain injury. Nevertheless, this form of monitoring remains largely a research tool at the present time.

Recently, some authors have reported its use as a potential triage tool for prehospital care in the management of TBI, as NIRS allows for rapid screening for intracranial hematoma. Two small EMS studies demonstrated that handheld NIRS devices may be feasible adjunct tools in this setting, particularly when CT scanners may not be readily available.

**Brain Tissue Oxygen Tension**

While the standard of care for patients with severe TBI includes ICP and CPP monitoring, this strategy does not always prevent secondary brain injury. Growing evidence suggests that monitoring local brain tissue oxygen tension (PbtO2) may be a useful adjunct to ICP monitoring in these patients. Normal values for PbtO2 are 20 to 40 mmHg, and critical levels are 8 to 10 mmHg. A recent clinical study sought to determine whether the addition of a PbtO2 monitor to guide therapy in severe traumatic brain injury was associated with improved patient outcomes. Twenty-eight patients with severe traumatic brain injury (GCS score ≤8) were enrolled in an observational study at a level I trauma center. These patients received invasive ICP and PbtO2 monitoring and were compared with 25 historical controls matched for age, injuries, and admission GCS score that had undergone ICP monitoring alone. Goals of therapy in both groups included maintaining an ICP <20 mmHg and a CPP >60 mmHg. Among patients with PbtO2 monitoring, therapy also was directed at maintaining PbtO2 >25 mmHg. The groups had similar mean daily ICP and CPP levels. The mortality rate in the historical controls treated with standard ICP and CPP management was 44%. Mortality was significantly lower in the patients who had therapy guided by PbtO2 monitoring in addition to ICP and CPP (25%; P < .05). The benefits of PbtO2 monitoring may include the early detection of brain tissue ischemia despite normal ICP and CPP. In addition, PbtO2-guided management may reduce potential adverse effects associated with therapies to maintain ICP and CPP.

**Conclusions**

Modern intensive care is predicated by the need and ability to continuously monitor a wide range of physiologic parameters. This capability has dramatically improved the care of critically ill patients and advanced the development of the specialty of critical care medicine. In some cases, the technological ability to measure such variables has surpassed our understanding of the significance or the knowledge of the appropriate intervention to ameliorate such pathophysiologic changes. In addition, the development of less invasive monitoring methods has been promoted by the recognition of complications associated with invasive monitoring devices. The future portends the continued development of noninvasive monitoring devices along with their application in an evidenced-based strategy to guide rational therapy.

**References**

Entries highlighted in bright blue are key references.


# INTRODUCTION

Minimally invasive surgery describes an area of surgery that crosses all traditional disciplines, from general surgery to neurosurgery. It is not a discipline unto itself, but more a philosophy of surgery, a way of thinking. Minimally invasive surgery is a means of performing major operations through small incisions, often using miniaturized, high-tech imaging systems, to minimize the trauma of surgical exposure. Some believe that the term minimal access surgery more accurately describes the small incisions generally necessary to gain access to surgical sites in high-tech surgery, but John Wickham’s term minimally invasive surgery (MIS) is widely used because it describes the paradox of postmodern high-tech surgery—small holes, big operations.

Robotic surgery today is practiced using a single platform (Intuitive, Inc, Sunnyvale, CA) and should better be termed computer-enhanced surgery because the term robotics assumes autonomous action that is not a feature of the da Vinci robotic system. Instead, the da Vinci robot couples an ergonomic workstation that features stereoscopic video imaging and intuitive micromanipulators (surgeon side) with a set of arms delivering specialized laparoscopic instruments enhanced with more degrees of freedom than are allowed by laparoscopic surgery alone (patient side). A computer between the surgeon side and patient side removes surgical tremor and scales motion to allow precise microsurgery, which is helpful for microdissection and difficult anastomoses.

Single-incision laparoscopic surgery (SILS), also called laparoendoscopic single-site surgery (LESS), is a recent addition to the armamentarium of the minimally invasive surgeon. As public awareness has grown, so too has its spread outside of larger institutions. SILS challenges the well-established paradigm of standard laparoscopic surgery by placing multiple trocars within the fascia at the umbilicus or through a single multichannel trocar at the umbilicus. The manipulation of tightly spaced instruments across the fulcrum of the abdominal wall requires that the surgeon either operate in a crossed hands fashion or use specialized curved instruments to avoid clashing outside the body while working intra-abdominally. The primary advantage of SILS is the reduction to one surgical scar. Greater efficacy, safety, and cost savings have yet to be fully elucidated in the increasing number of procedures that are being attempted in this manner. The advent of a robotic SILS platform now enables the computer reassignment of the surgeon’s hands, thus eliminating the difficult ergonomic challenges making the technique far more accessible.

Natural orifice transluminal endoscopic surgery (NOTES) is an extension of interventional endoscopy. Using the mouth, anus, vagina, and urethra (natural orifices), flexible endoscopes are passed through the wall of the esophagus, stomach, colon,
bladder, or vagina entering the mediastinum, the pleural space, or the peritoneal cavity. The advantage of this method of minimal access is principally the elimination of the scar associated with laparoscopy or thoracoscopy. Other advantages have yet to be elucidated, including pain reduction, need for hospitalization, and cost savings.

HISTORICAL BACKGROUND

Although the term minimally invasive surgery is relatively recent, the history of its component parts is nearly 100 years old. What is considered the newest and most popular variety of MIS, laparoscopy, is in fact the oldest. Primitive laparoscopy, placing a cystoscope within an inflated abdomen, was first performed by Kelling in 1901. Illumination of the abdomen required hot elements at the tip of the scope and was dangerous. In the late 1950s, Hopkins described the rod lens, a method of transmitting light through a solid quartz rod with no heat and little light loss. Around the same time, thin quartz fibers were discovered to be capable of trapping light internally and conducting it around corners, opening the field of fiber optics and allowing the rapid development of flexible endoscopes. In the 1970s, the application of flexible endoscopy grew faster than that of rigid endoscopy except in a few fields such as gynecology and orthopedics. By the mid-1970s, rigid and flexible endoscopes made a rapid transition from diagnostic instruments to therapeutic ones. The explosion of video-assisted surgery in the past 20 years was a result of the development of compact, high-resolution, charge-coupled devices (CCDs) that could be mounted on the internal end of flexible endoscopes or on the external end of a Hopkins telescope. Coupled with bright light sources, fiber-optic cables, and high-definition video monitors, the videolaparoscope has changed our understanding of surgical anatomy and reshaped surgical practice.

Flexible endoscopic imaging started in the 1960s with the first bundling of many quartz fibers into bundles, one for illumination and one for imaging. The earliest upper endoscopes revolutionized the diagnosis and treatment of gastroesophageal reflux and peptic ulcer disease and made possible early detection of upper and lower gastrointestinal (GI) cancer at a stage that could be cured. The first endoscopic surgical procedure was the colonoscopic polypectomy, developed by Shinya and Wolfe, two surgeons from New York City. The percutaneous endoscopic gastrostomy (PEG) invented by Gauderer and Ponsky may have been the first NOTES procedure, reported in 1981. Endoscopic pancreatic pseudocyst drainage is thought to be the next NOTES procedure developed; however, there was little energy and money put into the development of NOTES until a number of gastroenterologists claimed the ability to remove the gallbladder with a flexible endoscope, using a transgastric technique. With this pronouncement, the surgical community took notice and seized the momentum for NOTES research and development. Today most intra-abdominal NOTES procedures remain within the realm of research or incorporate a hybrid laparoscopic technique outside of highly specialized centers. Clinically the transvaginal approach has been studied the most extensively. Evaluation of 551 female patients from the German NOTES registry has shown conversion and complication rates similar to conventional laparoscopic surgery for cholecystectomy and appendectomy procedures. Endoscopic mucosal resection (EMR) of early-stage esophageal and gastric lesions has revolutionized the management of these malignancies. The peroral endoscopic myotomy (POEM) procedure for achalasia is showing clinical efficacy and gaining popularity.

As the race to minimize the size and increase the functionality of laparoscopic instruments progressed, the notion of using fewer access points to accomplish the same operations resulted in the development of single-incision laparoscopic surgery (SILS), synonymously termed laparoendoscopic single-site surgery (LESS). Viewed as a progression of laparoscopic surgery, SILS has recently garnered greater enthusiasm over its transvical NOTES counterpart. Currently the single-incision technique is used regularly across a wide variety of surgical areas including general, urologic, gynecologic, colorectal, and bariatric surgery. Although optical imaging produced the majority of MIS procedures, other (traditionally radiologic) imaging technologies allowed the development of innovative procedures in the 1970s. Fluoroscopic imaging allowed the adoption of percutaneous vascular procedures, the most revolutionary of which was balloon angioplasty. Balloon-based procedures spread into all fields of medicine used to open clogged lumens with minimal access. Stents were then developed that were used in many disciplines to keep the newly ballooned segment open. The culmination of fluoroscopic balloon and stent proficiency is exemplified by the transvenous intrahepatic portosystemic shunt and by the aortic stent graft, which has nearly replaced open elective abdominal aortic aneurysm repair.

MIS procedures using ultrasound imaging have been limited to fairly crude exercises, such as fragmenting kidney stones and freezing liver tumors, because of the relatively low...
resolution of ultrasound devices. Newer, high-resolution ultrasound methods with high-frequency crystals may act as a guide while performing minimally invasive resections of individual layers of the intestinal wall.

Axial imaging, such as computed tomography (CT), has allowed the development of an area of MIS that often is not recognized because it requires only a CT scanner and a long needle. CT-guided drainage of abdominal fluid collections and percutaneous biopsy of abnormal tissues are minimally invasive means of performing procedures that previously required a celiotomy. CT-guided percutaneous radiofrequency (RF) ablation has emerged as a useful treatment for primary and metastatic liver tumors. This procedure also is performed laparoscopically under ultrasound guidance.9

A powerful, noninvasive method of imaging that will allow the development of the least invasive—and potentially noninvasive—surgery is magnetic resonance imaging (MRI). MRI is an extremely valuable diagnostic tool, but it is only slowly coming to be of therapeutic value. One obstacle to the use of MRI for MIS is that image production and refreshment of the image as a procedure progresses are slow. Another is that all instrumentation must be nonmetallic when working with the powerful magnets of an MRI scanner. Moreover, MRI magnets are bulky and limit the surgeon’s access to the patient. Open magnets have been developed that allow the surgeon to stand between two large MRI coils, obtaining access to the portion of the patient being scanned. The advantage of MRI, in addition to the superb images produced, is that there is no radiation exposure to patient or surgeon. Some neurosurgeons are accumulating experience using MRI to perform frameless stereotactic surgery.

Robotic surgery has been dreamed about for some time, and many science fiction–like devices have been developed over the years to provide mechanical assistance for the surgeon. The first computer-assisted robot was designed to accurately drill femoral shaft bone for wobble-free placement of hip prostheses. Although the concept was appealing, the robot proved no better than a skilled orthopedic surgeon and was a good deal slower. Following this, the first and only two commercially successful robots for laparoscopic surgery were developed in California. Computer Motion, founded by Yulun Wang in Santa Barbara, used National Science Foundation funds to create a mechanical arm, the Aesop robot, which held and moved the laparoscope with voice, foot, or hand control. In Northern California, a master-slave system first developed for surgery on the multinational space station by Philip Green was purchased by Fred Moll and Lonnie Smith, and then reengineered with the surgeon in mind to create a remarkably intuitive computer-enhanced surgical platform. The company, Intuitive Surgical, was aptly named, and their primary product, the da Vinci robot, is currently the only major robotic platform on the market, although competitors are rapidly emerging in the horizon. Although eschewed by many experienced laparoscopists, the da Vinci achieved a toehold among many skilled surgeons who found that the robot could facilitate MIS procedures that were difficult with standard laparoscopic procedures. The latest iteration of the da Vinci Xi platform released in 2014 features high-definition, three-dimensional vision and a dual-console capability allowing greater visualization, assistance, and instruction capabilities. Additionally, the new overhead boom design facilitates anatomical access from virtually any position enabling complex multiquadrant surgeries.

### PHYSIOLOGY AND PATHOPHYSIOLOGY OF MINIMALLY INVASIVE SURGERY

Even with the least invasive of the MIS procedures, physiologic changes occur. Many minimally invasive procedures require minimal or no sedation, and there are few adverse consequences to the cardiovascular, endocrinologic, or immunologic systems. The least invasive of such procedures include stereotactic biopsy of breast lesions and flexible GI endoscopy. Minimally invasive procedures that require general anesthesia have a greater physiologic impact because of the anesthetic agent, the incision (even if small), and the induced pneumoperitoneum.

#### Laparoscopy

The unique feature of laparoscopic surgery is the need to lift the abdominal wall from the abdominal organs. Two methods have been devised for achieving this.10 The first, used by most surgeons, is a pneumoperitoneum. Throughout the early 20th century, intraperitoneal visualization was achieved by inflating the abdominal cavity with air, using a sphygmomanometer bulb.11 The problem with using air insufflation is that nitrogen is poorly soluble in blood and is slowly absorbed across the peritoneal surfaces. Air pneumoperitoneum was believed to be more painful than nitrous oxide (N2O) pneumoperitoneum, but less painful than carbon dioxide (CO2) pneumoperitoneum. Subsequently, CO2 and N2O were used for inflating the abdomen. N2O had the advantage of being physiologically inert and rapidly absorbed. It also provided better analgesia for laparoscopy performed under local anesthesia when compared with CO2 or air.12 Despite initial concerns that N2O would not suppress combustion, controlled clinical trials have established its safety within the peritoneal cavity.13 In addition, N2O has been shown to reduce the intraoperative end-tidal CO2 and minute ventilation required to maintain homeostasis when compared to CO2 pneumoperitoneum.13 The effect of N2O on tumor biology and the development of port site metastasis are unknown. As such, caution should be exercised when performing laparoscopic cancer surgery with this agent. Finally, the safety of N2O pneumoperitoneum in pregnancy has yet to be elucidated.

The physiologic effects of CO2 pneumoperitoneum can be divided into two areas: (a) gas-specific effects and (b) pressure-specific effects (Fig. 14-1). CO2 is rapidly absorbed across the peritoneal membrane into the circulation. In the circulation,
CO₂ creates a respiratory acidosi by the generation of carbonic acid.¹⁴ Body buffers, the largest reserve of which lies in bone, absorb CO₂ (up to 120 L) and minimize the development of hypercapnia or respiratory acidosis during brief endoscopic procedures.¹⁴ Once the body buffers are saturated, respiratory acidosis develops rapidly, and the respiratory system assumes the burden of keeping up with the absorption of CO₂ and its release from these buffers.

In patients with normal respiratory function, this is not difficult; the anesthesiologist increases the ventilatory rate or vital capacity on the ventilator. If the respiratory rate required exceeds 20 breaths per minute, there may be less efficient gas exchange and increasing hypercapnia.¹⁵ Conversely, if vital capacity is increased substantially, there is a greater opportunity for barotrauma and greater respiratory motion—induced disruption of the upper abdominal operative field. In some situations, it is advisable to evacuate the pneumoperitoneum or reduce the intra-abdominal pressure to allow time for the anesthesiologist to adjust for hypercapnia.¹⁶ Although mild respiratory acidosis probably is an insignificant problem, more severe respiratory acidosis leading to cardiac arrhythmias has been reported.¹⁷ Hypercapnia also causes tachycardia and increased systemic vascular resistance, which elevates blood pressure and increases myocardial oxygen demand.¹⁴,¹⁷

The pressure effects of the pneumoperitoneum on cardiovascular physiology also have been studied. In the hypovolemic individual, excessive pressure on the inferior vena cava and a reverse Trendelenburg position with loss of lower extremity muscle tone may cause decreased venous return and decreased cardiac output.¹⁴,¹⁸ This is not seen in the normovolemic patient. The most common arrhythmia created by laparoscopy is bradycardia. A rapid stretch of the peritoneal membrane often causes a vagovagal response with bradycardia and, occasionally, hypotension.¹⁹ The appropriate management of this event is desufflation of the abdomen, administration of vagolytic agents (e.g., atropine), and adequate volume replacement.²⁰

With the increased intra-abdominal pressure compressing the inferior vena cava, there is diminished venous return from the lower extremities. This has been well documented in the patient placed in the reverse Trendelenburg position for upper abdominal operations. Venous engorgement and decreased venous return promote venous thrombosis.²¹,²² Many series of advanced laparoscopic procedures in which deep venous thrombosis (DVT) prophylaxis was not used demonstrated the frequency of pulmonary embolus. This usually is an avoidable complication with the use of sequential compression stockings, subcutaneous heparin, or low molecular weight heparin.²³,²⁴ In short-duration laparoscopic procedures, such as appendectomy, hernia repair, or cholecystectomy, the risk of DVT may not be sufficient to warrant extensive DVT prophylaxis.

The increased pressure of the pneumoperitoneum is transmitted directly across the paralyzed diaphragm to the thoracic cavity, creating increased central venous pressure and increased filling pressures of the right and left sides of the heart. If the intra-abdominal pressures are kept under 20 mmHg, the cardiac output usually is well maintained.²²,²³ The direct effect of the pneumoperitoneum on increasing intrathoracic pressure increases peak inspiratory pressure, pressure across the chest wall, and also, the likelihood of barotrauma. Despite these concerns, disruption of blebs and consequent pneumothoraces are rare after uncomplicated laparoscopic surgery.²⁴ Pneumothoraces occurring with laparoscopic esophageal surgery may be very significant. The pathophysiology and management are discussed at the end of this section. Increased intra-abdominal pressure decreases renal blood flow, glomerular filtration rate, and urine output. These effects may be mediated by direct pressure on the kidney and the renal vein.²⁵,²⁶ The secondary effect of decreased renal blood flow is to increase plasma renin release, thereby increasing sodium retention. Increased circulating antidiuretic hormone levels also are found during the pneumoperitoneum, increasing free water reabsorption in the distal tubules.²⁷ Although the effects of the pneumoperitoneum on renal blood flow are immediately reversible, the hormonally mediated changes such as elevated antidiuretic hormone levels decrease urine output for up to 1 hour after the procedure has ended. Intraoperative oliguria is common during laparoscopy, but the urine output is not a reflection of intravascular volume status; intravenous (IV) fluid administration during an uncomplicated laparoscopic procedure should not be linked to urine output. Because insensible fluid losses through the open abdomen are eliminated with laparoscopy, the need for supplemental fluid during a laparoscopic surgical procedure should only keep up with venous pooling in the lower limbs, third-space losses into the bowel, and blood loss, which is generally less than occurs with an equivalent open operation.

The hemodynamic and metabolic consequences of pneumoperitoneum are well tolerated by healthy individuals for a prolonged period and by most individuals for at least a short period. Difficulties can occur when a patient with compromised cardiovascular function is subjected to a long laparoscopic procedure. It is during these procedures that alternative approaches should be considered or insufflation pressure reduced. Alternative gases that have been suggested for laparoscopy include the inert gases helium, neon, and argon. These gases are appealing because they cause no metabolic effects, but are poorly soluble in blood (unlike CO₂ and N₂O) and are prone to create gas emboli if the gas has direct access to the venous system.²² Gas emboli are rare but serious complications of laparoscopic surgery.²³,²⁴ They should be suspected if hypotension develops during insufflation. Diagnosis may be made by listening (with an esophageal stethoscope) for the characteristic “mill wheel” murmur. The treatment of gas embolism is to place the patient in a left lateral decubitus position with the head down to trap the gas in the apex of the right ventricle.²⁵ A rapidly placed central venous catheter then can be used to aspirate the gas out of the right ventricle.

In some situations, minimally invasive abdominal surgery can be performed without insufflation. This is possible with the assistance of an abdominal lift device that can be placed through a 10- to 12-mm trocar at the umbilicus.²⁷ These devices have the advantage of creating little physiologic derangement, but they are bulky and intrusive. The exposure and working room offered by lift devices also are inferior to those accomplished by pneumoperitoneum. Lifting the anterior abdominal wall reduces space available laterally and thereby displaces the bowel medially and anteriorly into the operative field. A pneumoperitoneum, with its well-distributed intra-abdominal pressure, provides better exposure. Abdominal lift devices also cause more postoperative pain, but they do allow the performance of MIS with standard (nonlaparoscopic) surgical instruments.

Endocrine responses to laparoscopic surgery are not always intuitive. Serum cortisol levels after laparoscopic operations are often higher than after the equivalent operation performed through an open incision.³⁰ The greatest difference
between the endocrine response of open and laparoscopic surgery is the more rapid equilibration of most stress-mediated hormone levels after laparoscopic surgery. Immune suppression also is less after laparoscopy than after open surgery. There is a trend toward more rapid normalization of cytokine levels after a laparoscopic procedure than after the equivalent procedure performed by celiotomy.31

Transhiatal mobilization of the distal esophagus is commonly performed as a component of many laparoscopic upper abdominal procedures. If there is compromise of the mediastinal pleura with resultant CO₂ pneumothorax, the defect should be enlarged so as to prevent a tension pneumothorax. Even with such a strategy, tension pneumothorax may develop, as mediastinal structures may seal the hole during inspiration, allowing the chest to fill during expiration. In addition to enlargement of the hole, a thoracostomy tube (chest tube) should be placed across the breach into the abdomen with intra-abdominal pressures reduced below 8 mmHg, or a standard chest tube may be placed. When a pneumothorax occurs with laparoscopic Nissen fundoplication or Heller myotomy, it is preferable to place an 18-French red rubber catheter with multiple side holes cut out of the distal end across the defect. At the end of the procedure, the distal end of the tube is pulled out a 10-mm port site (as the port is removed), and the pneumothorax is evacuated to a primitive water seal using a bowl of sterile water or saline. During laparoscopic esophagectomy, it is preferable to leave a standard chest tube, as residual intra-abdominal fluid will tend to be siphoned through the defect postoperatively if the tube is removed at the end of the case.

Thoracoscopy

The physiology of thoracic MIS (thoracoscopy) is different from that of laparoscopy. Because of the bony confines of the thorax, it is unnecessary to use positive pressure when working in the thorax.32 The disadvantages of positive pressure in the chest include decreased venous return, mediastinal shift, and the need to keep a firm seal at all trocar sites. Without positive pressure, it is necessary to place a double-lumen endotracheal tube so that the ipsilateral lung can be deflated when the operation starts. By collapsing the ipsilateral lung, working space within the thorax is obtained. Because insufflation is unnecessary in thoracoscopic surgery, it can be beneficial to use standard instruments via extended port sites in conjunction with thoracoscopic instruments. This approach is particularly useful when performing advanced procedures such as thoracoscopic anatomic pulmonary resection.

Extracavitary Minimally Invasive Surgery

Many MIS procedures create working spaces in extrathoracic and extraperitoneal locations. Laparoscopic inguinal hernia repair usually is performed in the anterior extraperitoneal Retzius space.33,34 Laparoscopic nephrectomy often is performed with retroperitoneal laparoscopy. Endoscopic retroperitoneal approaches to pancreatic necrosectomy have seen some limited use.35 Lower extremity vascular procedures and plastic surgical endoscopic procedures require the development of working space in unconventional planes, often at the level of the fascia, sometimes below the fascia, and occasionally in nonanatomic regions.36 Some of these techniques use insufflation of gas, but many use balloon inflation to develop the space, followed by low-pressure gas insufflation or lift devices to maintain the space (Fig. 14-2). These techniques produce fewer and less severe adverse physiologic consequences than does the pneumoperitoneum, but the insufflation of carbon dioxide into extraperitoneal locations can spread widely, causing subcutaneous emphysema and metabolic acidosis.

Anesthesia

Proper anesthesia management during laparoscopic surgery requires a thorough knowledge of the pathophysiology of the CO₂ pneumoperitoneum.20 The laparoscopic surgeon can influence cardiovascular performance by reducing or removing the CO₂ pneumoperitoneum. Insensible fluid losses are negligible, and therefore, IV fluid administration should not exceed that necessary to maintain circulating volume. MIS procedures are often outpatient procedures, so short-acting anesthetic agents are preferable. Because the factors that require hospitalization after laparoscopic procedures include the management of nausea, pain, and urinary retention, the anesthesiologist should minimize the use of agents that provoke these conditions and maximize the use of medications that prevent such problems. Critical to the anesthesia management of these patients is the use of nonnarcotic analgesics (e.g., ketorolac) when hemostasis allows it and the liberal use of antiemetic agents, including ondansetron and steroids.

The Minimally Invasive Team

From the beginning, the tremendous success of MIS was founded on the understanding that a team approach was
necessary. The many laparoscopic procedures performed daily range from basic to advanced complexity, and require that the surgical team have an intimate understanding of the operative conduct (Table 14-1). Minimally invasive procedures require complicated and fragile equipment that demands constant maintenance. In addition, multiple intraoperative adjustments to the equipment, camera, insufflator, monitors, and patient/surgeon position are made during these procedures. As such, a coordinated team approach is mandated to ensure patient safety and excellent outcomes. More and more, flexible endoscopes are used to guide or provide quality control for laparoscopic procedures. As NOTES, SILS, and robotic surgery become more common, hybrid procedures (laparoscopy and endoscopy) and complicated robotics cases will require a nursing staff capable of maintaining flexible endoscopes and understanding the operation of sophisticated technology.

A typical MIS team may consist of a laparoscopic surgeon and an operating room (OR) nurse with an interest in laparoscopic and endoscopic surgery. Adding dedicated assistants and circulating staff with an intimate knowledge of the equipment will add to and enhance team competency. Studies have demonstrated that having a designated laparoscopic team increases the efficiency and safety of laparoscopic surgery, which is translated into a benefit for the patient and the hospital.

**Room Setup and the Minimally Invasive Suite**

Nearly all MIS, whether using fluoroscopic, ultrasound, or optical imaging, incorporates a video monitor as a guide. Occasionally, two images are necessary to adequately guide the operation, as in procedures such as endoscopic retrograde cholangiopancreatography, laparoscopic common bile duct exploration, and laparoscopic ultrasonography. When two images are necessary, the images should be displayed on two adjacent video monitors or projected on a single screen with a picture-in-picture effect. The video monitor(s) should be set across the operating table from the surgeon. The patient should be interposed between the surgeon and the video monitor; ideally, the operative field also lies between the surgeon and the monitor. In pelviscopic surgery, it is best to place the video monitor at the patient’s feet, and in laparoscopic cholecystectomy, the monitor is placed at the 10 o’clock position (relative to the patient) while the surgeon stands on the patient’s left at the 4 o’clock position. The insufflating and patient-monitoring equipment ideally also is placed across the table from the surgeon so that the insufflating pressure and the patient’s vital signs and end-tidal CO₂ tension can be monitored.

The development of the minimally invasive surgical suite has been a tremendous contribution to the field of laparoscopy in that it has facilitated the performance of advanced procedures and techniques (Fig. 14-3). By having the core equipment (monitors, insufflators, and imaging equipment) located within mobile, ceiling-mounted consoles, the surgery team is able to accommodate and make small adjustments rapidly and continuously throughout the procedure. The specifically designed minimally invasive surgical suite serves to decrease equipment and cable disorganization, ease the movements of operative personnel around the room, improve ergonomics, and facilitate the use of advanced imaging equipment such as laparoscopic ultrasound. Although having a minimally invasive surgical suite available is very useful, it is not essential to successfully carry out advanced laparoscopic procedures.

**Patient Positioning**

Patients usually are placed in the supine position for laparoscopic surgery. When the operative field is the gastroesophageal junction or the left lobe of the liver, it is easiest to operate from between the legs. The legs may be elevated in Allen stirrups or abducted on leg boards to achieve this position. When pelvic procedures are performed, it usually is necessary to place the legs in Allen stirrups to gain access to the perineum. A lateral decubitus position with the table flexed provides the best access to the retroperitoneum when performing nephrectomy or adrenalectomy. For laparoscopic splenectomy, a 45° tilt of the patient provides excellent access to the lesser sac and the lateral peritoneal attachments to the spleen. For thoracoscopic surgery, the patient is placed in the lateral position with table flexion to open the intercostal spaces and the distance between the iliac crest and costal margin (Fig. 14-4). Additional consideration must be made in robotic operations to position the

---

**Table 14-1**

<table>
<thead>
<tr>
<th>BASIC CONSIDERATIONS</th>
<th>ADVANCED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laparoscopic surgical procedures</td>
<td></td>
</tr>
<tr>
<td>Appendixectomy</td>
<td>Nissen fundoplication</td>
</tr>
<tr>
<td>Cholecystectomy</td>
<td>Heller myotomy</td>
</tr>
<tr>
<td>Inguinal hernia repair</td>
<td>Paraesophageal hernia</td>
</tr>
<tr>
<td>Enteral access</td>
<td>Gastrectomy</td>
</tr>
<tr>
<td></td>
<td>Lysis of adhesions</td>
</tr>
<tr>
<td></td>
<td>Bile duct exploration</td>
</tr>
<tr>
<td></td>
<td>Colectomy</td>
</tr>
<tr>
<td></td>
<td>Splenectomy</td>
</tr>
<tr>
<td></td>
<td>Adrenalectomy</td>
</tr>
<tr>
<td></td>
<td>Nephrectomy</td>
</tr>
<tr>
<td></td>
<td>Lymph node dissection</td>
</tr>
<tr>
<td></td>
<td>Robotics</td>
</tr>
<tr>
<td></td>
<td>Bariatric</td>
</tr>
<tr>
<td></td>
<td>Complex abdominal wall reconstruction</td>
</tr>
<tr>
<td></td>
<td>Esophagectomy</td>
</tr>
<tr>
<td></td>
<td>Hepatectomy</td>
</tr>
<tr>
<td></td>
<td>Pancreatectomy</td>
</tr>
<tr>
<td></td>
<td>Prostatectomy</td>
</tr>
<tr>
<td></td>
<td>Hysterectomy</td>
</tr>
</tbody>
</table>

**Figure 14-3.** An example of a typical minimally invasive surgery suite. All core equipment is located on easily movable consoles.
Figure 14-4. Proper padding and protection of pressure points is an essential consideration in laparoscopic and thoracoscopic approaches. In preparation for thoracoscopy, this patient is placed in left lateral decubitus position with the table flexed, which serves to open the intercostal spaces and increase the distance between the iliac crest and the inferior costal margin.

Patient appropriately before starting. Clashing of the robotic arms with surrounding equipment or each other can occur if not positioned correctly. This is more common in predecessors of the da Vinci Xi platform. Unless an operative table with integrated table motion is available, once the robot is docked to the patient the bed cannot be moved without undocking.

When the patient’s knees are to be bent for extended periods or the patient is going to be placed in a reverse Trendelenburg position for more than a few minutes, DVT prophylaxis should be used. Sequential compression devices should be placed on the lower extremities during laparoscopic procedures to increase venous return and provides inhibition of thromboplastin activation.

General Principles of Access
The most natural ports of access for MIS and NOTES are the anatomic portals of entry and exit. The nares, mouth, anus, vagina, and urethra are used to access the respiratory, GI, and urinary systems. The advantage of using these points of access is that no incision is required. The disadvantages lie in the long distances between the orifice and the region of interest. For NOTES procedures, the vagina may serve as point of access, entering the abdomen via the posterior cul-de-sac of the pelvis. Similarly, the peritoneal cavity may be reached through the side wall of the stomach or colon.

Access to the vascular system may be accomplished under local anesthesia by cutting down and exposing the desired vessel, usually in the groin. Increasingly, vascular access is obtained with percutaneous techniques using a small incision, a needle, and a guidewire, over which are passed a variety of different-sized access devices. This approach, known as the Seldinger technique, is most frequently used by general surgeons for placement of Hickman catheters, but it also is used to gain access to the arterial and venous system for performance of minimally invasive procedures. Guidewire-assisted, Seldinger-type techniques also are helpful for gaining access to the gut for procedures such as PEG, for gaining access to the biliary system through the liver, and for gaining access to the upper urinary tract.

In thoracoscopic surgery, the access technique is similar to that used for placement of a chest tube. In these procedures, general anesthesia and single lung ventilation are essential. A small incision is made over the top of a rib and, under direct vision, carried down through the pleura. The lung is collapsed, and a trocar is inserted across the chest wall to allow access with a telescope. Once the lung is completely collapsed, subsequent access may be obtained with direct puncture, viewing all entry sites through the videoendoscope. Because insufflation of the chest is unnecessary, simple ports that keep the small incisions open are all that is required to allow repeated access to the thorax.

Laparoscopic Access
The requirements for laparoscopy are more involved because the creation of a pneumoperitoneum requires that instruments of access (trocars) contain valves to maintain abdominal inflation.

Two methods are used for establishing abdominal access during laparoscopic procedures. The first, direct puncture laparoscopy, begins with the elevation of the relaxed abdominal wall with two towel clips or a well-placed hand. A small incision is made in the umbilicus, and a specialized spring-loaded (Veress) needle is placed in the abdominal cavity (Fig. 14-5).

Figure 14-5. A. Tip of spring loaded (Veress) needle. B. Veress needle held at its serrated collar with a thumb and forefinger. At the umbilicus, the abdominal wall is grasped with fingers or penetrating towel clip to elevate the abdominal wall away from the underlying structures.
With the Veress needle, two distinct pops are felt as the surgeon passes the needle through the abdominal wall fascia and the peritoneum. The umbilicus usually is selected as the preferred point of access because, in this location, the abdominal wall is quite thin, even in obese patients. The abdomen is inflated with a pressure-limited insufflator. CO₂ gas usually is used, with maximal pressures in the range of 14 to 15 mmHg. During the process of insufflation, it is essential that the surgeon observe the pressure and flow readings on the monitor to confirm an intraperitoneal location of the Veress needle tip (Fig. 14-6). Laparoscopic surgery can be performed under local anesthesia, but general anesthesia is preferable. Under local anesthesia, N₂O is used as the insufflating agent, and insufflation is stopped after 2 L of gas is insufflated or when a pressure of 10 mmHg is reached.

After peritoneal insufflation, direct access to the abdomen is obtained with a 5- or 10-mm trocar. This can be performed through a radially dilating sheath placed over the Veress needle or an optical viewing trocar. In the latter technique, a camera is placed inside of a clear pyramidal trocar. Direct puncture entry is observed as the trocar is passed through the abdominal wall. The critical issues for safe direct-puncture laparoscopy include the use of a vented stylet for the trocar, or a trocar with a safety shield or dilating tip. An optical viewing trocar can be used without prior insufflation; however, proper recognition of the abdominal wall layers is critical to avoid entry into the mesentery or underlying structures. In all direct puncture entry the trocar must be pointed away from the sacral promontory and the great vessels. Patient position should be surveyed before trocar placement to ensure a proper trajectory.

Occasionally, the direct peritoneal access (Hasson) technique is advisable. With this technique, the surgeon makes a small incision just below the umbilicus and under direct vision locates the abdominal fascia. Two Kocher clamps are placed on the fascia, and with curved Mayo scissors, a small incision is made through the fascia and underlying peritoneum. A finger is placed into the abdomen to make sure that there is no adherent bowel. A sturdy suture is placed on each side of the fascia and secured to the wings of a specialized trocar, which is then passed directly into the abdominal cavity (Fig. 14-7). Rapid insufflation can make up for some of the time lost with the initial dissection. This technique is preferable for the abdomen of patients who have undergone previous operations in which small bowel may be adherent to the undersurface of the abdominal wound. The close adherence of bowel to the peritoneum in the previously operated abdomen does not eliminate the possibility of intestinal injury but should make great vessel injury extremely unlikely. Because of the difficulties in visualizing the abdominal region immediately adjacent to the primary trocar, it is recommended that the telescope be passed through a secondary trocar to inspect the site of initial abdominal access. Secondary punctures are made with 5- and 10-mm trocars. For safe access to the abdominal cavity, it is critical to visualize all sites of trocar entry. At the completion of the operation, all trocars are removed under direct vision, and the insertion sites are inspected for bleeding. If bleeding occurs, direct pressure with an instrument from another trocar site or balloon tamponade with a Foley catheter placed through the trocar site generally stops the bleeding within 3 to 5 minutes. When this is not successful, a full-thickness abdominal wall suture has been used successfully to tamponade trocar site bleeding.

It is generally agreed that 5-mm trocars need no site suturing. Ten-millimeter trocars placed off the midline, through a radially dilating sheath or above the transverse mesocolon do not typically require repair. Conversely, if the fascia has been dilated to allow the passage of the gallbladder or other organ, it should be repaired at the fascial level with interrupted sutures. The port site may be closed with suture delivery systems similar to crochet needles enabling mass closure of the abdominal wall. This is especially helpful in obese patients where direct fascial closure may be challenging, through a small skin incision. Failure to close lower abdominal trocar sites that are 10 mm in diameter or larger can lead to an incarcerated hernia.

**Access for Subcutaneous and Extraperitoneal Surgery**

There are two methods for gaining access to nonanatomic spaces. For retroperitoneal locations, balloon dissection is effective. This access technique is appropriate for the extraperitoneal repair of inguinal hernias and for retroperitoneal surgery for adrenalectomy, nephrectomy, lumbar discectomy, pancreatic necrosectomy, or para-aortic lymph node dissection. The
initial access to the extraperitoneal space is performed in a way similar to direct puncture laparoscopy, except that the last layer (the peritoneum) is not traversed. Once the transversalis fascia has been punctured, a specialized trocar with a balloon on the end is introduced. The balloon is inflated in the extraperitoneal space to create a working chamber. The balloon then is deflated, and a Hasson trocar is placed. An insufflation pressure of 10 mmHg usually is adequate to keep the extraperitoneal space open for dissection and will limit subcutaneous emphysema. Higher gas pressures force CO₂ into the soft tissues and may contribute to hypercarbia. Extraperitoneal endosurgery provides less working space than laparoscopy but eliminates the possibility of intestinal injury, intestinal adhesion, herniation at the trocar sites, and ileus. These issues are important for laparoscopic hernia repair because extraperitoneal approaches prevent the small bowel from sticking to the prosthetic mesh.34

Subcutaneous surgery has been most widely used in cardiac, vascular, and plastic surgery.36 In cardiac surgery, subcutaneous access has been used for saphenous vein harvesting, and in vascular surgery for ligation of subfascial perforating veins (Linton procedure). With minimally invasive techniques, the entire saphenous vein above the knee may be harvested through a single incision (Fig. 14-8).45,46 Once the saphenous vein is located, a long retractor that holds a 5-mm laparoscope allows the coaxial dissection of the vein and coagulation or clipping of each side branch. A small incision above the knee also can be used to ligate perforating veins in the lower leg.

Subcutaneous access also is used for plastic surgery procedures.46 Minimally invasive approaches are especially well suited to cosmetic surgery, in which attempts are made to hide the incision. It is easier to hide several 5-mm incisions than one long incision. The technique of blunt dissection along fascial planes combined with lighted retractors and endoscope-holding retractors is most successful for extensive subcutaneous surgery. Some prefer gas insufflation of these soft tissue planes. The primary disadvantage of soft tissue insufflation is that subcutaneous emphysema can be created.

**Hand-Assisted Laparoscopic Access**

Hand-assisted laparoscopic surgery is thought to combine the tactile advantages of open surgery with the minimal access of laparoscopy and thoracoscopy. This approach commonly is used to assist with difficult cases before conversion to celiotomy is necessary. Additionally, hand-assisted laparoscopic surgery is used to help surgeons negotiate the steep learning curve associated with advanced laparoscopic procedures.47 This technology uses an entryway for the hand that preserves the pneumoperitoneum and enables laparoscopic visualization in combination with the use of minimally invasive instruments (Fig. 14-9). Formal investigation of this modality has been limited primarily to case reports and small series and has focused primarily on solid organ and colon surgery.

Intraperitoneal, intrathoracic, and retroperitoneal access for robotic surgery adheres to the principles of laparoscopic and thoracoscopic access; however, the port size for the primary puncture is 12 mm to allow placement of the stereo laparoscope. Remaining trocars are 8 mm.

**Natural Orifice Transluminal Endoscopic Surgery Access**

Multiple studies have shown safety in the performance of NOTES procedures. Transvaginal, transvesicle, transanal, transcolonic, transgastric, and transoral approaches have all been attempted with varying success. The ease of decontamination, entry, and closure of these structures create variable challenges. The transvaginal approach for resection of the uterus has been employed for many years by gynecologists and has been modified by laparoscopists with great success. Extraction of the gallbladder, kidney, bladder, large bowel, and stomach can be

**Figure 14-8.** With two small incisions, virtually the entire saphenous vein can be harvested for bypass grafting.

**Figure 14-9.** This is an example of hand-assisted laparoscopic surgery during left colectomy. The surgeon uses a hand to provide retraction and counter tension during mobilization of the colon from its retroperitoneal attachments, as well as during division of the mesocolon. This technique is particularly useful in the region of the transverse colon.
Figure 14-10. Submucosal tunnel technique for transesophageal mediastinoscopy. (Reproduced with permission from Khashab MA, Kalloo AN. NOTES: current status and new horizons, Gastroenterology. 2012 Apr;142(4):704-710.e1.)

performed via the vagina. The esophagus can be traversed to enter the mediastinum. Leaving the orifice or organ of entry with an endoscope requires the use of an endoscopic needle knife followed by submucosal tunneling or direct puncture and balloon dilation (Fig. 14-10). Closure has been performed using endoscopic clips or sutures with advanced endoscopic platforms.

**Single-Incision Laparoscopic Surgery Access**

There is no standardized approach for SILS, and access techniques vary by surgeon preference. Traditionally, a single skin incision is made directly through the umbilical scar ranging from 1 to 3 cm. Through this single incision, multiple low-profile trocars can be placed separately into the fascia to allow insufflation, camera, and working instruments. The advantage of this technique is that conventional laparoscopic tools can be employed. The disadvantage becomes apparent when an extraction site is needed. A variety of specialized multilumen trocars are on the market that can be placed through the umbilical ring (Fig. 14-11A,B). The advantages of these devices include faster access, improved safety, minimization of air leaks, and platform-derived instrument triangulation. The major disadvantage is cost.

**Port Placement**

Trocars for the surgeon’s left and right hand should be placed at least 10 cm apart. For most operations, it is possible to orient the telescope between these two trocars and slightly back from them. The ideal trocar orientation creates an equilateral triangle between the surgeon’s right hand, left hand, and the telescope, with 10 to 15 cm on each leg. If one imagines the target of the operation (e.g., the gallbladder or gastroesophageal junction) oriented at the apex of a second equilateral triangle built on the first, these four points of reference create a diamond (Fig. 14-12). The surgeon stands behind the telescope, which provides optimal ergonomic orientation but frequently requires that a camera operator (or mechanical camera holder) reach between the surgeon’s hands to guide the telescope. SILS is challenging for even the experienced laparoscopist because it violates most of the aforementioned ergonomic principles. Having only a single point of entry into the abdominal cavity creates an inherently crowded port and hand position. The inability to space trocars severely limits the ability to triangulate the left- and right-hand instruments. As a result, the surgeon must often work in a crossed hands fashion (Fig. 14-13). Additionally, the axis of the camera view is often in line with the working instruments, making visualization difficult without a deflectable tip laparoscope.

The position of the operating table should permit the surgeon to work with both elbows in at the sides, with arms bent 90° at the elbow. It usually is necessary to alter the operating table position with left or right tilt with the patient in the Trendelenburg or reverse Trendelenburg position, depending on the operative field.
**Imaging Systems**

Two methods of videoendoscopic imaging are widely used. Both methods use a camera with a charge-coupled device (CCD), which is an array of photosensitive sensor elements (pixels) that convert the incoming light intensity to an electric charge. The electric charge is subsequently converted into a color image.

With videoendoscopy, the CCD chip is placed on the internal end of a long, flexible endoscope. With older flexible endoscopes, thin quartz fibers are packed together in a bundle, and the CCD camera is mounted on the external end of the endoscope. Most standard GI endoscopes have the CCD chip at the distal end, but small, delicate choledochoscopes and nephrosopes are equipped with fiber-optic bundles. Distally mounted CCD chips have been developed for laparoscopy but remain very expensive and therefore have not become as widely used.

Video cameras come in two basic designs. Nearly all laparoscopic cameras contain a red, green, and blue input, and are identical to the color cameras used for television production. An additional feature of many video cameras is digital enhancement. Digital enhancement detects edges, areas where there are drastic color or light changes between two adjacent pixels. By enhancing this difference, the image appears sharper and surgical resolution is improved. New laparoscopic cameras contain a high-definition (HD) chip, which increases the lines of resolution from 480 to 1080 lines. To enjoy the benefit of the clarity of HD video imaging, HD monitors also are necessary.

Priorities in a video imaging system for MIS are illumination first, resolution second, and color third. Without the first two attributes, video surgery is unsafe. Illumination and resolution are as dependent on the telescope, light source, and light cable as on the video camera used. Imaging for laparoscopy, thoracoscopy, and subcutaneous surgery uses a rigid metal telescope, usually 30 cm in length. Longer telescopes are available for obese patients and for reaching the mediastinum and deep in the pelvis from a periumbilical entry site. The standard...
telescope contains a series of quartz optical rods and focusing lenses.\textsuperscript{55} Telescopes vary in size from 2 to 12 mm in diameter. Because light transmission is dependent on the cross-sectional area of the quartz rod, when the diameter of a rod/lens system is doubled, the illumination is quadrupled. Little illumination is needed in highly reflective, small spaces such as the knee, and a very small telescope will suffice. When working in the abdominal cavity, especially if blood is present, the full illumination of a 10-mm telescope usually is necessary.

Rigid telescopes may have a flat or angled end. The flat end provides a straight view ($0^\circ$), and the angled end provides an oblique view ($30^\circ$ or $45^\circ$).\textsuperscript{52} Angled telescopes allow greater flexibility in viewing a wider operative field through a single trocar site (Fig. 14-14A); rotating an angled telescope changes the field of view. The use of an angled telescope has distinct advantages for most videoendoscopic procedures, particularly in visualizing the common bile duct during laparoscopic cholecystectomy or visualizing the posterior esophagus or the tip of the spleen during laparoscopic fundoplication. Flexible tip laparoscopes offer even greater optical freedom.

Light is delivered to the endoscope through a fiber-optic light cable. These light cables are highly inefficient, losing $>90\%$ of the light delivered from the light source. Extremely bright light sources (300 watts) are necessary to provide adequate illumination for laparoscopic surgery.

The quality of the videoendoscopic image is only as good as the weakest component in the imaging chain (Fig. 14-15). Therefore, it is important to use a video monitor that has a resolution equal to or greater than the camera being used.\textsuperscript{55} Resolution is the ability of the optical system to distinguish between line pairs. The larger the number of line pairs per millimeter, the sharper and more detailed the image. Most high-resolution monitors have up to 700 horizontal lines. HD television can deliver up to eight times more resolution than standard monitors; when combined with digital enhancement, a very sharp and well-defined image can be achieved.\textsuperscript{52,55} A heads-up display is a high-resolution liquid crystal monitor that is built into eyewear worn by the surgeon.\textsuperscript{56} This technology allows the surgeon to view the endoscopic image and operative field simultaneously. The proposed advantages of heads-up display include a high-resolution monocular image, which affords the surgeon mobility and reduces vertigo and eyestrain. However, this technology has not yet been widely adopted.

Interest in three-dimensional (3-D) laparoscopy has waxed and waned. 3-D laparoscopy provides the additional depth of field that is lost with two-dimensional endosurgery and improves performance of novice laparoscopists performing complex tasks of dexterity, including suturing and knot tying.\textsuperscript{57} The advantages of 3-D systems are less obvious to experienced Brunicardi_Ch14_p0453-p0478.indd   464 01/03/19   4:58 PM
laparoscopists. Additionally, because 3-D systems require the flickering of two similar images, which are resolved with special glasses, the images’ edges become fuzzy and resolution is lost. The optical accommodation necessary to rectify these slightly differing images is tiring and may induce headaches when one uses these systems for a long period of time. The da Vinci robot uses a specialized laparoscope with two optical bundles on opposite sides of the telescope. A specialized binocular eyepiece receives input from two CCD chips, each capturing the image from one of the two quartz rod lens systems, thereby creating true 3-D imaging without needing to employ active or passive technologies that have made 3-D laparoscopy so disappointing.

Single-incision laparoscopy presents new challenges to visualization of the operative field. In the traditional laparoscope, the light source enters the scope at a 90° angle. That position coupled with a bulky scope handle creates crowding in an already limited space. Additionally, because the scope and instruments enter the abdomen at the same point, an adequate perspective is often unobtainable even with a 30° scope. The advent of increased length laparoscopes with lighting coming from the end and a deflectable tip now allows the surgeon to recreate a sense of internal triangulation with little compromise externally. The ability to move the shaft of the scope off line while maintaining the same image provides a greater degree of freedom for the working ports.

Energy Sources for Endoscopic and Endoluminal Surgery

Many MIS procedures use conventional energy sources, but the benefits of bloodless surgery to maintain optimal visualization have spawned new ways of applying energy. The most common energy source is RF electrosurgery using an alternating current with a frequency of 500,000 cycles/s (Hz). Tissue heating progresses through the well-known phases of coagulation (60°C [140°F]), vaporization and desiccation (100°C [212°F]), and carbonization (>200°C [392°F]).68

The two most common methods of delivering RF electrosurgery are with monopolar and bipolar electrodes. With monopolar electrosurgery, a remote ground plate on the patient’s leg or back receives the flow of electrons that originate at a point source, the surgical electrode. A fine-tipped electrode causes a high current density at the site of application and rapid tissue heating. Monopolar electrosurgery is inexpensive and easy to modulate to achieve different tissue effects.59 A short-duration, high-voltage discharge of current (coagulation current) provides extremely rapid tissue heating. Lower-voltage, higher-wattage current (cutting current) is better for tissue desiccation and vaporization. When the surgeon desires tissue division with the least amount of thermal injury and least coagulation necrosis, a cutting current is used.

With bipolar electrosurgery, the electrons flow between two adjacent electrodes. The tissue between the two electrodes is heated and desiccated. There is little opportunity for tissue cutting when bipolar current is used alone, but the ability to coapt the electrodes across a vessel provides the best method of small-vessel coagulation without thermal injury to adjacent tissues.60 Advanced laparoscopic device manufacturers have leveraged the ability to selectively use bipolar energy and combined it with compressive force and a controllable blade to create a number of highly functional dissection and vessel-sealing tools (Fig. 14-16).

To avoid thermal injury to adjacent structures, the laparoscopic field of view must include all uninsulated portions of the electrosurgical electrode. In addition, the integrity of the insulation must be maintained and assured. Capacitive coupling occurs when a plastic trocar insulates the abdominal wall from the current; in turn, the current is bled off of a metal sleeve or laparoscope into the viscera54 (Fig. 14-17A). This may result in thermal necrosis and a delayed fecal fistula. Another potential mechanism for unrecognized visceral injury may occur with the direct coupling of current to the laparoscope and adjacent bowel59 (Fig. 14-17B).

Another method of delivering RF electrosurgery is argon beam coagulation. This is a type of monopolar electrosurgery in which a uniform field of electrons is distributed across a tissue surface by the use of a jet of argon gas. The argon gas jet distributes electrons more evenly across the surface than does spray electrofulguration. This technology has its greatest application for coagulation of diffusely bleeding surfaces such as the cut edge of liver or spleen. It is of less value in laparoscopic procedures because the increased intra-abdominal pressures created by the argon gas jet can increase the chances of a gas embolus. It is paramount to vent the ports and closely monitor insufflation pressure when using this source of energy within the context of laparoscopy.

With endoscopic endoluminal surgery, RF alternating current in the form of a monopolar circuit represents the mainstay for procedures such as snare polypectomy, sphincterotomy, lower esophageal sphincter ablation, and biopsy.61,62 A grounding (return) electrode is necessary for this form of energy. Bipolar electrocoagulation is used primarily for thermal hemostasis. The electrosurgical generator is activated by a foot pedal so the endoscopist may keep both hands free during the endoscopic procedure.

Gas, liquid, and solid-state lasers have been available for medical application since the mid-1960s.63 The CO₂ laser (wavelength 10.6 µm) is most appropriately used for cutting...
and superficial ablation of tissues. It is most helpful in locations unreachable with a scalpel such as excision of vocal cord granulomas. The CO₂ laser beam must be delivered with a series of mirrors and is therefore somewhat cumbersome to use. The next most popular laser is the neodymium yttrium-aluminum garnet (Nd:YAG) laser. Nd:YAG laser light is 1.064 µm (1064 nm) in wavelength. It is in the near-infrared portion of the spectrum and, like CO₂ laser light, is invisible to the naked eye. A unique feature of the Nd:YAG laser is that 1064-nm light is poorly absorbed by most tissue pigments and therefore travels deep into tissue. Deep tissue penetration provides deep tissue heating (Fig. 14-18). For this reason, the Nd:YAG laser is capable of the greatest amount of tissue destruction with a single application.

Such capabilities make it the ideal laser for destruction of large fungating tumors of the rectosigmoid, tracheobronchial tree, or esophagus. A disadvantage is that the deep tissue heating may cause perforation of a hollow viscus.

When it is desirable to coagulate flat lesions in the cecum, a different laser should be chosen. The frequency-doubled Nd:YAG laser, also known as the KTP laser (potassium thionyl phosphate crystal is used to double the Nd:YAG frequency), provides 532-nm light. This is in the green portion of the spectrum, and at this wavelength, selective absorption by red pigments in tissue (such as hemangiomas and arteriovenous malformations) is optimal. The depth of tissue heating is intermediate, between those of the CO₂ and the Nd:YAG lasers. Coagulation (without vaporization) of superficial vascular lesions can be obtained without intestinal perforation.

In flexible GI endoscopy, the CO₂ and Nd:YAG lasers have largely been replaced by heater probes and endoluminal stents. The heater probe is a metal ball that is heated to a temperature (60–100°C [140–212°F]) that allows coagulation of bleeding lesions without perforation.

Photodynamic therapy is a palliative treatment for obstructing cancers of the GI tract. Patients are given an IV dose of porfimer sodium, which is a photosensitizing agent that is taken up by malignant cells. Two days after administration, the drug is endoscopically activated using a laser. The activated porfimer sodium generates oxygen free radicals, which kill the tumor cells. The tumor is later endoscopically debrided. The use of this modality for definitive treatment of early cancers is limited.

A unique application of laser technology provides extremely rapid discharge (<10⁻⁶ s) of large amounts of energy (>10⁷ volts). These high-energy lasers, of which the pulsed dye laser has seen the most clinical use, allow the conversion of light energy to mechanical disruptive energy in the form of a shock wave. Such energy can be delivered through a quartz fiber, and with rapid repetitive discharges, can provide sufficient shock-wave energy to fragment kidney stones and gallstones. Shock waves also may be created with miniature electric spark-plug discharge systems known as electrohydraulic lithotriptors. These devices...
Lasers have the advantage of pigment selectivity, but electrohydraulic lithotriptors are more popular because they are substantially less expensive and are more compact.

Methods of producing shock waves or heat with ultrasonic energy are also of interest. Extracorporeal shockwave lithotripsy creates focused shock waves that intensify as the focal point of the discharge is approached. When the focal point is within the body, large amounts of energy are capable of fragmenting stones.

Slightly different configurations of this energy can be used to provide focused internal heating of tissues. Potential applications of this technology include the ability to noninvasively produce sufficient internal heating to destroy tissue without an incision.

A third means of using ultrasonic energy is to create rapidly oscillating instruments that are capable of heating tissue with friction; this technology represents a major step forward in energy technology. An example of its application is the laparoscopic coagulation shears device (Harmonic Scalpel), which is capable of coagulating and dividing blood vessels by first occluding them and then providing sufficient heat to weld the blood vessel walls together and to divide the vessel (Fig. 14-19). This nonelectric method of coagulating and dividing tissue with a minimal amount of collateral damage has facilitated the performance of numerous endosurgical procedures. It is especially useful in the control of bleeding from medium-sized vessels that are too big to manage with monopolar electrocautery. The ability to clamp tissue between an active blade and passive blade allows annealing of tissues followed by cutting.

**Instrumentation**

Hand instruments for MIS usually are duplications of conventional surgical instruments made longer, thinner, and smaller at the tip. It is important to remember that when grasping tissue with laparoscopic instruments, a greater force is applied over a smaller surface area, which increases the risk for perforation or injury.

Certain conventional instruments such as scissors are easy to reproduce with a diameter of 3 to 5 mm and a length of 20 to 45 cm, but other instruments such as forceps and clamps cannot provide remote access. Different configurations of graspers were developed to replace the various configurations of surgical forceps and clamps. Standard hand instruments are 5 mm in diameter and 30 cm in length, but smaller and shorter hand instruments are now available for pediatric surgery, for microlaparoscopic surgery, and for arthroscopic procedures. A unique laparoscopic hand instrument is the monopolar electrical hook. This device usually is configured with a suction and irrigation apparatus to eliminate smoke and blood from the operative field. The monopolar hook allows tenting of tissue over a bare metal wire with subsequent coagulation and division of the tissue.

Instrumentation for NOTES is still evolving, but many long micrograspers, microscissors, electrocautery adapters, suturing devices, clip applicators, and visceral closure devices are in design and application. These instruments often require an entirely different endoscopic platform requiring manipulation by a surgeon and assistant to accomplish complex maneuvers. Techniques such as mucosotomy, hydrodissection, and clip application require specialized training. The sheer size of the instrumentation often requires an overtube to allow easy exchange throughout the procedure. Instrumentation for SILS seeks to restore the surgeon’s ability to triangulate the left and right hands through variation in length, mechanical articulation, or curved design. Additionally, a lower profile camera head helps reduce the instrument crowding that occurs at the single point of abdominal entry.

**Robotic Surgery**

The term robot defines a device that has been programmed to perform specific tasks in place of those usually performed by people. The devices that have earned the title “surgical robots” would be more aptly termed computer-enhanced surgical devices, as they are controlled entirely by the surgeon for the purpose of improving performance. The first computer-assisted surgical device was the laparoscopic camera holder (Aesop, Computer Motion, Goleta, CA), which enabled the surgeon to maneuver the laparoscope either with a hand control, foot control, or voice activation. Randomized studies with such camera holders demonstrated a reduction in operative time, steadier image, and a reduction in the number of required laparoscope cleanings. This device had the advantage of eliminating the need for a human camera holder, which served to free valuable OR personnel for other duties. This technology has now been eclipsed by simpler systems using passive positioning of the camera with a mechanical arm, but the benefits of a steadier image and fewer members of the OR team remain.

The major revolution in robotic surgery was the development of a master-slave surgical platform that returned the wrist to laparoscopic surgery and improved manual dexterity by developing an ergonomically comfortable work station, with 3-D imaging, tremor elimination, and scaling of movement (e.g., large, gross hand movements can be scaled down to allow suturing with microsurgical precision) (Fig. 14-20). The most recent iteration of the robotic platform features a second surgical console enabling greater assisting and teaching opportunities. The surgeon is physically separated from the operating table, and the working arms of the device are placed over the patient (Fig. 14-21). An assistant remains at the bedside and changes the instruments as needed, providing retraction as needed to facilitate the procedure. The robotic platform (da Vinci, Intuitive Surgical, Sunnyvale, CA) was initially greeted with some skepticism by expert laparoscopists, as it was difficult to prove additional value for operations performed with the da Vinci robot. Not only were the operations longer and the equipment more expensive, but additional quality could not be demonstrated. Two randomized controlled trials compared robotic and conventional laparoscopic approaches to Nissen fundoplication. In both of these trials, the operative time was longer for robotic surgery, and there was no difference in ultimate outcome. Similar results were achieved for laparoscopic cholecystectomy. Nevertheless, the increased dexterity provided by the da

![Figure 14-19. Ultrasonic shear. When closed vibration of black (active blade) against white (passive blade) cuts and cauterizes intervening tissue.](image-url)
The tidal wave of enthusiasm for robotic surgery came when most minimally invasive urologists declared robotic prostatectomy to be preferable to laparoscopic and open prostatectomy. The great advantage—it would appear—of robotic prostatectomy is the ability to visualize and spare the pelvic nerves responsible for erectile function. In addition, the creation of the neocystourethrotomy, following prostatectomy, was greatly facilitated by needle holders and graspers with a wrist in them. Female pelvic surgery with the da Vinci robot is also reaching wide appeal. The magnified imaging provided makes this approach ideal for microsurgical tasks such as reanastomosis of the Fallopian tubes. In general surgery, there is emerging
popularity for the use of the robotic platform for revisional bariatric surgery and complex abdominal wall reconstruction. The ability to close the defect before placement of mesh in ventral hernia repairs or to perform complex transversus abdominus release herniorrhaphy is revolutionizing MIS hernia repair.

The final frontier for computer-enhanced surgery is the promise of telesurgery, in which the surgeon is a great distance from the patient (e.g., combat or space). This application has rarely been used, as the safety provided by having the surgeon at bedside cannot be sacrificed to prove the concept. However, remote laparoscopic cholecystectomy has been performed when a team of surgeons located in New York performed a cholecystectomy on a patient located in France.

Endoluminal and Endovascular Surgery

The fields of vascular surgery, interventional radiology, neuroradiology, gastroenterology, general surgery, pulmonology, and urology all encounter clinical scenarios that require the urgent restoration of luminal patency. Based on this need, fundamental techniques have been pioneered that are applicable to all specialties and virtually every organ system. As a result, all minimally invasive surgical procedures, from coronary artery angioplasty to palliation of pancreatic malignancy, involve the use of access devices, catheters, guidewires, balloon dilators, stents, and other devices (e.g., lasers, atherectomy catheters) that are capable of opening up the occluded biologic cylinder (Table 14-2).

Endoluminal balloon dilators may be inserted through an endoscope, or they may be fluoroscopically guided. Balloon dilators all have low compliance—that is, the balloons do not stretch as the pressure within the balloon is increased. The high pressures achievable in the balloon create radial expansion of the narrowed vessel or orifice, usually disrupting the atherosclerotic plaque, the fibrotic stricture, or the muscular band (e.g., esophageal achalasia).

Once the dilation has been attained, it is frequently beneficial to hold the lumen open with a stent. Stenting is particularly valuable in treating malignant lesions and atherosclerotic occlusions or aneurysmal disease (Fig. 14-22). Stenting is also of value to seal leaky cylinders, including aortic dissections, traumatic vascular injuries, leaking GI anastomoses, and fistulas. Stenting usually is not applicable for long-term management of benign GI strictures except in patients with limited life expectancy (Fig. 14-23).

A variety of stents are available that are divided into six basic categories: plastic stents, metal stents, drug-eluting stents (to decrease fibrovascular hyperplasia), covered metal stents, anchored stent grafts, and removable covered plastic stents (Fig. 14-24). Plastic stents came first and are used widely as endoprosthesis for temporary bypass of obstructions in the biliary or urinary systems. Metal stents generally are delivered over a balloon and expanded with the balloon to the desired size. These metal stents usually are made of titanium or nitinol and are still used in coronary stenting. A chemotherapeutic agent was added to coronary stents several years ago to decrease endothelial proliferation. These drug-eluting stents provide greater long-term patency but require long-term anticoagulation with antiplatelet agents to prevent thrombosis.

Table 14-2
Modalities and techniques of restoring luminal patency

<table>
<thead>
<tr>
<th>MODALITY</th>
<th>TECHNIQUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core out</td>
<td>Photodynamic therapy Laser Coagulation Endoscopic biopsy forceps Chemical Ultrasound</td>
</tr>
<tr>
<td>Fracture</td>
<td>Ultrasound Endoscopic biopsy Balloon</td>
</tr>
<tr>
<td>Dilate</td>
<td>Balloon Bougie Angioplasty Endoscope</td>
</tr>
<tr>
<td>Bypass</td>
<td>Transvenous intrahepatic portosystemic shunt Surgical (synthetic or autologous conduit)</td>
</tr>
<tr>
<td>Stent</td>
<td>Self-expanding metal stent Plastic stent</td>
</tr>
</tbody>
</table>

Figure 14-22. The deployment of a metal stent across an isolated vessel stenosis is illustrated. (Reproduced with permission from Hunter JG, Sackier JM, eds. Minimally Invasive Surgery. New York: McGraw-Hill; 1993:235.)
Silastic or other materials may prevent tumor ingrowth but also makes stent migration more likely. In an effort to minimize stent migration, stents have been incorporated with hooks and barbs at the proximal end of the stent to anchor it to the wall of the vessel. Endovascular stenting of aortic aneurysms has nearly replaced open surgery for this condition. Lastly, self-expanding plastic stents have been developed as temporary devices to be used in the GI tract to close internal fistulas and bridge leaking anastomoses.

**Natural Orifice Transluminal Endoscopic Surgery**

The use of the flexible endoscope to enter the GI, urinary, or reproductive tracts and then traverse the wall of the structure to enter the peritoneal cavity, the mediastinum, or the chest has strong appeal to patients wishing to avoid scars and pain caused by abdominal wall trauma. In truth, transluminal surgery has been performed in the stomach for a long time, either from the inside out (e.g., percutaneous, PEG, and transgastric pseudocyst drainage) or from the outside in (e.g., laparoscopic-assisted intragastric tumor resection). The catalyzing events for NOTES were the demonstration that a porcine gallbladder could be removed with a flexible endoscope passed through the wall of the stomach and then removed through the mouth and the demonstration in a series of 10 human cases from India of the ability to perform transgastric appendectomy. Since that time, a great deal of money has been invested by endoscopic and MIS companies to help surgeons and gastroenterologists explore this new territory. Systemic inflammatory markers such as C-reactive protein, tumor necrosis factor-α, interleukin (IL)-1β, and IL-6 have been shown to be similar in transgastric and transcolonic NOTES when compared to laparoscopy in porcine models. Concerns about the safety of transluminal access and limitations in equipment remain the greatest barriers to expansion. To date, the most headline-grabbing procedures have been the transvaginal and transgastric removal of the gallbladder (Fig. 14-25). To ensure safety, all human cases thus far have involved laparoscopic assistance to aid in retraction and ensure adequate closure of the stomach or vagina. To date, thousands of transvaginal and transgastric procedures have been performed internationally, with two large registries demonstrating noninferiority to conventional laparoscopy. The fact that the vast majority of these procedures are being done transvaginally creates an obvious limitation in applicability.

The rapid growth of endoscopic technology catalyzed by NOTES has already spun off new technologies capable of performing a wide variety of endoscopic surgical procedures from EMR, to ablation of Barrett’s esophagus, to creation of competent antireflux valves in patients with gastroesophageal reflux disease.

Peroral esophageal myotomy (POEM) has shown promise as a NOTES treatment for esophageal achalasia. In this procedure, a 1.5- to 2-cm mucosotomy is created within the anterior esophagus 10 cm proximal to the gastroesophageal junction. A submucosal tunnel is then created using a combination of electrocautery, hydrodissection, and carbon dioxide insufflation. The scope is advanced beyond the gastroesophageal junction, and a circular myotomy is performed avoiding disruption of the longitudinal fibers. The mucosotomy is then closed using endoscopic clips (Fig. 14-26). Over 1000 clinical POEM cases have been performed worldwide. Data from expert NOTES surgeons suggest that this selective myotomy avoids abdominal trauma.
and minimally disrupts the normal anatomic characteristics of the gastroesophageal junction while providing significant relief of symptoms. Randomized clinical trials and long-term follow-up need to be performed to further evaluate efficacy.

Although this application is still considered experimental, there is little doubt that when equivalent operations can be performed with less pain, fewer scars, and less disability, patients will flock to it. NOTES procedures are associated with an increased mental workload and significant learning curve for even experienced surgical endoscopists. Surgeons should engage only when they can perform these procedures with the safety and efficacy demanded by our profession.

**Single-Incision Laparoscopic Surgery**

As a surgical technique, SILS seems to be a natural progression from conventional laparoscopic surgery. As surgeons sought to reduce the number and size of abdominal wall trocars and NOTES procedures necessitated laparoscopic surveillance, the idea of a hybridization took off. An incision in the umbilicus, a preexisting scar, is thought to be less painful, have fewer wound complications, lead to quicker return to activity, and have a better cosmetic appearance than conventional laparoscopy. Perhaps one of the earliest examples of SILS is the application of laparoscopic instrumentation to resect lesions in the rectum or sigmoid colon. Using the anus as the portal of entry, transanal endoscopic microsurgery (TEMS) employs a specialized multichannel trocar to reach lesions located 8 to 18 cm away from the anal verge (Fig. 14-27).

More deformable versions of these complex trocars have been developed with features to allow insufflation and be amenable to maintaining a seal within the natural orifice of the umbilicus (see Fig. 14-11). Ports typically contain three or four channels. The latter often affords the ability to place a dedicated retractor.

There are many challenges faced by the operating surgeon in SILS procedures. These include crowded trocar placement, lack of triangulation of left- and right-hand instruments, frequent crossing or clashing of instruments, limited visualization, and limited retraction ability. These challenges are mitigated by surgeon’s experience and the development of specialized instruments. Articulating or curved instruments of varying lengths and an extended length can improve working space. Curved instruments are typically reusable and offer less clutter than their more sophisticated counterparts, providing some cost reduction (Fig. 14-28). A low-profile HD scope with or without a deflectable tip can improve visualization greatly. Even with such instrumentation, the learning curve is very steep, particularly when the surgeon is forced to work in a cross-handed technique. The accomplished SILS surgeon will possess a tool bag of innovative...
strategies to retract structures like the gallbladder away from the operative field. These tricks may range from the use of percutaneous needlescopic instruments to the application of transfascial sutures. Expert consensus recommendations for efficient SILS are shown in Tables 14-3 and 14-4. When performing SILS procedures, it is imperative to follow proven tenets of operative conduct such as visualizing the “critical view” of safety in a laparoscopic cholecystectomy. As safety should always be the paramount concern, the addition of extra trocars or conversion to traditional laparoscopy should not be considered a failure.

Contraindications include those true of traditional laparoscopy. Relative contraindications include previous surgery and high body mass index (BMI). Patients with a high BMI or central obesity can pose a challenge because the umbilicus may be located far from operative target. Size and morphology of the target organ should always be considered when doing SILS. Many studies have demonstrated equivalency to standard laparoscopic procedures regarding intraoperative and postoperative complications. However, it is questionable what the full benefit of the dramatic reduction in ergonomics and the increase in complexity provide beyond an improved cosmetic appearance. This is in large part due to the already improved benefits of laparoscopic surgery.

A meta-analysis performed by Ahmed and colleagues in 2010 found the conversion rate from SILS to conventional laparoscopy to be 0% to 24% for cholecystectomies, 0% to 41%

---

Table 14-3

<table>
<thead>
<tr>
<th>Expert panel recommendations for accomplishing single-incision laparoscopic surgery efficiently</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multichannel port preferably to be placed intraumbilically, but an extraumbilical approach can be used in certain cases</td>
</tr>
<tr>
<td>Extra ports should be used where there is a clinical need</td>
</tr>
<tr>
<td>When applicable, sutures can be useful for added retraction</td>
</tr>
<tr>
<td>Closure should be accomplished using sutures of absorbable material placed either continuously or interrupted</td>
</tr>
<tr>
<td>Skin should be closed with absorbable sutures or glue</td>
</tr>
</tbody>
</table>

for appendectomies, and 0% to 33% for nephrectomies. The most common complications were intra-abdominal abscesses and wound infections. Existing and emerging robotics platforms may provide the bridge necessary to bypass the significant technical skills learning curve required to operate through a single site (Fig. 14-29).

**SPECIAL CONSIDERATIONS**

**Pediatric Laparoscopy**

The advantages of MIS in children may be more significant than in the adult population. MIS in the adolescent is little different from that in the adult, and standard instrumentation and trocar positions usually can be used. However, laparoscopy in the infant and young child requires specialized instrumentation. The instruments are shorter (15–20 cm), and many are 3 mm in diameter rather than 5 mm. Because the abdomen of the child is much smaller than that of the adult, a 5-mm telescope provides sufficient illumination for most operations. The development of 5-mm clippers and bipolar devices has obviated the need for 10-mm trocars in pediatric laparoscopy. Because the abdominal wall is much thinner in infants, a pneumoperitoneum pressure of 8 mmHg can provide adequate exposure. DVT is rare in children, so prophylaxis against thrombosis probably is unnecessary. A wide variety of pediatric surgical procedures are frequently performed with MIS access, from pull-through procedures for congenital diaphragmatic hernias to repair of congenital diaphragmatic hernias.

**Laparoscopy During Pregnancy**

Concerns about the safety of laparoscopic cholecystectomy or appendectomy in the pregnant patient have been thoroughly investigated and are readily managed. Access to the abdomen in the pregnant patient should take into consideration the height of the uterine fundus, which reaches the umbilicus at 20 weeks. In order not to damage the uterus or its blood supply, most surgeons feel that the open (Hasson) approach should be used in favor of direct puncture laparoscopy. The patient should be positioned slightly on the left side to avoid compression of the vena cava by the uterus. Because pregnancy poses a risk for thromboembolism, sequential compression devices are essential for all procedures. Fetal acidosis induced by maternal hypercarbia also has been raised as a concern. The arterial pH of the fetus follows the pH of the mother linearly; and therefore, fetal acidosis may be prevented by avoiding a respiratory acidosis in the mother. The pneumoperitoneum pressure induced by laparoscopy is not a safety issue either as it has been proved that mid-pregnancy uterine contractions provide a much greater pressure in utero than a pneumoperitoneum of 15 mmHg. More than 100 cases of laparoscopic cholecystectomy in pregnancy have been reported with uniformly good results. The operation should be performed during the second trimester of pregnancy if possible. Protection of the fetus against intraoperative X-rays

---

**Table 14-4**

**Expert panel recommendations for single-incision laparoscopic surgery equipment and instrumentation**

<table>
<thead>
<tr>
<th>RECOMMENDED EQUIPMENT/INSTRUMENTATION</th>
<th>BENEFIT TO SURGEON</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slimline instruments with low-profile design</td>
<td>Reduces internal and external clashing</td>
</tr>
<tr>
<td>Varied-length instruments</td>
<td>Reduces extracorporeal clashing</td>
</tr>
<tr>
<td>Longer instruments</td>
<td>Advantageous for reaching the surgical field</td>
</tr>
<tr>
<td>Articulating (or prebent) instruments</td>
<td>Restore triangulation</td>
</tr>
<tr>
<td>Small-diameter, low-profile angle scope</td>
<td>Reduces clashing by providing additional space</td>
</tr>
<tr>
<td>High-definition camera</td>
<td>Achieves high-quality images for intraoperative visualization</td>
</tr>
</tbody>
</table>

is imperative. Some believe it advisable to track fetal pulse rates with a transvaginal ultrasound probe; however, the significance of fetal tachycardia or bradycardia is a bit unclear in the second trimester of pregnancy. To be prudent, however, heart rate decelerations reversibly associated with pneumoperitoneum creation might signal the need to convert to open cholecystectomy or appendectomy.

**Minimally Invasive Surgery and Cancer Treatment**

MIS techniques have been used for many decades to provide palliation for the patient with an obstructive cancer. Laser treatment, intracavitary radiation, stenting, and dilation are outpatient techniques that can be used to reestablish the continuity of an obstructed esophagus, bile duct, ureter, or airway. MIS techniques also have been used in the staging of cancer. Mediastinoscopy is still used occasionally before thoracotomy to assess the status of the mediastinal lymph nodes. Laparoscopy also is used to assess the liver in patients being evaluated for pancreatic, gastric, or hepatic resection. New technology and greater surgical skills allow for accurate minimally invasive staging of cancer. Occasionally, it is appropriate to perform palliative measures (e.g., laparoscopic gastrojejunostomy to bypass a pancreatic cancer) at the time of diagnostic laparoscopy if diagnostic findings preclude attempts at curative resection.

Initially controversial, the role of MIS to provide a safe curative treatment of cancer has proven to be no different from the principles of open surgery. All gross and microscopic tumor should be removed (an R0 resection), and an adequate lymphadenectomy should be performed to allow accurate staging. Generally, this number has been 10 to 15 lymph nodes, although there is still debate as to the value of more extensive lymphadenectomy. All of the major abdominal cancer operations have been performed with laparoscopy. Of the three major cancer resections of GI cancer (liver lobe, pancreatic head, and esophagus), only esophagectomy is routinely performed by a fair number of centers. Laparoscopic hepatectomy has attracted a loyal following, and distal pancreatectomy frequently is performed with laparoscopic access. In Japan, laparoscopic-assisted gastrectomy has become quite popular for early gastric cancer, an epidemic in Japan far exceeding that of colon cancer in North America and Northern Europe. The most common cancer operation performed laparoscopically is segmental colectomy, which has proven itself safe and efficacious in a multicenter, controlled, randomized trial.

**Considerations in the Elderly and Infirm**

Laparoscopic cholecystectomy has made possible the removal of a symptomatic gallbladder in many patients previously thought to be too elderly or too ill to undergo a laparotomy. Older patients are more likely to require conversion to laparotomy because of disease chronicity.

Operations on these patients require close monitoring of anesthesia. The intraoperative management of these patients may be more difficult with laparoscopic access than with open access. The advantage of MIS lies in what happens after the operation. Much of the morbidity of surgery in the elderly is a result of impaired mobility. In addition, pulmonary complications, urinary tract sepsis, DVT, pulmonary embolism, congestive heart failure, and myocardial infarction often are the result of improper fluid management and decreased mobility. By allowing rapid and early mobilization, laparoscopic surgery has made possible the safe performance of procedures in the elderly and infirm.

**Cirrhosis and Portal Hypertension**

Patients with hepatic insufficiency pose a significant challenge for any type of surgical intervention. The ultimate surgical outcome in this population relates directly to the degree of underlying hepatic dysfunction. Often, this group of patients has minimal reserve, and the stress of an operation will trigger complete hepatic failure or hepatorenal syndrome. These patients are at risk for major hemorrhage at all levels, including trocar insertion, operative dissection in a field of dilated veins, and secondary to an underlying coagulopathy. Additionally, ascitic leak from a port site may occur, leading to bacterial peritonitis. Therefore, a watertight port site closure should be carried out in all patients.

It is essential that the surgeon be aware of the severity of hepatic cirrhosis as judged by a Model of End-Stage Liver Disease (MELD) score or Child’s classification. Additionally, the presence of portal hypertension is a relative contraindication to laparoscopic surgery until the portal pressures are reduced with portal decompression. For example, if a patient has an incarcerated umbilical hernia and ascites, a preoperative paracentesis or transjugular intrahepatic portosystemic shunt procedure in conjunction with aggressive diuresis may be considered. Because these patients commonly are intravascularly depleted, insufflation pressures should be reduced to prevent a decrease in cardiac output, and minimal amounts of Na+-sparing IV fluids should be given.

**Economics of Minimally Invasive Surgery**

Minimally invasive surgical procedures reduce the costs of surgery most when length of hospital stay can be shortened and return to work is quickened. For example, shorter hospital stays can be demonstrated in laparoscopic cholecystectomy, Nissen fundoplication, splenectomy, and adrenalectomy. Procedures such as inguinal herniorrhaphy that are already performed as outpatient procedures are less likely to provide cost savings. Procedures that still require a 4- to 7-day hospitalization, such as laparoscopic-assisted colectomy, are less likely to deliver a lower bottom line than their open surgery counterparts. Nonetheless, with responsible use of disposable instrumentation and a commitment to the most effective use of the inpatient setting, most laparoscopic procedures can be made less expensive than their conventional equivalents.

**Education and Skill Acquisition**

Historically, surgeons in training (residents, registrars, and fellows) acquired their skills in minimally invasive techniques through a series of operative experiences of graded complexity. This training occurred on patients. Although such a paradigm did not compromise patient safety, learning in the OR is costly. In addition, the recent worldwide constraint placed on resident work hours makes it attractive to teach laparoscopic skills outside of the OR.

Skills labs started at nearly every surgical training center in the 1990s with low fidelity box-type trainers. These were rudimentary simulated abdominal cavities with a video camera, monitor, trocars, laparoscopic instruments, and target models. These targets were often as simple as a pegboard and rubber rings, or a latex drain to practice suturing and knot tying. Virtual reality training devices present a unique opportunity to improve and enhance experiential learning in endoscopy and laparoscopy.
for all surgeons. This technology has the advantage of enabling objective measurement of psychomotor skills, which can be used to
determine progress in skill acquisition and, ultimately, technical competency. Several of these devices have been validated as a means of measuring proficiency in skill performance. More
importantly, training on virtual reality platforms has proven to translate to improved operative performance in randomized trials. Currently, surgical skills labs are mandatory for Residency Review Committee credentialing. Successful completion of the Fundamentals of Laparoscopic Surgery (FLS) technical and cognitive examination became a mandatory prerequisite for the American Board of Surgery (ABS) qualification examination in general surgery in 2010. The Fundamentals of Endoscopic Surgery (FES) became a prerequisite to ABS qualification in 2015. In the future, institutions may require simulator training to document specific entrustable professional activities (EPA) related to laparoscopic procedures before privileging in the OR. A Fundamentals of Robotic Surgery (FRS) high stakes exam is on the horizon for future surgical trainees. The American College of Surgeons has taken a leadership position in accrediting skills labs across the world as American College of Surgeons–accredited educational institutes.

**Telementoring**

In response to the Institute of Medicine’s call for the development of unique technologic solutions to deliver health care to rural and underserved areas, surgeons are beginning to explore the feasibility of telementoring. Teleconsultation or telementoring is two-way audio and visual communication between two geographically separated providers. This communication can take place in the office setting or directly in the OR when complex scenarios are encountered. Although local communication channels may limit its performance in rural areas, the technology is available and currently is being used, especially in states and provinces with large geographically remote populations.

**Innovation and Introduction of New Procedures**

The revolution in minimally invasive general surgery, which occurred in 1990, created ethical challenges for the profession. The problem was this: If competence is gained from experience, how was the surgeon to climb the competence curve (otherwise known as the learning curve) without injuring patients? If it was indeed impossible to achieve competence without making mistakes along the way, how should one effectively communicate this to patients such that they understand the weight of their decisions? Even more fundamentally important is determining the path that should be followed before one recruits the first patient for a new procedure.

Although procedure development is fundamentally different than drug development (i.e., there is great individual variation in the performance of procedures, but no difference between one tablet and the next), adherence to a process similar to that used to develop a new drug is a reasonable path for a surgical innovator. At the outset, the surgeon must identify the problem that is not solved with current surgical procedures. For example, although the removal of a gallbladder through a Kocher incision is certainly effective, it creates a great deal of disability, pain, and scarification. As a result of those issues, many patients with very symptomatic biliary colic delayed operation until life-threatening complications occurred. Clearly, there was a need for developing a less invasive approach (Fig. 14-30).

Once the opportunity has been established, the next step involves a search through other disciplines for technologies and techniques that might be applied. Again, this is analogous to the drug industry, where secondary drug indications have often turned out to be more therapeutically important than the primary indication for drug development. The third step is in vivo studies in the most appropriate animal model. These types of studies are controversial because of the resistance to animal experimentation, and yet without such studies, many humans would be injured or killed during the developmental phase of medical drugs, devices, and techniques. These steps often are called the preclinical phase of procedure development.

The decision as to when such procedures are ready to come out of the lab is a difficult one. Put simply, the procedure should be reproducible, provide the desired effect, and not have serious side effects. Once these three criteria are reached, the time for human application has arrived. Before the surgeon discusses the new procedure with patients, it is important to achieve full institutional support. Involvement of the medical board, the chief of the medical staff, and the institutional review board is essential before commencing on a new procedure. These bodies are responsible for the use of safe, high-quality medical practices within their institution, and they will demand that great caution and all possible safeguards are in place before proceeding.

The dialogue with the patient who is to be first must be thorough, brutally honest, and well documented. The psychology

![Figure 14-30](image-url) The progress of general surgery can be reflected by a series of performance curves. General anesthesia and sterile technique allowed the development of maximally invasive open surgery over the last 125 years. Video optics allowed the development of minimally invasive surgery over the last 25 years. Noninvasive (seamless) surgery will result when a yet undiscovered transformational event allows surgery to occur without an incision, and perhaps without anesthesia.
that allows a patient to decide to be first is quite interesting, and may, under certain circumstances, require psychiatric evaluation. Certainly, if a dying cancer patient has a chance with a new drug, this makes sense. Similarly, if the standard surgical procedure has a high attendant morbidity and the new procedure offers a substantially better outcome, the decision to be first is understandable. On the other hand, when the benefits of the new approach are small and the risks are largely unknown, a more complete psychological profile may be necessary before proceeding.

For new surgical procedures, it generally is wise to assemble the best possible operative team, including a surgeon experienced with the old technique, and assistants who have participated in the earlier animal work. This initial team of experienced physicians and nurses should remain together until full competence with the procedure is attained. This may take 10 procedures, or it may take 50 procedures. The team will know that it has achieved competence when the majority of procedures take the same length of time and the team is relaxed and sure of the flow of the operation. This will complete phase I of the procedure development.

In phase II, the efficacy of the procedure is tested in a nonrandomized fashion. Ideally, the outcome of new techniques must be as good as or better than the procedure that is being replaced. This phase should occur at several medical centers to prove that good outcomes are achievable outside of the pioneering institution. These same requirements may be applied to the introduction of new technology into the OR. The value equation requires that the additional measurable procedure quality exceeds the additional measurable cost to the patient or healthcare system. In phase III, a randomized trial pits the new procedure against the old.

Once the competence curve has been climbed, it is appropriate for the team to engage in the education of others. During the ascension of the competence curve, other learners in the institution (i.e., surgical residents) may not have the opportunity to participate in the first case series. Although this may be difficult for them, the best interest of the patient must be put before the education of the resident.

The second stage of learning occurs when the new procedure has proven its value and a handful of experts exist, but the majority of surgeons have not been trained to perform the new procedure. In this setting, it is relatively unethical for surgeons to forge ahead with a new procedure in humans as if they had spent the same amount of time in intensive study that the first team did. The fact that one or several surgical teams were able to perform an operation does not ensure that all others with the same medical degrees can perform the operation with equal skill. It behooves the learners to contact the experts and request their assistance to ensure an optimal outcome at the new center. Although it is important that the learners contact the experts, it is equally important that the experts be willing to share their experience with their fellow professionals. As well, the experts should provide feedback to the learners as to whether they feel the learners are equipped to forge ahead on their own. If not, further observation and assistance from the experts are required. Although this approach may sound obvious, it is fraught with difficulties. In many situations, ego, competitiveness, and monetary concerns have short-circuited this process and led to poor patient outcomes. To a large extent, MIS has recovered from the black eye it received early in development, when inadequately trained surgeons caused an excessive number of significant complications.

If innovative procedures and technologies are to be developed and applied without the mistakes of the past, surgeons must be honest when they answer these questions: Is this procedure safe? Would I consider undergoing this procedure if I developed a surgical indication? Is the procedure as good as or better than the procedure it is replacing? Do I have the skills to apply this procedure safely and with equivalent results to the more experienced surgeon? Answering these questions in the affirmative should be a professional obligation. A negative response should motivate the surgeon to seek an alternative procedure or outside assistance before subjecting a patient to the new procedure.

REFERENCES
Entries highlighted in bright blue are key references.

70. Aiono S, Gilbert JM, Soin B, Fainay PA, Gordan A. Controlled trial of the introduction of a robotic camera assistant (EndoAssist) for laparoscopic cholecystectomy. Surg Endosc. 2002;16(9):1267-1270.
Chapter 15

Molecular Biology, The Atomic Theory of Disease, and Precision Surgery
Xin-Hua Feng, Xia Lin, Xinran Li, Juehua Yu, John Nemunaitis, and F. Charles Brunicardi

OVERVIEW OF MOLECULAR CELL BIOLOGY

The beginning of modern medicine can be traced back to centuries ago when physicians and scientists began studying human anatomy from cadavers in morgues and animal physiology following hunting expeditions. Gradually, from the study of animals and plants in greater detail and the discovery of microbes, scientific principles governing life led to the emergence of the biologic sciences. As biologic science developed and expanded, scientists and physicians began to utilize its principles to solve challenges of human diseases while continuing to explore the fundamentals of life in greater detail. With ever-evolving state-of-the-art scientific tools, our understanding of how cells, tissues, organs, and entire organisms function, down to the level of molecular and subatomic structure, has resulted in modern biology with an enormous impact on modern healthcare and the discovery of amazing treatments for disease at an exponential pace. Significant progress has been made in molecular studies of organ development, cell signaling, and gene regulation. The advent of recombinant DNA technology, polymerase chain reaction (PCR) techniques, and next-generation genomic sequencing, which resulted in the sequencing of the human genome, holds the potential to have a transformational influence on healthcare and society this century by not only broadening our understanding of the pathophysiology of disease, but also by bringing about necessary changes in personalized medicine.

Today’s practicing surgeons are becoming increasingly aware that many modern surgical procedures rely on the information gained through molecular research (i.e., precision surgery). Genomic information, such as deleterious *BRCA* and *RET* proto-oncogene mutations, is being used to help direct prophylactic procedures to remove potentially harmful tissues before they do damage to patients. Molecular engineering has led to cancer-specific gene therapy that could serve in the near future as a more effective adjunct to surgical debulking of tumors than radiation or chemotherapy, so surgeons will benefit from a clear introduction to how basic biochemical and biologic principles relate to the developing area of molecular biology. This chapter reviews the current information on modern molecular biology for the surgical community.

Basic Concepts of Molecular Research

The modern era of molecular biology, which has been mainly concerned with how genes govern cell activity, began in 1953 when James D. Watson and Francis H. C. Crick made one of the greatest scientific discoveries by deducing the double-helical structure of deoxyribonucleic acid (DNA).1,2 The year 2003 marked the 50th anniversary of this great discovery. In the same year, the Human Genome Project completed with sequencing approximately 20,000 to 25,000 genes and 3 billion base pairs in human DNA.3 Before 1953, one of the most mysterious aspects of biology was how genetic material was precisely duplicated from one generation to the next. Although DNA had been implicated as genetic material, it was given rise to the notion that a template was involved in the transfer of information between generations, and thus confirmed the suspicion that DNA carried an organism’s hereditary information.

Within cells, DNA is packed tightly into chromosomes. One important feature of DNA as genetic material is its ability to encode important information for all of a cell’s functions (Fig. 15-1). Based on the principles of base complementarity, scientists also discovered how information in DNA is accurately transferred into the protein structure. DNA serves as a template for RNA synthesis, termed *transcription*, including messenger RNA (mRNA, or the protein-encoding RNA), ribosomal RNA (rRNA), and transfer RNA (tRNA). mRNA carries the information from DNA to make proteins, termed *translation*, with the assistance of rRNA and tRNA. Each of these steps is precisely
controlled in such a way that genes are properly expressed in each cell at a specific time and location. In recent years, new classes of noncoding RNAs (ncRNA), for example, microRNA (or miRNA), piwi-interacting RNA (or piRNA), and long intergenic noncoding RNA (or lincRNA), have been identified. Although the number of ncRNAs encoded in the human genome is unknown and a lot of ncRNAs have not been validated for their functions, ncRNAs have been associated to regulate gene expression through posttranscriptional gene regulation such as mRNA degradation or epigenetic regulation such as chromatin structure modification and DNA methylation induction. Consequently, the differential gene activity in a cell determines its actions, properties, and functions.

**Molecular Approaches to Surgical Research**

Rapid advances in molecular and cellular biology over the past half century have revolutionized the understanding of disease and will radically transform the practice of surgery. In the future, molecular techniques will be increasingly applied to surgical disease and will lead to new strategies for the selection and implementation of operative therapy. Surgeons should be familiar with the fundamental principles of molecular and cellular biology so that emerging scientific breakthroughs can be translated into improved care of the surgical patient.

The greatest advances in the field of molecular biology have been in the areas of analysis and manipulation of DNA. Since Watson and Crick’s discovery of DNA structure, an intensive effort has been made to unlock the deepest biologic secrets of DNA. Among the avalanche of technical advances, one discovery in particular has drastically changed the world of molecular biology: the uncovering of the enzymatic and microbiologic techniques that produce recombinant DNA. Recombinant DNA technology involves the enzymatic manipulation of DNA and, subsequently, the cloning of DNA. DNA molecules are cloned for a variety of purposes including safeguarding DNA samples, facilitating sequencing, generating probes, and expressing recombinant proteins in one or more host organisms. DNA can be produced by a number of means, including restricted digestion of an existing vector, PCR, and cDNA synthesis. As DNA cloning techniques have developed over the last quarter century, researchers have moved from studying DNA to studying the functions of proteins, and from cell and animal models to molecular therapies in humans. Expression of recombinant proteins provides a method for analyzing gene regulation, structure, and function. In recent years, the uses for recombinant proteins have expanded to include a variety of new applications, including gene therapy and biopharmaceuticals. The basic molecular approaches for modern surgical research include DNA cloning, cell manipulation, disease modeling in animals, and clinical trials in human patients.

**FUNDAMENTALS OF MOLECULAR AND CELL BIOLOGY**

**DNA and Heredity**

DNA forms a right-handed, double-helical structure that is composed of two antiparallel strands of unbranched polymeric deoxyribonucleotides linked by phosphodiester bonds between the 5′ carbon of one deoxyribose moiety to the 3′ carbon of the next (Fig. 15-2). DNA is composed of four types of deoxyribonucleotides: adenine (A), cytosine (C), guanine (G), and thymine (T). The nucleotides are joined together by phosphodiester bonds. In the double-helical structure deduced by Watson and Crick, the two strands of DNA are complementary to each other.

**Figure 15-1.** The flow of genetic information from DNA to protein to cell functions. The process of transmission of genetic information from DNA to RNA is called transcription, and the process of transmission from RNA to protein is called translation. Proteins are the essential controlling components for cell structure, cell signaling, and metabolism. Genomics and proteomics are the study of the genetic composition of a living organism at the DNA and protein level, respectively. The study of the relationship between genes and their cellular functions is called functional genomics.
Building blocks of DNA

Sugar phosphate Base Nucleotide DNA strand

Sugar Phosphate

Double-stranded DNA DNA double helix

Hydrogen-bonded pairs

A-T
G-C
C-G
T-A

Sugar-phosphate backbone

Figure 15-2. Schematic representation of a DNA molecule forming a double helix. DNA is made up of four types of nucleotides, which are linked covalently into a DNA strand. A DNA molecule is composed of two DNA strands held together by hydrogen bonds between the pair bases. The arrowheads at the ends of the DNA strands indicate the polarities of the two strands, which run antiparallel to each other in the DNA molecule. The diagram at the bottom left of the figure shows the DNA molecule straightened out. In reality, the DNA molecule is twisted into a double helix, of which each turn of DNA is made up of 10.4 nucleotide pairs, as shown on the right.

Because of size, shape, and chemical composition, A always pairs with T, and C with G, through the formation of hydrogen bonds between complementary bases that stabilize the double helix.

Recognition of the hereditary transmission of genetic information is attributed to the Austrian monk, Gregor Mendel. His seminal work, ignored upon publication until its rediscovery in 1900, established the laws of segregation and of independent assortment. These two principles established the existence of paired elementary units of heredity and defined the statistical laws that govern them. DNA was isolated in 1869, and a number of important observations of the inherited basis of certain diseases were made in the early part of the 20th century. Although today it appears easy to understand how DNA replicates, before the 1950s the idea of DNA as the primary genetic material was not appreciated. The modern era of molecular biology began in 1944 with the demonstration that DNA was the substance that carried genetic information. The first experimental evidence that DNA was genetic material came from simple transformation experiments conducted in the 1940s using Streptococcus pneumoniae. One strain of the bacteria could be converted into another by incubating it with DNA from the other, just as the treatment of the DNA with deoxyribonuclease would inactivate the transforming activity of the DNA. Similarly, in the early 1950s, before the discovery of the double-helical structure of DNA, the entry of viral DNA and not the protein into the host bacterium was believed to be necessary to initiate infection by the bacterial virus or bacteriophage. Key historical events concerning genetics are outlined in Table 15-1.

For cells to pass on the genetic material (DNA) to each progeny, the amount of DNA must be doubled. Watson and Crick recognized that the complementary base-pair structure of DNA implied the existence of a template-like mechanism for the copying of genetic material. The transfer of DNA material from the mother cell to daughter cells takes place during somatic cell division (also called mitosis). Before a cell divides, DNA must be precisely duplicated. During replication, the two strands of DNA separate, and each strand creates a new complementary strand by precise base-pair matching (Fig. 15-3). The two, new, double-stranded DNAs carry the same genetic information, which can then be passed on to two daughter cells. Proofreading mechanisms ensure that the replication process occurs in a highly accurate manner. The fidelity of DNA replication is absolutely crucial to maintaining the integrity of the genome from generation to generation. However, mistakes can still occur during this process, resulting in mutations, which may lead to a change of the DNA’s encoded protein and, consequently, a change of the cell’s behavior. The reliable dependence of many features of modern organisms on subtle changes in genome is linked to Mendelian inheritance and also contributes to the processes of Darwinian evolution. In addition, massive changes, so-called genetic instability, can occur in the genome of somatic cells such as cancer cells.

Gene Regulation

Living cells have the necessary machinery to enzymatically transcribe DNA into RNA and translate the mRNA into protein. This machinery accomplishes the two major steps required for gene expression in all organisms: transcription and translation (Fig. 15-4). However, gene regulation is far more complex, particularly in eukaryotic organisms. For example, many gene transcripts must be spliced to remove the intervening sequences. The sequences that are spliced off are called introns, which appear to be useless, but in fact may carry some regulatory information. The sequences that are joined together, and are eventually translated into protein, are called exons. Additional regulation of gene expression includes modification of mRNA, control of mRNA stability, and its nuclear export into cytoplasm (where it is assembled into ribosomes for translation). After mRNA is translated into protein, the levels and functions of the proteins can be further regulated posttranslationally. However, the following sections will mainly focus on gene regulation at transcriptional and translational levels.

Transcription. Transcription is the enzymatic process of RNA synthesis from DNA. In bacteria, a single RNA polymerase carries out all RNA synthesis, including that of mRNA, rRNA,
Figure 15-3. DNA replication. As the nucleotide A only pairs with T, and G with C, each strand of DNA can determine the nucleotide sequence in its complementary strand. In this way, double-helical DNA can be copied precisely.
MOLECULAR BIOLOGY, THE ATOMIC THEORY OF DISEASE, AND PRECISION SURGERY

CHAPTER 15

Nucleus Protein

DNA

Nuclear envelope

Figure 15-4. Four major steps in the control of eukaryotic gene expression. Transcriptional and posttranscriptional control determine the level of messenger RNA (mRNA) that is available to make a protein, while translational and posttranslational control determine the final outcome of functional proteins. Note that posttranscriptional and posttranslational controls consist of several steps.

and tRNA. Transcription often is coupled with translation in such a way that an mRNA molecule is completely accessible to ribosomes, and bacterial protein synthesis begins on an mRNA molecule even while it is still being synthesized. Therefore, a discussion of gene regulation with a look at the simpler prokaryotic system precedes that of the more complex transcription and posttranscriptional regulation of eukaryotic genes.

Transcription in Bacteria Initiation of transcription in prokaryotes begins with the recognition of DNA sequences by RNA polymerase. First, the bacterial RNA polymerase catalyzes RNA synthesis through loose binding to any region in the double-stranded DNA and then through specific binding to the promoter region with the assistance of accessory proteins called σ factors (sigma factors). A promoter region is the DNA region upstream of the transcription initiation site. RNA polymerase binds tightly at the promoter sites and causes the double-stranded DNA structure to unwind. Consequently, few nucleotides can be base-paired with the DNA template to begin transcription. Once transcription begins, the σ factor is released. The growing RNA chain may begin to peel off as the chain elongates. This occurs in such a way that there are always about 10 to 12 nucleotides of the growing RNA chains that are base-paired with the DNA template.

The bacterial promoter contains a region of about 40 bases that include two conserved elements called −35 region and −10 region. The numbering system begins at the initiation site, which is designated +1 position, and counts backward (in negative numbers) on the promoter and forward on the transcribed region. Although both regions on different promoters are not the same sequences, they are fairly conserved and very similar. This conservation provides the accurate and rapid initiation of transcription for most bacterial genes. It is also common in bacteria that one promoter serves to transcribe a series of clustered genes, called an operon. A single transcribed mRNA contains a series of coding regions, each of which is later independently translated. In this way, the protein products are synthesized in a coordinated manner. Most of the time, these proteins are involved in the same metabolic pathway, thus demonstrating that the control by one operon is an efficient system. After initiation of transcription, the polymerase moves along the DNA to elongate the chain of RNA, although at a certain point, it will stop. Each step of RNA synthesis, including initiation, elongation, and termination, will require the integral functions of RNA polymerase as well as the interactions of the polymerase with regulatory proteins.

Transcription in Eukaryotes Transcription mechanisms in eukaryotes differ from those in prokaryotes. The unique features of eukaryotic transcription are as follows: (a) Three separate RNA polymerases are involved in eukaryotes: RNA polymerase I transcribes the precursor of 5.8S, 18S, and 28S rRNAs; RNA polymerase II synthesizes the precursors of mRNA as well as microRNA; and RNA polymerase III makes tRNAs and 5S rRNAs. (b) In eukaryotes, the initial transcript is often the precursor to final mRNAs, tRNAs, and rRNAs. The precursor is then modified and/or processed into its final functional form. RNA splicing is one type of processing to remove the noncoding introns (the region between coding exons) on an mRNA. (c) In contrast to bacterial DNA, eukaryotic DNA often is packaged with histone and nonhistone proteins into chromatin. Transcription will only occur when the chromatin structure changes in such a way that DNA is accessible to the polymerase. (d) RNA is made in the nucleus and transported into cytoplasm, where translation occurs. Therefore, unlike bacteria, eukaryotes undergo uncoupled transcription and translation.

Eukaryotic gene transcription also involves the recognition and binding of RNA polymerase to the promoter DNA. However, the interaction between the polymerase and DNA is far more complex in eukaryotes than in prokaryotes. Because the majority of studies have been focused on the regulation and functions of proteins, this chapter primarily focuses on how protein-encoding mRNA is made by RNA polymerase II.

Translation. DNA directs the synthesis of RNA; RNA in turn directs the synthesis of proteins. Proteins are variable-length polypeptide polymers composed of various combinations of 20 different amino acids and are the working molecules of the cell. The process of decoding information on mRNA to synthesize proteins is called translation (see Fig. 15-1). Translation takes place in ribosomes composed of rRNA and ribosomal proteins. The numerous discoveries made during the 1950s made it easy to understand how DNA replication and transcription involve base-pairing between DNA and DNA or DNA and RNA. However, at that time, it was still impossible to comprehend how mRNA transfers the information to the protein-synthesizing machinery. The genetic information on mRNA is composed of
arranged sequences of four bases that are transferred to the linear arrangement of 20 amino acids on a protein. Amino acids are characterized by a central carbon unit linked to four side chains: an amino group (–NH₂), a carboxy group (–COOH), a hydrogen, and a variable (–R) group. The amino acid chain is assembled via peptide bonds between the amino group of one amino acid and the carboxy group of the next. Because of this decoding, the information carried on mRNA relies on tRNA. Translation involves all three RNAs. The precise transfer of information from mRNA to protein is governed by \textit{genetic code}, the set of rules by which codons are translated into an amino acid (Table 15-2). A \textit{codon}, a triplet of three bases, codes for one amino acid. In this case, random combinations of the four bases form \(4 \times 4 \times 4\), or 64, because 64 codes are more than enough for 20 amino acids, most amino acids are coded by more than one codon. The start codon is AUG, which also corresponds to methionine; therefore, almost all proteins begin with this amino acid. The sequence of nucleotide triplets that follows the start codon signal is termed the \textit{reading frame}. The codons on mRNA are sequentially recognized by tRNA adaptor proteins. Specific enzymes termed \textit{aminoacyl-tRNA synthetases} link a specific amino acid to a specific tRNA. The translation of mRNA to protein requires the ribosomal complex to move stepwise along the mRNA until the initiator methionine sequence is identified. In concert with various protein initiator factors, the methionyl-tRNA is positioned on the mRNA and protein synthesis begins. Each new amino acid is added sequentially by the appropriate tRNA in conjunction with proteins called \textit{elongation factors}. Protein synthesis proceeds in the amino-to-carboxy-terminus direction.

The biologic versatility of proteins is astounding. Among many other functions, proteins serve as enzymes that catalyze critical biochemical reactions, carry signals to and from the extracellular environment, and mediate diverse signaling and regulatory functions in the intracellular environment. They also transport ions and various small molecules across plasma membranes. Proteins make up the key structural components of cells and the extracellular matrix and are responsible for cell motility. The unique functional properties of proteins are largely determined by their structure (Fig. 15-5).

\textbf{Regulation of Gene Expression.} The human organism is made up of a myriad of different cell types that, despite their vastly different characteristics, contain the same genetic material. This cellular diversity is controlled by the \textit{genome} and is accomplished by tight regulation of gene expression. This leads to the synthesis and accumulation of different complements of RNA and, ultimately, to the proteins found in different cell types. For example, muscle and bone express different genes or the same genes at different times. Moreover, the choice of which genes are expressed in a given cell at a given time depends on signals received from its environment. There are multiple levels at which gene expression can be controlled along the pathway from DNA to RNA to protein (see Fig. 15-4). \textit{Transcriptional control} refers to the mechanism for regulating when and how often a gene is transcribed. Splicing of the primary RNA transcript (\textit{RNA processing control}) and selection of completed

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|c|c|c|c|}
\hline
\textbf{Secon BASE IN CODON} & \textbf{U} & \textbf{C} & \textbf{A} & \textbf{G} & \textbf{U} & \textbf{A} & \textbf{G} & \textbf{C} & \textbf{O} & \textbf{D} \\
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
AUC & Ile & [I] & ACC & Thr & [T] & AAC & Asn & [N] & AGG & Arg & [R] & G \\
\hline
\hline
\hline
\hline
\hline
\hline
\end{tabular}
\caption{The genetic code}
\end{table}

\begin{tabular}{|c|c|c|c|c|c|c|c|c|c|c|}
\hline
\textbf{Second Base in Codon} & \textbf{U} & \textbf{C} & \textbf{A} & \textbf{G} & \textbf{U} & \textbf{A} & \textbf{G} & \textbf{C} & \textbf{O} & \textbf{D} \\
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
AUC & Ile & [I] & ACC & Thr & [T] & AAC & Asn & [N] & AGG & Arg & [R] & G \\
\hline
\hline
\hline
\hline
\hline
\hline
\end{tabular}

A = adenine; C = cytosine; G = guanine; U = uracil; Ala = alanine; Arg = arginine; Asn = asparagine; Asp = aspartic acid; Cys = cysteine; Glu = glutamic acid; Gln = glutamine; Gly = glycine; His = histidine; Ile = isoleucine; Leu = leucine; Lys = lysine; Met = methionine; Phe = phenylalanine; Pro = proline; Ser = serine; Thr = threonine; Trp = tryptophan; Tyr = tyrosine; Val = valine. Letter in [ ] indicates single letter code for amino acid.
mRNAs for nuclear export (RNA transport control) represent additional potential regulatory steps. The mRNAs in the cytoplasm can be selectively translated by ribosomes (translational control) or selectively stabilized or degraded (mRNA degradation control). Finally, the resulting proteins can undergo selective activation, inactivation, or compartmentalization (protein activity control).

Because a large number of genes are regulated at the transcriptional level, regulation of gene transcripts (i.e., mRNA) often is referred to as gene regulation in a narrow definition. Each of the steps during transcription is properly regulated in eukaryotic cells. Because genes are differentially regulated from one another, one gene can be differentially regulated in different cell types or at different developmental stages. Therefore, gene regulation at the level of transcription is largely context dependent. However, there is a common scheme that applies to transcription at the molecular level (Fig. 15-6). Each gene promoter possesses unique sequences called TATA boxes that can be recognized and bound by a large complex containing RNA polymerase II, forming the basal transcription machinery. Usually located upstream of the TATA box (but sometimes longer distances) are a number of regulatory sequences referred to as enhancers that are recognized by regulatory proteins called transcription factors. These transcription factors specifically bind to the enhancers, often in response to environmental or developmental cues, and cooperate with each other and with basal transcription factors to initiate transcription. Regulatory sequences that negatively regulate the initiation of transcription also are present on the promoter DNA. The transcription factors that bind to these sites are called repressors, in contrast to the activators that activate transcription. The molecular interactions between transcription factors and promoter DNA, as well as between the cooperative transcription factors, are highly regulated and context-dependent. Specifically, the recruitment of transcription factors to the promoter DNA occurs in response to physiologic signals. A number of structural motifs in these DNA-binding transcription factors facilitate this recognition and interaction. These include the helix-turn-helix, the homeodomain motif, the zinc finger, the leucine zipper, and the helix-loop-helix motifs.

Human Genome

Genome is a collective term for all genes present in one organism. The human genome contains DNA sequences of 3 billion base pairs, carried by 23 pairs of chromosomes. The human genome has an estimated 25,000 to 30,000 genes, and overall it is 99.9% identical in all people. Approximately 3 million locations where single-base DNA differences exist have been identified and termed single nucleotide polymorphisms. Single nucleotide polymorphisms may be critical determinants of human variation in disease susceptibility and responses to environmental factors.

The completion of the human genome sequence in 2003 represented another great milestone in modern science. The Human Genome Project created the field of genomics, which is the study of genetic material in detail (see Fig. 15-1). The medical field is building on the knowledge, resources, and technologies emanating from the human genome to further the understanding of the relationship of the genes and their mutations to human health and disease. This expansion of genomics into human health applications resulted in the field of genomic medicine.

The emergence of genomics as a science will transform the practice of medicine and surgery in this century. This breakthrough has allowed scientists the opportunity to gain remarkable insights into the lives of humans. Ultimately, the goal is to use this information to develop new ways to treat, cure, or even prevent the thousands of diseases that afflict humankind. In the 21st century, work will begin to incorporate the information embedded in the human genome sequence into surgical practices. By doing so, the genomic information can be used for diagnosing and predicting disease and disease susceptibility. Diagnostic tests can be designed to detect errant genes in patients suspected of having particular diseases or of being at risk for developing them. Furthermore, exploration into the function of each human gene is now possible, which will shed
light on how faulty genes play a role in disease causation. This knowledge also makes possible the development of a new generation of therapeutics based on genes. Drug design is being revolutionized as researchers create new classes of medicines based on a reasoned approach to the use of information on gene sequence and protein structure function rather than the traditional trial-and-error method. Drugs targeted to specific sites in the body promise to have fewer side effects than many of today’s medicines. Finally, other applications of genomics will involve the transfer of genes to replace defective versions or the use of gene therapy to enhance normal functions such as immunity.

Proteomics refers to the study of the structure and expression of proteins as well as the interactions among proteins encoded by a human genome (see Fig. 15-1). A number of Internet-based repositories for protein sequences exist, including Swiss-Prot (www.expasy.ch). These databases allow comparisons of newly identified proteins with previously characterized sequences to allow prediction of similarities, identification of splice variants, and prediction of membrane topology and posttranslational modifications. Tools for proteomic profiling include two-dimensional gel electrophoresis, time-of-flight mass spectrometry, matrix-assisted laser desorption/ionization, and protein microarrays. Structural proteomics aims to describe the three-dimensional structure of proteins that is critical to understanding function. Functional genomics seeks to assign a biochemical, physiologic, cell biologic, and/or developmental function to each predicted gene. An ever-increasing arsenal of approaches, including transgenic animals, RNA interference (RNAi), and various systematic mutational strategies, will allow dissection of functions associated with newly discovered genes. Although the potential of this field of study is vast, it is in its early stages.

It is anticipated that a genomic and proteomic approach to human disease will lead to a new understanding of pathogenesis that will aid in the development of effective strategies for early diagnosis and treatment. For example, identification of altered protein expression in organs, cells, subcellular structures, or protein complexes may lead to development of new biomarkers for disease detection. Moreover, improved understanding of how protein structure determines function will allow rational design of drugs for therapeutic efficacy and potential toxicity. A genomic and proteomic approach to human disease will lead to a new understanding of pathogenesis that will aid in the development of effective strategies for early diagnosis and treatment. For example, identification of altered protein expression in organs, cells, subcellular structures, or protein complexes may lead to development of new biomarkers for disease detection. Moreover, improved understanding of how protein structure determines function will allow rational design of drugs for therapeutic efficacy and potential toxicity.

Cell Cycle and Apoptosis

Every organism is composed of many different cell types at different developmental stages. Some cell types continue to grow, while some cells stop growing after a developmental stage or resume growth after a break. For example, embryonic stem cells grow continuously, while nerve cells and striated muscle cells stop dividing after maturation. Cell cycle is the process for every cell including DNA replication and protein synthesis, DNA segregation in half, and package DNA and protein in two newly formed cells to enable passage of identical genetic information from one parental cell to two daughter cells. Thus, the cell cycle is the fundamental mechanism to maintain tissue homeostasis. A cell cycle comprises four periods: G1 (first gap phase before DNA synthesis), S (synthesis phase when DNA replication occurs), G2 (the gap phase before mitosis), and M (mitosis, the phase when two daughter cells with identical DNA are generated) (Fig. 15-7). After a full cycle, the daughter cells enter G1 again, and when they receive appropriate signals, undergo another cycle, and so on. The machinery that drives cell cycle progression is made up of a group of enzymes called cyclin-dependent kinases (CDKs). Cyclin expression fluctuates during the cell cycle, and cyclins are essential for CDK activities and form complexes with CDK. The cyclin A/CDK1 and cyclin B/CDK1 drive the progression for the M phase, while cyclin A/CDK2 is the primary S phase complex. Early G1, cyclin D/CDK4/6 or late G1, cyclin E/CDK2 controls the G1-S transition. There also are negative regulators for CDK termed CDK inhibitors, which inhibit the assembly or activity of the cyclin-CDK complex. Expression of cyclins and CDK inhibitors often is regulated by developmental and environmental factors.

The cell cycle is connected with signal transduction pathways as well as gene expression. Although the S and M phases rarely are subjected to changes imposed by extracellular signals, the G1 and G2 phases are the primary periods when cells decide whether or not to move on to the next phase. During the G1 phase, cells receive green- or red-light signals, S phase entry or G1 arrest, respectively. Growing cells proliferate only when supplied with appropriate mitogenic growth factors. Cells become committed to entry of the cell cycle only toward the end of G1. Mitogenic signals stimulate the activity of early G1 CDKs (e.g., cyclin D/CDK4) that inhibit the activity of pRb protein and activate the transcription factor called E2F to induce the expression of batteries of genes essential for G1-S progression. Meanwhile, cells also receive antiproliferative signals such as those from tumor suppressors. These antiproliferative signals also act in the G1 phase to stop cells’ progress into the S phase by inducing CKI production. For example, when DNA is damaged, cells will repair the damage before entering the S phase. Therefore, G1 contains one of the most important checkpoints for cell cycle progression. If the analogy is made that CDK is to a cell as an engine is to a car, then cyclins and CKI are the gas pedal and brake, respectively. Accelerated proliferation or
improper cell cycle progression with damaged DNA would be disastrous. Genetic gain-of-function mutations in oncogenes (that often promote expression or activity of the cyclin/CDK complex) or loss-of-function mutations in tumor suppressor (that stimulate production of CKI) are causal factors for malignant transformation.

In addition to cell cycle control, cells use genetically programmed mechanisms to kill cells. This cellular process, called apoptosis or programmed cell death, is essential for the maintenance of tissue homeostasis (Fig. 15-8).

Normal tissues undergo proper apoptosis to remove unwanted cells, those that have completed their jobs or have been damaged or improperly proliferated. Apoptosis can be activated by many physiologic stimuli such as death receptor signals (e.g., Fas or cytokine tumor necrosis factor), growth factor deprivation, DNA damage, and stress signals. Two major pathways control the biochemical mechanisms governing apoptosis: the death receptor and mitochondrial. However, recent advances in apoptosis research suggest an interconnection of the two pathways. What is central to the apoptotic machinery is the activation of a cascade of proteinases called caspases. Similar to CDK in the cell cycle, activities and expression of caspases are well controlled by positive and negative regulators. The complex machinery of apoptosis must be tightly controlled. Perturbations of this process can cause neoplastic transformation or other diseases.

**Signal Transduction Pathways**

Gene expression in a genome is controlled in a temporal and spatial manner, at least in part by signaling pathways. A signaling pathway generally begins at the cell surface and, after a signaling relay by a cascade of intracellular effectors, ends up in the nucleus (Fig. 15-9). All cells have the ability to sense changes in their external environment. The bioactive substances to which cells can respond are many and include proteins, short peptides, amino acids, nucleotides/nucleosides, steroids, retinoids, fatty acids, and dissolved gases. Some of these substances are lipophilic and thereby can cross the plasma membrane by
Control and specificity through simple protein-protein interactions—referred to as adhesions—is a common feature of signal transduction pathways in cells. Signaling also involves catalytic activities of signaling molecules, such as protein kinases/phosphatases, that modify the structures of key signaling proteins. Upon binding and/or modification by upstream signaling molecules, downstream effectors undergo a conformational (allosteric) change and, consequently, a change in function. The signal that originates at the cell surface and is relayed by the cytoplasmic proteins often ultimately reaches the transcriptional apparatus in the nucleus. It alters the DNA binding and activities of transcription factors that directly turn genes on or off in response to the stimuli. Abnormal alterations in signaling activities and capacities in otherwise normal cells can lead to diseases such as cancer.

Advances in biology in the last two decades have dramatically expanded the view on how cells are wired with signaling pathways. In a given cell, many signaling pathways operate simultaneously and crosstalk with one another. A cell generally may react to a hormonal signal in a variety of ways: (a) by changing its metabolite or protein, (b) by generating an electrical current, or (c) by contracting. Cells continually are subject to multiple input signals that simultaneously and sequentially activate multiple receptor- and non–receptor-mediated signal transduction pathways, which form a signaling network. Although the regulators responsible for cell behavior are rapidly identified as a result of genomic and proteomic techniques, the specific functions of the individual proteins, how they assemble, and the networks that control cellular behavior remain to be defined. An increased understanding of cell regulatory pathways—and how they are disrupted in disease—will likely reveal common themes based on protein interaction domains that direct associations of proteins with other polypeptides, phospholipids, nucleic acids, and other regulatory molecules. Advances in the understanding of signaling networks will require methods of investigation that move beyond traditional “linear” approaches into medical informatics and computational biology. The bewildering biocomplexity of such networks mandates multidisciplinary and transdisciplinary research collaboration. The vast amount of information that is rapidly emerging from genomic and proteomic data mining will require the development of new modeling methodologies within the emerging disciplines of medical mathematics and physics.

Signaling pathways often are grouped according to the properties of signaling receptors. Many hydrophobic signaling molecules are able to diffuse across plasma membranes and directly reach specific cytoplasmic targets. Steroid hormones, thyroid hormones, retinoids, and vitamin D are examples that exert their activity upon binding to structurally related receptor proteins that are members of the nuclear hormone receptor superfamily. Ligand binding induces a conformational change that enhances transcriptional activity of these receptors. Most extracellular signaling molecules interact with transmembrane protein receptors that couple ligand binding to intracellular signals, leading to biologic actions.

There are three major classes of cell-surface receptors: transmitter-gated ion channels, seven-transmembrane G-protein–coupled receptors (GPCRs), and enzyme-linked receptors. The superfamily of GPCRs is one of the largest families of proteins, representing over 800 genes of the human genome. Members of this superfamily share a characteristic seven-transmembrane configuration. The ligands for these receptors are diverse and include hormones, chemokines, neurotransmitters, proteinases, inflammatory mediators, and even sensory signals such as odorants and photons. Most GPCRs signal through heterotrimeric G proteins, which are guanine-nucleotide regulatory complexes. Thus, the receptor serves as the receiver, the G protein serves as the transducer, and the enzyme serves as the effector arm. Enzyme-linked receptors possess an extracellular ligand-recognition domain and a cytosolic domain that either has intrinsic enzymatic activity or directly links with an enzyme. Structurally, these receptors usually have only one transmembrane-spanning domain. Of at least five forms of enzyme-linked receptors classified by the nature of the enzyme activity to which they are coupled, the growth factor receptors such as tyrosine kinase receptor or serine/threonine kinase receptors mediate diverse cellular events including cell growth, differentiation, metabolism, and survival/apoptosis. Dysregulation (particularly mutations) of these receptors is thought to underlie conditions of abnormal cellular proliferation in the context of cancer. The following sections will further review two examples of growth factor signaling pathways and their connection with human diseases.

Insulin Pathway and Diabetes. The discovery of insulin in the early 1920s is one of the most dramatic events in the treatment of human disease. Insulin is a peptide hormone that is secreted by the β-cell of the pancreas. Insulin is required for the growth and metabolism of most mammalian cells, which contain cell-surface insulin receptors (InsR). Insulin binding to InsR activates the kinase activity of InsR. InsR then adds phosphoryl groups, a process referred to as phosphorylation, and subsequently activates its immediate intracellular effector, called insulin receptor substrate (IRS). IRS plays a central role in coordinating the signaling of insulin by activating distinct signaling pathways, the PI3K-Akt pathway and MAPK pathway, both of which possess multiple protein kinases that can control transcription, protein synthesis, and glycolysis (Fig. 15-10).

The primary physiologic role of insulin in glucose homeostasis, which is accomplished through the stimulation of glucose uptake into insulin-sensitive tissues such as fat and skeletal muscle. Defects in insulin synthesis/secretion and/or responsiveness are major causal factors in diabetes, one of the leading causes of death and disability in the United States, affecting an estimated 16 million Americans. Type 2 diabetes accounts for about 90% of all cases of diabetes. Clustering of type 2 diabetes in certain families and ethnic populations points to a strong genetic background for the disease. More than 90% of affected individuals have insulin resistance, which develops when the body is no longer able to respond correctly to insulin circulating in the blood. Although relatively little is known about the biochemical basis of this metabolic disorder, it is clear that the insulin-signaling pathways malfunction in this disease. It is also known that genetic mutations in the InsR or IRS cause type 2 diabetes, although which one is not certain. The majority of type 2 diabetes cases may result from defects in downstream-signaling components in the insulin-signaling pathway.
Figure 15-10. Insulin-signaling pathway. Insulin is a peptide growth factor that binds to and activates the heterotetrameric receptor complex (InsR). InsR possesses protein tyrosine kinase activity and is able to phosphorylate the downstream insulin receptor substrate (IRS). Phosphorylated IRS serves as a scaffold and controls the activation of multiple downstream pathways for gene expression, cell survival, and glucose metabolism. Inactivation of the insulin pathway can lead to type 2 diabetes.

Type 2 diabetes also is associated with declining β-cell function, resulting in reduced insulin secretion; these pathways are under intense study. A full understanding of the basis of insulin resistance is crucial for the development of new therapies for type 2 diabetes. Furthermore, apart from type 2 diabetes, insulin resistance is a central feature of several other common human disorders, including atherosclerosis and coronary artery disease, hypertension, and obesity.

**Transforming Growth Factor-β (TGF-β) Pathway and Cancers.** Growth factor signaling controls cell growth, differentiation, and apoptosis. Although insulin and many mitogenic growth factors promote cell proliferation, some growth factors and hormones inhibit cell proliferation. TGF-β is one of them. The balance between mitogens and TGF-β plays an important role in controlling the proper pace of cell cycle progression. The growth inhibition function of TGF-β signaling in epithelial cells plays a major role in maintaining tissue homeostasis.

The TGF-β superfamily comprises a large number of structurally related growth and differentiation factors that act through a receptor complex at the cell surface (Fig. 15-11). The complex consists of transmembrane serine/threonine kinases. The receptor signals through activation of heterotrimeric complexes of intracellular effectors called SMADs (which are contracted from homologous Caenorhabditis elegans Sma and Drosophila Mad, two evolutionarily conserved genes for TGF-β signaling). Upon phosphorylation by the receptors, SMAD complexes translocate into the nucleus, where they bind to gene promoters and cooperate with specific transcription factors to regulate the expression of genes that control cell proliferation and differentiation. For example, TGF-β strongly induces the transcription of a gene called p15INK4B (a type of CKI) and, at the same time, reduces the expression of many oncogenes such as c-Myc. The outcome of the altered gene expression leads to the inhibition of cell cycle progression. Meanwhile, the strength and duration of TGF-β signaling is fine-tuned by a variety of positive or negative modulators, including protein phosphatases. Therefore, controlled activation of TGF-β signaling is an intrinsic mechanism for cells to ensure controlled proliferation.

Resistance to TGF-β’s anticancer action is one hallmark of human cancer cells. TGF-β receptors and SMADs are identified as tumor suppressors. The TGF-β signaling circuit can be disrupted in a variety of ways and in different types of human tumors. Some lose TGF-β responsiveness through downregulation or mutations of their TGF-β receptors. The cytoplasmic SMAD4 protein, which transduces signals from ligand-activated TGF-β receptors to downstream targets, may be eliminated through mutation of its encoding gene. The locus encoding cell cycle inhibitor p15INK4B may be deleted. Alternatively, the immediate downstream target of its actions, cyclin-dependent kinase 4 (CDK4), may become unresponsive to the inhibitory actions of p15INK4B because of mutations that block p15INK4B binding. The resulting cyclin D/CDK4 complexes constitutively inactivate tumor suppressor pRb by hyperphosphorylation. Finally, functional pRb, the end target of this pathway, may be lost through mutation of its gene. For example, in pancreatic and colorectal cancers, 100% of cells derived from these cancers carry genetic defects in the TGF-β signaling pathway. Therefore, the antiproliferative pathway converging onto pRb...
and the cell division cycle is, in one way or another, disrupted in a majority of human cancer cells. Besides cancer, dysregulation of TGF-β signaling also has been associated with other human diseases such as Marfan’s syndrome and thoracic aortic aneurysm.

**Gene Therapy and Molecular Drugs in Cancer**

Modern advances in the use of molecular biology to manipulate genomes have greatly contributed to the understanding of the molecular basis for how cells live, die, or differentiate. Given the fact that human diseases arise from improper changes in the genome, the continuous understanding of how the genome functions will make it possible to tailor medicine on an individual basis. Although significant hurdles remain, the course toward therapeutic application of molecular biology already has been mapped out by many proof-of-principle studies in the literature. In this section, cancer is used as an example to elaborate some therapeutic applications of molecular biology. Modern molecular medicine includes gene therapy and molecular drugs that target genes or gene products within human cells.

Cancer is a complex disease, involving uncontrolled growth and spread of tumor cells (Fig. 15-12). Cancer development depends on the acquisition and selection of specific characteristics that set the tumor cell apart from normal somatic cells. Cancer cells have defects in regulatory circuits that govern normal cell proliferation and homeostasis. Many lines of evidence indicate that tumorigenesis in humans is a multistep process and that these steps reflect genetic alterations that drive the progressive transformation of normal human cells into highly malignant derivatives. The genomes of tumor cells are invariably altered at multiple sites, having suffered disruption through lesions as subtle as point mutations and as obvious as changes in chromosome complement. A succession of genetic changes, each conferring one or another type of growth advantage, leads to the progressive conversion of normal human cells into cancer cells.

Cancer research in the past 20 years has generated a rich and complex body of knowledge, revealing cancer to be a disease involving dynamic changes in the genome. The causes of cancer include genetic predisposition, environmental influences, infectious agents, and aging. These transform normal cells into cancerous ones by derailing a wide spectrum of regulatory pathways including signal transduction pathways, cell cycle machinery, or apoptotic pathways.15,16 The early notion that cancer was caused by mutations in genes critical for the control of cell growth implied that genome stability is important for preventing oncogenesis. There are two classes of cancer genes in which alteration has been identified in human and animal cancer cells: oncogenes, with dominant gain-of-function mutations, and tumor suppressor genes, with recessive loss-of-function mutations. In normal cells, oncogenes promote cell growth by activating cell cycle progression, whereas tumor suppressors counteract oncogenes’ functions. Therefore, the balance between oncogenes and tumor suppressors maintains a well-controlled state of cell growth.

During the development of most types of human cancer, cancer cells can break away from primary tumor masses, invade adjacent tissues, and hence travel to distant sites where they form new colonies. This spreading process of tumor cells, called metastasis, is the cause of 90% of human cancer deaths. Metastatic cancer cells that enter the bloodstream can reach virtually all tissues of the body. Bones are one of the most common places for these cells to settle and start growing again. Bone metastasis is one of the most frequent causes of pain in people with cancer. It also can cause bones to break and create other symptoms and problems for patients.

The progression in the knowledge of cancer biology has been accelerating in recent years. All of the scientific knowledge...
acquired through hard work and discovery has made it possible for cancer treatment and prevention. As a result of explosive new discoveries, some modern treatments were developed. The success of these therapies, together with traditional treatments such as surgical procedures, is further underscored by the fact that in 2002 the cancer rate was reduced in the United States. Current approaches to the treatment of cancer involve killing cancer cells with toxic chemicals, radiation, or surgery. Alternatively, several new biologic- and gene-based therapies are aimed at enhancing the body’s natural defenses against invading cancers. Understanding the biology of cancer cells has led to the development of designer therapies for cancer prevention and treatment. Gene therapy, immune system modulation, genetically engineered antibodies, and molecularly designed chemical drugs are all promising fronts in the war against cancer.

Immunotherapy. The growth of the body is controlled by many natural signals through complex signaling pathways. Some of these natural agents have been used in cancer treatment and have been proven effective for fighting several cancers through the clinical trial process. These naturally occurring biologic agents, such as interferons, interleukins, and other cytokines, can now be produced in the laboratory. These agents, as well as the synthetic agents that mimic the natural signals, are given to patients to influence the natural immune response agents either by directly altering the cancer cell growth or by acting indirectly to help healthy cells control the cancer. One of the most exciting applications of immunotherapy has come from the identification of certain tumor targets called antigens and the aiming of an antibody at these targets. This was first used as a means of localizing tumors in the body for diagnosis and was more recently used to attack cancer cells. Trastuzumab (Herceptin) is an example of such a drug. Trastuzumab is a monoclonal antibody that neutralizes the mitogenic activity of cell-surface growth factor receptor HER-2, which is overexpressed in approximately 25% of breast cancers. HER-2—overexpressing tumors tend to grow faster and generally are more likely to recur than tumors that do not overproduce HER-2. Trastuzumab is designed to attack cancer cells that overexpress HER-2 by slowing or preventing the growth of these cells, resulting in increased survival of HER-2—positive breast cancer patients. Another significant example is the administration of interleukin-2 (IL-2) to patients with metastatic melanoma or kidney cancer, which has been shown to mediate the durable regression of metastatic cancer. IL-2, a cytokine produced by human helper T lymphocytes, has a wide range of immune regulatory effects, including the expansion of lymphocytes following activation by a specific antigen. Although IL-2 has no direct impact on cancer cells, the impact of IL-2 on cancers in vivo derives from its ability to expand lymphocytes with antitumor activity. The expanded lymphocyte pool enables recognition of the antigen on cancer cells. Thus, the molecular identification of cancer antigens has opened new possibilities for the development of effective immunotherapies for patients with cancer. Clinical studies using immunization with peptides derived from cancer antigens have shown that high levels of lymphocytes with antitumor activity can be produced in cancer-bearing patients. Highly avid antitumor lymphocytes can be isolated from immunized patients and grown in vitro for use in cell-transfer therapies.

Chemotherapy. The primary function of anticancer chemicals is to block different steps involved in cell growth and replication. These chemicals often block a critical chemical reaction in a signal transduction pathway or during DNA replication or gene expression. For example, STI571, also known as Gleevec, is one of the first molecularly targeted drugs based on the changes that cancer causes in cells. STI571 offers promise for the treatment of chronic myeloid leukemia (CML) and may soon surpass interferon-γ as the standard treatment for the disease. In CML, STI571 is targeted at the Bcr-Abl kinase, an activated oncogene product in CML (Fig. 15-13). Bcr-Abl is an over activated protein kinase resulting from a specific genetic abnormality generated by chromosomal translocation that is found in the cells of patients with CML. STI571-mediated inhibition of Bcr-Abl kinase activity not only prevents cell growth of Bcr-Abl–transformed leukemic cells, but also induces apoptosis. Clinically, the drug quickly corrects the blood cell abnormalities caused by the leukemia in a majority of patients, achieving a complete disappearance of the leukemic blood cells and the return of normal blood cells. Additionally, the drug appears to have some effect on other cancers including certain breast tumors and gastrointestinal (GI) stromal tumors, a very rare type of stomach cancer.

Gene Therapy. Gene therapy is an experimental treatment that involves genetically altering a patient’s own tumor cells or lymphocytes (cells of the immune system, some of which can attack cancer cells). For years, the concept of gene therapy has held promise as a new, potentially potent weapon to attack cancer. Although a rapid progression in the understanding of the molecular and clinical aspects of gene therapy has been witnessed in the past decade, gene therapy treatment has not yet been shown to be superior to standard treatments in humans.

Figure 15-13. Mechanism of STI571 as a molecular drug. Bcr-Abl is an over activated oncogene product resulting from a specific genetic abnormality generated by chromosomal translocation that is found in cells of patients with chronic myeloid leukemia. Bcr-Abl is an activated protein kinase and thus requires adenosine triphosphate (ATP) to phosphorylate substrates, which in turn promote cell proliferation. STI571 is a small molecule that competes with the ATP-binding site and thus blocks the transfer of phosphoryl group to substrate. PO₄ = phosphate; Tyr = tyrosine.
Several problems must be resolved to transform it into a clinically relevant form of therapy. The major issues that limit its translation to the clinic are improving the selectivity of tumor targeting, improving the delivery to the tumor, and the enhancement of the transduction rate of the cells of interest. In most gene therapy trials for malignant diseases, tumors can be accessed and directly injected (in situ gene therapy). In situ gene therapy also offers a better distribution of the vector virus throughout the tumor. Finally, a combination of gene therapy strategies will be more effective than the use of a single gene therapy system. An important aspect of effective gene therapy involves the choice of appropriate genes for manipulation. Genes that promote the production of messenger chemicals or other immune-active substances can be transferred into the patient’s cells. These include genes that inhibit cell cycle progression, induce apoptosis, enhance host immunity against cancer cells, block the ability of cancer cells to metastasize, and cause tumor cells to undergo suicide. Recent development of RNAi technology, which uses a loss-of-function approach to block gene functions, ensures a new wave of hopes for gene therapy. Nonetheless, gene therapy is still experimental and is being studied in clinical trials for many different types of cancer. The mapping of genes responsible for human cancer is likely to provide new targets for gene therapy in the future. The preliminary results of gene therapy for cancer are encouraging, and as advancements are made in the understanding of the molecular biology of human cancer, the future of this rapidly developing field holds great potential for treating cancer.

It is noteworthy that the use of multiple therapeutic methods has proven more powerful than a single method. The use of chemotherapy after surgery to destroy the few remaining cancerous cells in the body is called adjuvant therapy. Adjuvant therapy was first tested and found to be effective in breast cancer. It was later adopted for use in other cancers. A major discovery in chemotherapy is the advantage of multiple chemotherapeutic agents (known as combination or cocktail chemotherapy) over single agents. Some types of fast-growing leukemias and lymphomas (tumors involving the cells of the bone marrow and lymph nodes) responded extremely well to combination chemotherapy, and clinical trials led to gradual improvement of the drug combinations used. Many of these tumors can be cured today by combination chemotherapy. As cancer cells carry multiple genetic defects, the use of combination chemotherapy, immunotherapy, and gene therapies may be more effective in treating cancers.

**Stem Cell Research**

Stem cell biology represents a cutting-edge scientific research field with potential clinical applications. It may have an enormous impact on human health by offering hope for curing human diseases such as diabetes mellitus, Parkinson’s disease, neurologic degeneration, and congenital heart disease. Stem cells are endowed with two remarkable properties (Fig. 15-14). First, stem cells can proliferate in an undifferentiated but pluripotent state and, as a result, can self-renew. Second, they have the ability to differentiate into many specialized cell types. There are two groups of stem cells: embryonic stem (ES) cells and adult stem cells.

Human ES cells (hESCs) are derived from early preimplantation embryos called blastocysts (5 days postfertilization) and are capable of generating all differentiated germ layers in the body by chimera assays or 2-D/3-D differentiation in a dish—ectoderm, mesoderm, and endoderm—and therefore are considered pluripotent. There are two pluripotent states associated with hESCs, one of which is the classic culture with bFGF (basic fibroblast growth factor) and knock out replacer (KSR), termed as “primed” pluripotent state. More recently, “naïve” hESC culture methods have been introduced based on mouse studies, by supplementing 2i inhibitors (MEK1 and GSK3β inhibitors) into the medium in addition to bFGF.

Adult stem cells are present in and can be isolated from adult tissues. They often are tissue specific and only can generate the cell types comprising a particular tissue in the body; therefore, they are considered multipotent. However, in some cases, they can transdifferentiate into cell types found in other tissues, called transdifferentiation. For example, hematopoietic stem cells are adult stem cells. They reside in bone marrow and are capable of generating all cell types of the blood and immune system. Another example is mesenchymal stem cells (MSCs), which is initially identified in the bone marrow (BM) to support hematopoietic stem cell homeostasis. In addition to BM, MSCs are also present in adipose tissue, umbilical cord, placenta, amniotic fluid, dental pulp, skeletal muscle, tendons, and synovial, etc., and are reported to obtain the ability to differentiate into osteogenic, chondrogenic, and adipogenic lineages in vitro. Due to their characteristics of easy acquisition (from adipose tissue, for example), strong ex vivo proliferation, immune-modulatory function, and ability to migrate to damaged tissue, MSCs have been utilized in regenerative medicine.

Stem cells can be grown in culture and be induced to differentiate into a particular cell type, either in vitro or in vivo. With the recent and continually increasing improvement in culturing stem cells, scientists are beginning to understand the molecular mechanisms of stem cell self-renewal and differentiation in response to environmental cues. It is believed that discovery of the signals that control self-renewal vs differentiation will be extremely important for the therapeutic use of stem cells in treating disease. It is possible that success in the study of the changes in signal transduction pathways in stem cells will lead to the development of therapies to replace diseased or damaged cells in the body using stem cell derivatives. Recently, stem cell research has been transformed by the discovery from the Shinya Yamanaka group and the James Thomsen group, who have found that a simple genetic manipulation can reprogram adult differentiated cells back into pluripotent stem cells. This exciting discovery not only bypasses the ethical issues of using early embryos to generate ES cells, but also ensures a...
potentially limitless source of patient-specific stem cells for tissue engineering and regenerative medicine.

**The Atomic Theory of Disease**

As early as the 5th century B.C., the ancient Greek Democritus first proposed that matter is composed of indivisible particles called “atoms.” In the 17th and 18th centuries, Isaac Newton described the expansion of gases as rushed atoms into empty space. The existence of atoms was doubted until the discovery of subatomic particles in the 20th century, which demonstrated that the atom was actually divisible into protons, neutrons, and electrons.

Over 100 years after this discovery, direct impacts from utilizing subatomic particles were revealed. This began with the discovery of the X-ray. The most advanced and well-applied atomic technologies include accurate imaging such as X-ray computed tomography (CT) scan, magnetic resonance imaging (MRI), single photon emission computed tomography (SPECT), and positron emission tomography (PET). Additional applications include radiation oncology, which utilizes ionizing particles to treat malignant diseases by inducing double-strand DNA breaks resulting in programmed cell death.

The staggering advances in anatomy, physiology, and molecular biology over the past centuries have led us to our current state in which the atom is now the anatomy of the 21st century. As 99% of the body is composed of six elements (oxygen, carbon, hydrogen, nitrogen, calcium, and phosphorus), the next great advance in medicine will be bridging the subatomic, molecular, and genomic levels by forming an atomic theory of disease, which states that alterations in the composition of subatomic particles are the root cause of disease. The atomic theory of disease would include genetic alterations at the atomic/subatomic level that are akin to single nucleotide polymorphisms (SNPs), in which alleles for a gene differ on the exact nucleotide in a single location, which can change the ultimate protein structure. This can lead to subtle changes in function or dramatic results that cause pathology. We hypothesize that on a subatomic level, there could potentially be polymorphisms as well, in which there are subtle changes in the sea of subatomic particles. Isotopes, discovered 100 years ago, would fall into this category of subatomic polymorphism, as they differ in the number of neutrons present in the atom. Differences in other particles may not change the mass of the atom, but may alter some of the characteristics of the atom. This is where the basic human variations originate because an atomic polymerism would result in particular genetic change. Remarkably, somatic point mutations of *KRAS* and *P53* are caused by a single proton shift in cytosine, which is known as tautomerization of cytosine. In turn, the tautomerized cytosine binds with adenine, and not guanine, thus leading to point mutations in *KRAS* and *P53*, which are well known driver mutations for many cancers. This is another example of alterations in subatomic particles that directly cause disease.

A known example of a change in the subatomic milieu of an element leading to a disease process is that of methemoglobinemia, a disorder characterized by an overabundance of methemoglobin. Methemoglobin contains an oxidized form of iron (carrying an extra electron), as opposed to the reduced form in normal hemoglobin. This results in a shift in the oxygen-hemoglobin dissociation curve to the left, causing hypoxia. Methemoglobinemia can be congenital, due to a defect in an enzyme that normally reduces methemoglobin back to hemoglobin, or acquired, caused by breakdown products of drugs that can oxidize hemoglobin. Although there is less than 1% of methemoglobin normally present in human tissues, affecting local blood flow and inflammation through its effects on nitric oxide and heme, large quantities can lead to respiratory failure and death.

Another example would be exposure to external energy such as radiation that leads to instability of nuclear genome. In Chernobyl and Fukushima, radiolabeled food was metabolized and incorporated into body cells and decay to emit gamma radiation, causing DNA damage. This radiation damage occurs primarily at a subatomic level from a radiobiologic point of view with a direct or indirect ionization of atoms. The clinical results depend on the tissue characteristics and the equilibrium between the damage applied to normal and diseased tissues.

**TECHNOLOGIES OF MOLECULAR AND CELL BIOLOGY**

**DNA Cloning**

Since the advent of recombinant DNA technology three decades ago, hundreds of thousands of genes have been identified. Recombinant DNA technology is the technology that uses advanced enzymatic and microbiologic techniques to manipulate DNA. Pure pieces of any DNA can be inserted into bacterial DNA or other carrier DNA such as plasmids to produce recombinant DNA in bacteria. In this way, DNA can be reconstructed, amplified, and used to manipulate the functions of individual cells or even organisms. This technology, often referred to as DNA cloning, is the basis of all other DNA analysis methods. It is only with the awesome power of recombinant DNA technology that the completion of the Human Genome Project was possible. It also has led to the identification of the entire gene complements of organisms such as viruses, bacteria, worms, flies, and plants.

*Molecular cloning* refers to the process of cloning a DNA fragment of interest into a DNA vector that ultimately is delivered into bacterial or mammalian cells or tissues (Fig. 15-15). This represents a very basic technique that is widely used in almost all areas of biomedical research. DNA vectors often are called *plasmids*, which are extrachromosomal molecules of DNA that vary in size and can replicate and be transmitted from bacterial cell to cell. Plasmids can be propagated either in the cytoplasm or after insertion, as part of the bacterial chromosome in *Escherichia coli*. The process of molecular cloning involves several steps of manipulation of DNA. First, the vector plasmid DNA is cleaved with a restriction enzyme to create compatible ends with the foreign DNA fragment to be cloned. The vector and the DNA fragment are then joined in vitro by a DNA ligase. Alternatively, DNA cloning can be simply done through the so-called *Gateway Technology* that allows for the rapid and efficient transfer of DNA fragments between different cloning vectors while maintaining reading frame and orientation, without the use of restriction endonucleases and DNA ligase. The technology, which is based on the site-specific recombination system of bacteriophage λ, is simple, fast, robust, and automatic and thus compatible for high-throughput DNA cloning.

Finally, the ligation product or the Gateway reaction product is introduced into competent host bacteria; this procedure is called *transformation*, which can be done by either calcium/heat shock or electroporation. Precautions must be taken in every step of cloning to generate the desired DNA construct.
Figure 15-15. Generation of recombinant DNA. The vector is a circular DNA molecule that is capable of replicating in *Escherichia coli* cells. Insert DNA (often your favorite gene) is ligated to the vector after ends of both DNA are properly treated with restriction enzymes. Ligated DNA (i.e., the recombinant plasmid DNA) is then transformed into *E. coli* cells, where it replicates to produce recombinant progenies. *E. coli* cells carrying the recombinant plasmid can be propagated to yield large quantities of plasmid DNA.

The vector must be correctly prepared to maximize the creation of recombinants; for example, it must be enzymatically treated to prevent self-ligation. Host bacteria must be made sufficiently competent to permit the entry of recombinant plasmids into cells. The selection of desired recombinant plasmid-bearing *E. coli* normally is achieved by the property of drug resistance conferred by the plasmid vectors. The plasmids encoding markers provide specific resistance to (i.e., the ability to grow in the presence of) antibiotics such as ampicillin, kanamycin, and tetracycline. The foreign component in the plasmid vector can be a mammalian expression cassette, which can direct expression of foreign genes in mammalian cells. The resulting plasmid vector can be amplified in *E. coli* to prepare large quantities of DNA for its subsequent applications such as transfection, gene therapy, transgenics, and knockout mice.

**Detection of Nucleic Acids and Proteins**

**Southern Blot Hybridization.** Southern blotting refers to the technique of transferring DNA fragments from an electrophoresis gel to a membrane support and the subsequent analysis of the fragments by hybridization with a radioactively or chemiluminescently labeled probe (Fig. 15-16). Southern blotting is named after E. M. Southern, who in 1975 first described the technique of DNA analysis. It enables reliable and efficient analysis of size-fractionated DNA fragments in an immobilized membrane support. Southern blotting is composed of several steps. It normally begins with the digestion of the DNA samples with appropriate restriction enzymes, which will discriminate wild-type and mutant DNA by size and the separation of DNA samples in an agarose gel by electrophoresis with appropriate DNA size markers, called the DNA ladder. The DNA gel is stained with a dye, usually ethidium bromide, and photographed with a ruler laid alongside the gel so that band positions can later be identified on the membrane. The DNA gel then is treated so the DNA fragments are denatured (i.e., strand separation). The DNA then is transferred onto a nitrocellulose membrane by capillary diffusion or under electricity. After immobilization, the DNA can be subjected to hybridization analysis, enabling bands with sequence similarity to a radioactively or chemiluminescently labeled probe to be identified.

The development of Southern transfer and the associated hybridization techniques made it possible for the first time to obtain information about the physical organization of single and multicopy sequences in complex genomes. The later application of Southern blotting hybridization to the study of restriction fragment length polymorphisms opened up new possibilities such as genetic fingerprinting and prenatal diagnosis of genetic diseases.

**Northern Blot Hybridization.** Northern blotting refers to the technique of size fractionation of RNA in a gel and the transferring of an RNA sample to a solid support (membrane) in such a manner that the relative positions of the RNA molecules are maintained. The resulting membrane then is hybridized with a labeled probe complementary to the mRNA of interest. Signals generated from detection of the membrane can be used to determine the size and abundance of the target RNA. In principle,
Northern blot hybridization is similar to Southern blot hybridization (and hence its name), with the exception that RNA, not DNA, is on the membrane. Although reverse-transcriptase PCR has been used in many applications (described in the next section, “Polymerase Chain Reaction”), Northern analysis is the only method that provides information regarding mRNA size and has remained a standard method for detection and quantitation of mRNA. The process of Northern hybridization involves several steps, as does Southern hybridization, including electrophoresis of RNA samples in an agarose-formaldehyde gel, transfer to a membrane support, and hybridization to a radioactively labeled DNA probe. Data from hybridization allow quantification of steady-state mRNA levels and, at the same time, provide information related to the presence, size, and integrity of discrete mRNA species. Thus, Northern blot analysis, also termed RNA gel blot analysis, commonly is used in molecular biology studies relating to gene expression.

**Polymerase Chain Reaction.** PCR is an in vitro method for the polymerase-directed amplification of specific DNA sequences using two oligonucleotide primers that hybridize to opposite strands and flank the region of interest in the target DNA (Fig. 15-17).27 One cycle of PCR reaction involves template denaturation, primer annealing, and the extension of

---

**Figure 15-17.** Amplification of DNA using the polymerase chain reaction (PCR) technique. Knowledge of the DNA sequence to be amplified is used to design two synthetic DNA oligonucleotides, each complementary to the sequence on one strand of the DNA double helix at opposite ends of the region to be amplified. These oligonucleotides serve as primers for in vitro DNA synthesis, which is performed by a DNA polymerase, and they determine the segment of the DNA that is amplified. A. PCR starts with a double-stranded DNA, and each cycle of the reaction begins with a brief heat treatment to separate the two strands (Step 1). After strand separation, cooling of the DNA in the presence of a large excess of the two primer DNA oligonucleotides allows these primers to hybridize to complementary sequences in the two DNA strands (Step 2). This mixture is then incubated with DNA polymerase and the four deoxyribonucleoside triphosphates so that DNA is synthesized, starting from the two primers (Step 3). The entire cycle is then begun again by a heat treatment to separate the newly synthesized DNA strands. B. As the procedure is performed over and over again, the newly synthesized fragments serve as templates in their turn, and, within a few cycles, the predominant DNA is identical to the sequence bracketed by and including the two primers in the original template. Of the DNA put into the original reaction, only the sequence bracketed by the two primers is amplified because there are no primers attached anywhere else. In the example illustrated in B, three cycles of reaction produce 16 DNA chains, eight of which (boxed in brown) are the same length as and correspond exactly to one or the other strand of the original bracketed sequence shown at the far left; the other strands contain extra DNA downstream of the original sequence, which is replicated in the first few cycles. After three more cycles, 240 of the 256 DNA chains correspond exactly to the original bracketed sequence, and after several more cycles, essentially all of the DNA strands have this unique length.
the annealed primers by DNA polymerase. Because the primer extension products synthesized in one cycle can serve as a template in the next, the number of target DNA copies nearly doubles at each cycle. Thus, a repeated series of cycles results in the exponential accumulation of a specific fragment in which the termini are sharply defined by the 5' ends of the primers. The introduction of the thermostable DNA polymerase (e.g., Taq polymerase) transforms the PCR into a simple and robust reaction. The reaction components (e.g., template, primers, Taq polymerase, 2'-deoxynucleoside 5'-triphosphates, and buffer) could all be assembled and the amplification reaction carried out by simply cycling the temperatures within the reaction tube. The specificity and yield in amplifying a particular DNA fragment by PCR reaction are affected by the proper setting of the reaction parameters (e.g., enzyme, primer, and Mg²⁺ concentration, as well as the temperature cycling profile). Modifying various PCR parameters to optimize the specificity of amplification yields more homogenous products, even in rare template reactions.

The emergence of the PCR technique has dramatically altered the approach to both fundamental and applied biological problems. The capability of amplifying a specific DNA fragment from a gene or the whole genome greatly advances the study of the gene and its function. It is simple, yet robust, speedy, and most of all, flexible. As a recombinant DNA tool, it underlies almost all of molecular biology. This revolutionary technique enabled the modern methods for the isolation of genes, construction of a DNA vector, introduction of alterations into DNA, and quantitation of gene expression, making it a fundamental cornerstone of genetic and molecular analysis.

**Immunoblotting and Immunoprecipitation.** Analyses of proteins are primarily carried out by antibody-directed immunologic techniques. For example, Western blotting, also called immunoblotting, is performed to detect protein levels in a population of cells or tissues, whereas immunoprecipitation is used to concentrate proteins from a larger pool. Using specific antibodies, microscopic analysis called immunofluorescence and immunohistochemistry is possible for the subcellular localization and expression of proteins in cells or tissues, respectively.

Immunoblotting refers to the process of identifying a protein from a mixture of proteins (Fig. 15-18). It consists of five steps: (a) sample preparation; (b) electrophoresis (separation of a protein mixture by sodium dodecyl sulfate-polyacrylamide gel electrophoresis); (c) transfer (the electrophoretic transfer of proteins from gel onto membrane support [e.g., nitrocellulose, nylon, or polyvinylidene difluoride]); (d) staining (the subsequent immunodetection of target proteins with specific antibody); and (e) development (colorimetric, chemiluminescent, and recently fluorescent visualization of the antibody-recognized protein). Thus, immunoblotting combines the resolution of gel electrophoresis with the specificity of immunochromic detection. Immunoblotting is a powerful tool used to determine a number of important characteristics of proteins. For example, immunoblotting analysis will determine the presence and the quantity of a protein in a given cellular condition and its relative molecular weight. Immunoblotting also can be used to determine whether posttranslational modification such as phosphorylation has occurred on a protein. Importantly, through immunoblotting analysis, a comparison of the protein levels and modification states in normal vs diseased tissues is possible.

Immunoprecipitation, another widely used immunochromic technique, is a method that uses antibody to enrich a protein of interest and any other proteins that are associated with it (Fig. 15-19). The principle of the technique lies in the property of a strong and specific affinity between antibodies and their antigens to locate and pull down target proteins in solution. Once the antibody-antigen (target protein) complexes are formed in the solution, they are collected and purified using small agarose beads with covalently attached protein A or protein G. Both protein A and protein G specifically interact with the antibodies,
thus forming a large immobilized complex of antibody-antigen bound to beads. The purified protein can then be analyzed by a number of biochemical methods. When immunoprecipitation is combined with immunoblotting, it can be used for the sensitive detection of proteins in low concentrations, which would otherwise be difficult to detect. Moreover, combined immunoprecipitation and immunoblotting analysis is very efficient in analyzing the protein-protein interactions or determining the posttranslational modifications of proteins. In addition, immunoprecipitated proteins can be used as preparative steps for assays such as intrinsic or associated enzymatic activities. The success of immunoprecipitation is influenced by two major factors: (a) the abundance of the protein in the original preparation and (b) the specificity and affinity of the antibody for this protein.

Recently, immunoprecipitation is even used to enrich modified DNA (for example, 5-methylcytosine) for bisulfite sequencing. Besides proteins of interest, specific antibodies can also be raised against specially modified DNA. Like the protein immunoprecipitation, modified DNA can be pulled down, taking advantage of the specificity and affinity of antibody to antigen.

**DNA Microarray.** Now that the human genome sequence is completed, the primary focus of biologists is rapidly shifting toward gaining an understanding of how genes function. One of the interesting findings about the human genome is that there are only approximately 25,000 to 30,000 protein-encoding genes. However, it is known that genes and their products function in a complicated and yet orchestrated fashion and that the surprisingly small number of genes from the genome sequence is sufficient to make a human being. Nonetheless, with the tens of thousands of genes present in the genome, traditional methods in molecular biology, which generally work on a one-gene-in-one-experiment basis, cannot generate the whole picture of genome function. In the past several years, a new technology called DNA microarray has attracted tremendous interest among biologists as well as clinicians. This technology promises to monitor the whole genome on a single chip so researchers can have a better picture of the interactions among thousands of genes simultaneously.

DNA microarray, also called gene chip, DNA chip, and gene array, refers to large sets of probes of known sequences orderly arranged on a small chip, enabling many hybridization reactions to be carried out in parallel in a small device (Fig. 15-20).28 Like Southern and Northern hybridization, the underlying principle of this technology is the remarkable ability of nucleic acids to form a duplex between two strands with complementary base sequences. DNA microarray provides a medium for matching known and unknown DNA samples based on base-pairing rules and automating the process of identifying the unknowns. Microarrays require specialized robotics and imaging equipment that spot the samples on a glass or nylon substrate, carry out the hybridization, and analyze the data generated. DNA microarrays containing different sets of genes from a variety of organisms are now commercially available, allowing biologists to simply purchase the chips and perform hybridization and data collection. The massive scale of microarray experiments requires the aid of computers. They are used during the capturing of the image of the hybridized target, the conversion of the image into usable measures of the extent of hybridization, and the interpretation of the extent of hybridization into a meaningful measure of the amount of the complementary sequence in the target. Some data-analysis packages are available commercially or can be found in the core facility of certain institutions.

DNA microarray technology has produced many significant results in quite different areas of application. There are two major application forms for the technology: identification of sequence (gene/gene mutation) in multiple regions of a genome and determination of expression level (abundance) of large numbers of genes simultaneously. For example, analysis of genomic DNA detects amplifications and deletions found in human tumors. Differential gene expression analysis also has uncovered networks of genes differentially present in cancers that cannot be distinguished by conventional means. Significantly, recent advancements in next-generation sequencing (e.g., Solexa and 454 technology) have demonstrated the precision and speed to analyze gene expression in any genome.

**Next-Generation Sequencing.**29,30 First-generation sequencing, also termed Sanger's sequencing, requires a single-stranded DNA template, a specific DNA primer, a DNA polymerase, normal deoxynucleoside triphosphates (dTTPs), and modified di-deoxynucleotidetriphosphates (ddNTPs). In the process of DNA sequencing, DNA polymerase adds random dNTP or ddNTP
Figure 15-20. DNA microarrays. DNA microarrays, also referred to as gene chips, have arrayed oligonucleotides or complementary DNAs (cDNAs) corresponding to tens or hundreds of distinct genes. DNA microarray is used to comparatively analyze gene expression in different cells or tissues. Messenger RNAs (mRNAs) extracted from different sources are converted into cDNAs, which are then labeled with different fluorescent dyes. The two fluorescent cDNA probes are mixed and hybridized to the same DNA microarrays. The ratio of red to green fluorescence at each spot on the chip represents the relative expression of levels of that gene between two different cells. In the example shown in the figure, cDNA from cell #1 is labeled with red fluorescence and that from cell #2 is labeled with green fluorescence. On the microarray, red spots demonstrate that the gene in the cell sample #1 is expressed at a higher level than the corresponding gene in cell sample #2. The green spots indicate that the gene in the cell sample #2 is expressed at a higher level than the corresponding gene in the cell sample #1. Yellow spots represent equal expression of the gene in both cell samples.

after the primer. If ddNTP is incorporated at the end of the chain, it terminates the reaction and results in DNA fragments of different sizes. The ddNTPs could be radioactively or fluorescently labeled for auto-machine detection. Usually, Sanger’s sequencing is able to read sequence below 1 kb with the quality deteriorating after 700 bp. The accuracy and success rate largely depends on the DNA polymerase used.

Recombinant DNA technology greatly impacts the completion of the Human Genome Project due to the invention of shotgun sequencing, which includes breaking the genome DNA into small pieces and randomly cloning those pieces into DNA vectors that are easily sequenced. Based on the overlapping sequence of each clone, computer analysis can be programmed to map and align the DNA sequence that will ultimately cover the whole human genome.

Based on shotgun sequencing, as the sequencing technologies advance, next-generation sequencing (NGS), also called second-generation sequencing, has become one of the most powerful tools to analyze DNA mutation, identify epigenetic modification, and profile gene expression or ncRNA expression. The next-generation sequencing process usually includes library construction, sequencing, and data analysis. There are three major NGS platforms: Roche 454, Life Technologies Ion Torrent, and Illumina Solexa. Take the Illumina next-generation sequencing as an example: DNA are shared or digested into small pieces and then used to generate a DNA library with adapters on both ends of each DNA piece. Then, the DNA library is diluted and loaded on a chamber of a slide, called a lane, for cluster amplification. Cycled fluorescent deoxyribonucleotide triphosphates (dNTPs) are then added to the chamber to enable DNA polymerization, resulting in different fluorescent emission representing different dNTP reading on different clusters, into a microscope. The fluorescent signal is transformed into sequencing data that will be aligned and mapped to a standard genome database. The advantages of next-generation sequencing include the following: no necessity of DNA cloning; fast and cost-effective; and a huge amount of data to give good depth and accuracy of the sequence.

Based on the applications, the most common next-generation sequencing technologies for whole-genome sequencing are whole-genome DNA sequencing, whole-genome bisulfite sequencing (BS-seq), RNA sequencing (RNA-seq), and chromatin immunoprecipitation (ChIP) sequencing (ChIP-seq). Whole-genome DNA sequencing is purely to sequence the DNA sequence of a genome without any preprocessing of the DNA, reflecting any deletions, replications, and mutations within the genomic DNA. Given that the genomic information for human is enormous, to achieve great depth and accuracy the genome needs to be sequenced multiple times to reach statistic significance and pass quality control. Therefore, whole-genome DNA sequencing is still considered to be costly for diagnosis as well as for research purposes. Under clinical settings, one of the most applied DNA sequencing technologies is whole-exome sequencing, i.e., using techniques to capture and analyze exons in all coding genes, given that most known diseases are due to mutations in spliced-in regions of coding genes. Whole-exome sequencing is mostly used for detecting single nucleotide variants and is less reliable in detecting insertion-deletion variants or small copy number variants.

BS-seq is commonly used to identify DNA methylation on the genome (5-methylcytosine [5mC]). The process always involves a bisulfite treatment of DNA before library
construction, during which the unmethylated cytosine will be transformed to a uracil, resulting in reading as a thymine in data output, whereas 5mC is protected and remains as cytosine in data output. Thus, 5mC and cytosine are distinguished this way. To perform whole-genome BS-seq, the library construction starts with a decent amount of DNA followed by a complete conversion of unmethylated cytosine. Because it is also counted as whole-genome sequencing, a certain depth of data needs to be achieved for accurate and convincing results. To develop an affordable genome-wide DNA methylation analysis, the reduced representation bisulfite sequencing (RRBS) approach has been applied to enrich CpG-dense regions of the genome by digesting genomic DNA using a methylation-insensitive restriction enzyme, usually MspI. This method covers a majority of promoters as well as some repetitive regions in the genome. RNA-seq is usually performed to analyze transcription for the same purpose as performing a microarray. However, RNA-seq is more accurate and provides more information such as splicing variants than traditional microarray. Usually, cDNA that is reversely transcribed from extracted RNA is used to generate libraries. Depending on the needs, mRNA and ncRNA can be enriched in different protocols for RNA extraction. Currently, techniques have been improved to perform RNA profiling on a single cell level. Single-cell RNA sequencing (scRNA-seq) allows expression analysis of individual cells in a population. One use is to examine heterogeneity in tumor cells. Although low copy number genes are usually poorly detected, scRNA-seq on a large number of cells can reveal rare cell subpopulations and uncommon RNA expression patterns among different subpopulations.

ChIP-seq is always used to map the location of a DNA-binding protein in the genome. Prior to library construction, ChIP is performed to enrich DNA bound by the protein of interest (POI). First, POI and DNA are cross-linked before sonication. Then, a specific antibody is used to pull down POI and attached DNA fragments. After the protein and DNA are reverse cross-linked, DNA is purified to make the ChIP-seq library. If using an antibody against particular transcription factor (TF), DNA sequences bound by this TF are pulled down and examined. The TF binding consensus sequences is then predicted, and if the TF binds to a promoter region, it is likely that the gene using this promoter is regulated by this TF. If using an antibody against an epigenetic modification, the modified regions are marked up into the genome to facilitate identification of potential epigenetic regulating mechanisms.

By using next-generation sequencing technology, any potential mutations in a patient can be scrutinized as well as any defects in epigenetic modification. By combining data from different kinds of sequencing (DNA-seq, RNA-seq, ChIP-seq), better understanding of mutation or transcription-caused diseases aligning with epigenetic regulation can be achieved, which will greatly facilitate the diagnosis of patients and personalization of medicine in a fast and economic way by preventing unnecessary medical costs and procedures.

Third-generation sequencing has emerged rapidly at the research level to involve single molecule real-time sequencing (SMRT). Although first developed and marketed by Pacific Biosciences (Paci Bio), Roche is now leading this technology. Third-generation sequencing allows amplification-free single-molecule sequencing with read length extension up to megabases and reduced sequencing coverage bias. It can be used to build the gap in the human genome (for example, low complexity regions), provide access to structural genomic variants, and simultaneously analyze genome-wide single-nucleotide methylation. So far, clinical application and research is heavily dependent on NGS, especially standardizing and reducing the cost of post-NGS analysis. Third-generation sequencing is under fast development and has been adapted to aid and append NGS.

**Cell Manipulations**

**Cell Culture.** Cell culture has become one of the most powerful tools in biomedical laboratories, as cultured cells are being used in a diversity of biologic fields ranging from biochemistry to molecular and cellular biology. Through their ability to be maintained in vitro, cells can be manipulated by the introduction of genes of interest (cell transfection) and be transferred into in vivo biologic receivers (cell transplantation) to study the biologic effect of the interested genes (Fig. 15-21). In common

---

**Figure 15-21.** Cell culture and transfection. A. Primary cells can be isolated from tissues and cultured in medium for a limited period of time. After genetic manipulations to overcome the cell aging process, primary cells can be immortalized into cell lines for long-term culture. B. DNA can be introduced into cells to produce recombinant gene products or to analyze the biologic functions of the gene.
laboratory settings, cells are cultured either as a monolayer (in which cells grow as one layer on culture dishes), considered 2-D, or in suspension or biomedical material skeleton such as hydrogel, considered 3-D.

It is important to know the wealth of information concerning cell culturing before attempting the procedure. For example, conditions of culture will depend on the cell types to be cultured (e.g., origins of the cells such as epithelial or fibroblasts, or primary vs immortalized/transformed cells). It is also necessary to use a cell type-specific culture medium that varies in combination of growth factors and serum concentrations. If primary cells are derived from human patients or animals, some commercial resources have a variety of culture media available for testing. Generally, cells are manipulated in a sterile hood, and the working surfaces are wiped with 70% to 80% ethyl alcohol solution. Cultured cells are usually maintained in a humidified 5% carbon dioxide incubator at 37°C (98.6°F) to maintain a pH value raging from 7.2 – 7.4 and need to be examined daily under an inverted microscope to check for possible contamination and confluency (the area cells occupy on the dish). In some cases, cells need to be maintained in hypoxia, and the oxygen input could be reduced to as low as 1%. As a general rule, cells should be fed with fresh medium every 2 to 3 days and split when they reach confluency. Depending on the growth rate of cells, the actual time and number of plates required to split cells in two varies from cell line to cell line. Splitting a monolayer requires the detachment of cells from plates by using a trypsin or collagenase treatment, of which concentration and time period vary depending on cell lines. If cultured cells grow continuously in suspension, they are split or subcultured by dilution.

Because cell lines may change their properties when cultured, it is not possible to maintain cell lines in culture indefinitely. Therefore, it is essential to store cells at various time passages for future use. The common procedure to use is cryopreservation. The solution for cryopreservation is usually fetal calf serum containing 10% dimethyl sulfoxide or glycerol, stored in liquid nitrogen (−196°C [−320.8°F]) for years of preservation. However, the viability and health of cells when thawed will decrease over time even in liquid nitrogen.

**Cell Transfection.** Cells are cultured for two reasons: to maintain and to manipulate them (see Fig. 15–21). The transfer of foreign macromolecules, such as nucleic acid, into living cells provides an efficient method for studying a variety of cellular processes and functions at the molecular level. DNA transfection has become an important tool for studying the regulation and function of genes. The cDNA to be expressed should be in a plasmid vector, behind an appropriate promoter working in mammalian cells (e.g., the constitutively active cytomegalovirus promoter or inducible promoter). Depending on the cell type, many ways of introducing DNA into mammalian cells have been developed. Commonly used approaches include calcium phosphate, electroporation, liposome-mediated transfection, the nonliposomal formulation, and the use of viral vectors. These methods have shown variable success when attempting to transfected a wide variety of cells. Transfection can be performed in the presence or absence of serum. It is suggested to test the transfection efficiency of cell lines of interest by comparing transfection with several different approaches. For a detailed transfection protocol, it is best to follow the manufacturer’s instructions for the particular reagent. General considerations for a successful transfection depend on several parameters, such as the quality and quantity of DNA and cell culture (type of cell and growth phase). To minimize variations in both of these in transfection experiments, it is best to use cells that are healthy, proliferate well, and are plated at a constant density.

Depending on the transfection method, DNA expression can be transient or stable. Using calcium phosphate and liposome-mediated transfection, after DNA is introduced into the cells it is normally maintained epitopically in cells and will be diluted while host cells undergo cell division. Therefore, functional assays should be performed 24 to 72 hours after transfection, also termed transient transfection. In many applications, it is important to study the long-term effects of DNA in cells by stable transfection. Thus, electroporation and viral vector are often used in these situations to enable integration of ectopic DNA into the host genome. Stable cell clones can be selected when plasmids carry an antibiotic-resistant marker. In the presence of antibiotics, only those cells that continuously carry the antibiotic-resistant marker (after generations of cell division) can survive. One application of stable transfection is the generation of transgenic or knockout mouse models, in which the transgene has to be integrated in the mouse genome in the ES cells, followed by microinjection of those transgenic ES cells into blastocysts to generate chimera mice. Stable cells also can be transplanted into host organs to test the effect of transgenic cells in vivo.

**Genetic Manipulations**

Understanding how genes control the growth and differentiation of the mammalian organism has been the most challenging topic of modern research. It is essential for us to understand how genetic mutations and chemicals lead to the pathologic condition of human bodies. The knowledge and ability to change the genetic program will inevitably make a great impact on society and have far-reaching effects on how we think of ourselves.

The mouse has become firmly established as the primary experimental model for studying how genes control mammalian development. Genetically altered mice are powerful tools to study the function and regulation of genes as well as modeling human diseases. The gene function can be studied by creating mutant mice through homologous recombination (gene knockout). A gene of interest (GOI) also can be introduced into the mouse (transgenic mouse) to study its effect on development or diseases. Because mouse models do not precisely represent human biology, genetic manipulations of human somatic or ES cells provide a great means for the understanding of the molecular networks in human cells in addition to mouse models. In all cases, the gene to be manipulated must first be cloned. Gene cloning has been made easy by recombinant DNA technology and the availability of human and mouse genomes (see “Human Genome”). The following section briefly describes the technologies and the principles behind combining both mouse genetics and human cell culture to explore gene function and disease mechanisms.

**Transgenic Mice.** During the past 20 years, DNA cloning and other techniques have allowed the introduction of new genetic material into the mouse germline. As early as 1980, the first genetic material was successfully introduced into the mouse germline by using pronuclear microinjection of DNA (Fig. 15–22). These animals, called transgenic, contain foreign DNA within their genomes. In simple terms, a transgenic mouse is created by the microinjection of ectopic DNA into the one-celled mouse embryo to induce integration, allowing the
efficient introduction of cloned genes into the following developing mouse somatic tissues, as well as into the germline.

**Designs of a Transgene** The transgenic technique has proven to be extremely important for basic investigations of gene regulation, creation of animal models of human disease, and genetic engineering of livestock. The design of a transgene construct is a simple task. Like constructs used in cell transfection, a simple transgene construct consists of a protein-encoding gene and a promoter that precedes it. The most common applications for the use of transgenic mice are similar to those in the cell culture system: (a) to study the functions of proteins encoded by the transgene, (b) to analyze the tissue-specific and developmental-stage–specific activity of a gene promoter, and (c) to generate reporter lines to facilitate biomedical studies. Examples of the first application include overexpression of oncogenes, growth factors, hormones, and other key regulatory genes, as well as genes of viral origins. Overexpression of the transgene normally represents gain-of-function mutations. The tissue distribution or expression of a transgene is determined primarily by cis-acting promoter enhancer elements within or in the immediate vicinity of the genes themselves. Thus, controlled expression of the transgene can be made possible by using an inducible or tissue-specific promoter. Furthermore, transgenic mice carrying dominant negative mutations of a regulatory gene have also been generated. For example, a truncated growth factor receptor that can bind to the ligand, but loses its catalytic activity when expressed in mice, can block the growth factor binding to the endogenous protein. In this way, the transgenic mice exhibit a loss of function of phenotype, possibly resembling the knockout of the endogenous gene. The second application of the transgenic expression is to analyze the gene promoter of interest. The gene promoter of interest normally is fused to a reporter gene that encodes β-galactosidase (also called LacZ), luciferase, or green fluorescence protein. Chemical staining of LacZ activity or detection of chemiluminescence/fluorescence can easily visualize the expression of the reporter gene. The third application originates from the second: when the activity of the promoter is known, a fluorescent reporter gene (such as GFP) will be driven by the tissue-specific promoter, therefore labeling a particular type of cells at a particular stage. This application is generally used to isolate a special cell type expressing the GFP reporter by fluorescence-activated cell sorting (FACS), as well as lineage-tracing experiments.

**Production of Transgenic Mice** The success of generating transgenic mice is largely dependent on the proper quality and concentration of the DNA supplied for microinjection. For DNA to be microinjected into mouse embryos, it should be linearized by restriction digestion to increase the chance of proper transgene integration. Concentration of DNA should be accurately determined. Mice that develop from injected eggs often are termed *founder* mice.

**Genotyping of Transgenic Mice** The screening of founder mice and the transgenic lines derived from the founders is accomplished by determining the integration of the injected gene into the genome. This normally is achieved by performing PCR or Southern blot analysis with a small amount of DNA extracted from the mouse tail. Once a given founder mouse is identified to be transgenic, it will be mated to begin establishing a transgenic line. Usually, for a given gene, more than one transgenic line is generated to assure that the phenotype is due to transgene but not to the interruption of the gene where the transgene integrates into.

**Analysis of Phenotype of Transgenic Mice** Phenotypes of transgenic mice are dictated by both the expression pattern and biologic functions of the transgene. Depending on the promoter and the transgene, phenotypes can be predictable or unpredictable. Elucidation of the functions of the transgene-encoded protein in vitro often offers some clue to what the protein might function to do in vivo. When a constitutively active promoter is used to drive the expression of transgenes, mice should express the gene in every tissue; however, this mouse model may not allow the identification and study of the earliest events in disease pathogenesis. Ideally, the use of tissue-specific or inducible promoter allows one to determine if the pathogenic protein leads to a reversible or irreversible disease process in a cell-autonomous manner. For example, rat insulin promoter can target transgene expression exclusively in the β-cells of pancreatic islets. The phenotype of insulin promoter-mediated transgenic mice is projected to affect the function of human β-cells.

**Gene Knockout in Mice.** The first recorded knockout mouse was created by Mario R. Capecchi, Sir Martin J. Evans, and Oliver Smithies in 1989. They were awarded the 2007 Nobel Prize in Physiology or Medicine. The isolation and genetic manipulation of mouse ES cells represents one of the most important milestones for modern genetic technologies. Several unique properties of ES cells, such as the pluripotency to differentiate into all germ layers in an embryo, including the germline, make them an efficient vehicle to introduce genetic alterations in mice. An important breakthrough from this idea is to generate gene-targeted mutation in mice, first by introducing the targeting vector into the ES cells, allowing selection...
in tissue culture

Altered version
of target gene
constructed by

genetic engineering

Let each cell
grow to form
a colony

Test for the rare
colony in which
the DNA fragment
has replaced
one copy of the
normal gene

ES cells with one copy of target gene replaced by mutant gene

Inject ES cells
into early embryo

Early embryo partly formed from ES cells

Introduce early embryo into pseudopregnant mouse

Somatic cells of offspring tested for presence of altered gene, and selected mice bred to test for gene in germline cells

Transgenic mouse with one copy of target gene replaced by altered gene in germline

A

B

Figure 15-23. Knockout mouse technology. Summary of the procedures used for making gene replacements in mice. In the first step (A), an altered version of the gene is introduced into cultured embryonic stem (ES) cells. Only a few rare ES cells will have their corresponding normal genes replaced by the altered gene through a homologous recombination event. Although the procedure is often laborious, these rare cells can be identified and cultured to produce many descendants, each of which carries an altered gene in place of one of its two normal corresponding genes. In the next step of the procedure (B), these altered ES cells are injected into a very early mouse embryo; the cells are incorporated into the growing embryo, and a mouse produced by such an embryo will contain some somatic cells that carry the altered gene. Some of these mice also will contain germline cells that contain the altered gene. When bred with a normal mouse, some of the progeny of these mice will contain the altered gene in all of their cells. If two such mice are in turn bred (not shown), some of the progeny will contain two altered genes (one on each chromosome) in all of their cells. If the original gene alteration completely inactivates the function of the gene, these mice are known as knockout mice. When such mice are missing genes that function during development, they often die with specific defects long before they reach adulthood. These defects are carefully analyzed to help decipher the normal function of the missing gene.


for successful homologous recombination in a dish, then introducing the selected ES clone into the blastocysts, and finally recovering animals bearing the mutant allele from the germline (Fig. 15-23). This not only makes mouse genetics a powerful approach to address important gene functions, but also identifies the mouse as a great system to model human disease.

Targeting Vector The basic concept in building a target vector to knock out a gene is to use two segments of homologous sequence to a GOI that flank a part of the gene essential for functions (e.g., the coding region). In the targeting vector, a positive selectable marker (e.g., the neo gene) is placed between the homology arms. Upon the homologous recombination between the arms of the vector and the corresponding genomic regions of the GOI in ES cells, the positive selectable marker will replace the essential segment of the target gene, thus creating a null allele. In addition, a negative selectable marker also can...
be used alone or in combination with the positive selectable marker, but must be placed outside of the homologous arms to enrich for homologous recombination. To create a conditional knockout (i.e., gene knockout in a spatiotemporal fashion), site-specific recombinases such as the popular cre-loxP system are used. If the consensus loxP sequences that are recognized by cre recombinases are properly designed into targeting loci, controlled expression of the recombinase as a transgene can result in the site-specific recombination at the right time and in the right place (i.e., cell type or tissue). This method, often referred to as conditional knockout, is markedly useful to prevent developmental complications and to introduce null mutations in the adult mouse that would otherwise be lethal. To bring in additional control to tissue-specific cre, an inducible cre could be adopted on top of the tissue-specific promoter. The most popular inducible cre includes CreER: CreER encodes a cre fused with estrogen receptor located to cytosol. With the signaling of tamoxifen, CreER is released and then translocates into the nucleus to induce the recombination of loxP. Therefore, the timing of recombination could be precisely determined by controlling the time of administering tamoxifen. Overall, this cre-loxP system allows for spatial and temporal control over transgene expression and takes advantage of inducers with minimal pleiotropic effects.

**Introduction of the Targeting Vector into ES Cells** ES cell lines can be obtained from other investigators or commercial sources or established from blastocyst-stage embryos. To maintain ES cells at their full developmental potential, optimal growth conditions should be provided in culture. If culture conditions are inappropriate or inadequate, ES cells may acquire genetic lesions or alter their gene expression patterns and consequently decrease their pluripotency. Excellent protocols are available in public domains or in mouse facilities in most institutions.

To alter the genome of ES cells, the targeting vector DNA then is transfected into ES cells. Electroporation is the most widely used and the most efficient transfection method for ES cells. Similar procedures for stable cell transfection are used for selecting ES cells that carry the targeting vector. High-quality, targeting-vector DNA free of contaminating chemicals is first linearized and then electroporated into ES cells. Stable ES cells are selected in the presence of a positive selectable antibiotic drug. After a certain period of time and depending on the type of antibiotics, all sensitive cells die, and the resistant cells grow into individual colonies of the appropriate size for subcloning by picking. It is extremely important to minimize the time during which ES cells are in culture between selection and injection into blastocysts. Before injecting the ES cells, DNA is prepared from ES colonies to screen for positive ES cells that exhibit the correct integration or homologous recombination of the targeting vector. Positive ES colonies are then expanded and used for creation of chimeras.

**Creation of the Chimera** A chimeric organism is one in which cells originate from more than one embryo source. Here, chimeric mice are denoted as those that contain some tissues from the ES cells with an altered genome. When these ES cells give rise to the lineage of the germ layer, the germ cells carrying the altered genome can be passed on to the offspring, thus creating the germline transmission from ES cells. There are two methods for introducing ES cells into preimplantation-stage embryos: injection and aggregation. The injection of embryonic cells directly into the cavity of blastocysts is one of the fundamental methods for generating chimeras, but aggregation chimeras also have become an important alternative for transmitting the ES cell genome into mice. Since every tissue type of a chimera should contain cells from different origins, the mixture of recognizable markers (e.g., coat color) that are specific to the donor mouse and the ES cells can be used to identify chimeric mice. However, most experimenters probably use existing mouse core facilities already established in some institutions or contract a commercial vendor for the creation of a chimera.

**Genotyping and Phenotyping of Knockout Animals** The next step is to analyze whether germline transmission of targeted mutation occurs in mice. DNA from a small amount of tissue from offspring of the chimera is extracted and subjected to genomic PCR or Southern blot DNA hybridization. Positive mice (i.e., those with properly integrated targeting vector into the genome) will be used for the propagation of more knockout mice for phenotype analysis. When the knockout genes are crucial for early embryogenesis, mice often die in utero, an occurrence called embryonic lethality. When this happens, only the phenotype of the homozygous (both alleles ablated) knockout mouse embryos and the phenotype of the heterozygous (only one allele ablated) adult mice can be studied. Because most researchers are interested in the phenotype of adult mice, in particular when using mice as disease models, it is recommended to create the conditional knockout using the cre-loxP system so that the GOI can be knocked out at will.

To date, more than 5000 genes have been disrupted by homologous recombination and transmitted through the germline. The phenotypic studies of these mice provide ample information about the functions of these genes in growth and differentiation of organisms and during development of human diseases.

**RNA Interference.** Although gene ablation in animal models provides an important means to understand the in vivo functions of GOI, animal models may not adequately represent human biology. Alternatively, gene targeting can be used to knock out genes in human cells, including human ES cells. Gene targeting in human ES cells by homologous recombination has extremely low efficiency, although there are more new techniques emerging at increasing the targeting efficiency. A number of recent advances have made gene targeting in somatic cells as easy as in murine ES cells. However, gene targeting (knocking out both alleles) in somatic cells is a time-consuming process.

Development of RNAi technology in the past few years has provided a more promising approach to understanding the biologic functions of human genes in human cells. RNAi is an ancient natural mechanism by which small, double-stranded RNA (dsRNA) acts as a guide for an enzyme complex that destroys complementary RNA and downregulates gene expression in a sequence-specific manner. Although the mechanism by which dsRNA suppresses gene expression is not entirely understood, experimental data provide important insights. In nonmammalian systems such as Drosophila, it appears that longer dsRNA is processed into 21–23 nt dsRNA (called small interfering RNA or siRNA) by an enzyme called Dicer containing RNase III motifs. The siRNA apparently then acts as a guide sequence within a multicompartment nuclease complex to target complementary mRNA for degradation. Because long dsRNA induces a potent antiviral response pathway in mammalian cells, short siRNAs are used to perform gene silencing experiments in mammalian cells (Fig. 15-24).
For siRNA studies in mammalian cells, researchers have used two 21-mer RNAs with 19 complementary nucleotides and 3′ terminal noncomplementary dimers of thymidine or uridine. The antisense siRNA strand is fully complementary to the mRNA target sequence. Target sequences for an siRNA are identified visually or by software.

The target 19 nucleotides should be compared to an appropriate genome database to eliminate any sequences with significant homology to other genes. Those sequences that appear to be specific to the GOI are the potential siRNA target sites. A few of these target sites are selected for siRNA design. The antisense siRNA strand is the reverse complement of the target sequence. The sense strand of the siRNA is the same sequence as the target mRNA sequence. A deoxythymidine dimer is routinely incorporated at the 3′ end of the sense strand siRNA, although it is unknown whether this noncomplementary dinucleotide is important for the activity of siRNAs.

There are two ways to introduce siRNA to knock down gene expression in human cells:

1. RNA transfection: siRNA can be made chemically or using an in vitro transcription method. Like DNA oligos, chemically synthesized siRNA oligos can be commercially ordered. However, synthetic siRNA is expensive, and several siRNAs may have to be tried before a particular gene is successfully silenced. In vitro transcription provides a more economic approach. Both short and long RNA can be synthesized using bacteriophage RNA polymerase T7, T3, or SP6. In the case of long dsRNAs, RNase such as recombinant Dicers will be used to process the long dsRNA into a mixture of 21–23 nt siRNA. siRNA oligos or mixtures can be transfected into a few characterized cell lines such as HeLa (human cervical carcinoma) and 293T cells (human kidney carcinoma). Transfection of siRNA directly into primary cells may be difficult.

2. DNA transfection: Expression vectors for expressing siRNA have been made using RNA polymerase III promoters such as U6 and H1. These promoters precisely transcribe a hairpin structure of dsRNA, which will be processed into siRNA in the cell (see Fig. 15-24). Therefore, properly designed DNA oligos corresponding to the desired siRNA will be inserted downstream of the U6 or H1 promoter. There are two advantages of the siRNA expression vectors over siRNA oligos. First, it is easier to transfect DNA into cells. Second, stable populations of cells can be generated that maintain the long-term silencing of target genes. Furthermore, the siRNA expression cassette can be incorporated into a retroviral or adenoviral vector to provide a wide spectrum of applications in gene therapy.

There has been a fast and fruitful development of RNAi tools for in vitro and in vivo use in mammals. These novel approaches, together with future developments, will be crucial to put RNAi technology to use for effective disease therapy or to exert the awesome power of mammalian genetics. Therefore, the applications of RNAi to human health are enormous. siRNA can be applied as a new tool for sequence-specific regulation of gene expression in functional genomics and biomedical studies. With the availability of the human genome sequences, RNAi approaches hold tremendous promise for unleashing the dormant potential of sequenced genomes.

There is the potential of sequenced genomes. RNAi approaches hold tremendous promise for unleashing the dormant potential of sequenced genomes.

Figure 15-24. RNA interference in mammalian cells. Small interfering RNA (siRNA) can be produced from a polymerase III–driven expression vector. Such a vector first synthesizes a 19–29 nt double-stranded (ds)RNA stem and a loop (labeled as shRNA in the figure), and then the RNase complex called Dicer processes the hairpin RNA into a small dsRNA (labeled as siRNA in the figure). siRNA can be chemically synthesized and directly introduced into the target cell. In the cell, through RNA-induced silencing complex (RISC), siRNA recognizes and degrades target messenger RNAs (mRNAs).
the human genome blueprint, the promise of gene therapy and molecular therapies, and the existence of stem cells have captured the imagination of the public and the biomedical community. Aside from their potential in curing human diseases, these emerging technologies also have provoked many political, economic, religious, and ethical discussions. As more is discerned about the technologic scientific advances, more attention must also be paid to concerns for their inherent risks and social implications. It is important for surgeons to play a leadership role in the emergence of personalized medicine and surgery, as surgeons have access to the diseased tissues. Surgeons should be establishing collaborations with the genomic and molecular scientists to develop genomic biobanks in order to study the genome and molecular signaling of the disease tissues that will help with an understanding of the underlying cause of an individual’s disease and ultimately lead to effective, targeted therapies. Surgeons must take this enormous opportunity to collaborate with basic and clinical scientists to develop the field of precision medicine and surgery this century.

Bifunctional RNAi Technology. Over the last 20 years, the field has worked to define oncogene and nononcogene addiction, discriminate between driver and passenger genes, and appreciate the complexity of complex, robust, network interactions. These insights have led to a preliminary understanding of therapeutically relevant sensitivity and resistance pathway signal patterns requiring multiple target modulation. However, this knowledge has not been effectively or reproducibly clinically translated. Clinical response is usually far greater when a combination of single-target molecular therapy is administered. However, it must also be realized that targeting two or more pathways may also increase the toxicity profile, particularly if target specificity is limited. When attempted, off-target toxicity has been demonstrated with combination small-molecule therapy. In contrast, multitargeting bifunctional short hairpin (bi-shRNA) DNA vectors are designed to limit off-target effect given the high specificity for the genes they are designed to target.

Exogenously applied hairpin constructs can be designed to be incorporated into cleavage-dependent RISC or cleavage-independent RISC complexes, or both. The concept of a bifunctional shRNA is to increase knockdown efficiency without loss of sequence specificity by engaging both siRNA and miRNA-like (i.e., common biogenic pathway but complementary to target sequence) RISCs, thereby concurrently activating nucleolytic (Ago2-RISC) and nonnucleolytic (Ago1, 3, 4 ± Ago2-RISC) processes. Each bi-shRNA contains both a matched stem sequence to promote Ago2-mediated passenger strand cleavage and a second partial mismatched stem sequence for cleavage-independent passenger strand departure. Thus, functionality of the effectors is set by programmed passenger strand guided RISC loading rather than Ago subset distribution in the cancer cell. Both component Ago2 and Ago (1, 2, 4 ± 3) RNAi moieties are fully complementary to the mRNA target sequence. Preliminary data indicate reduced “off-target effects” by shRNA compared with target-identical siRNAs. More than two mismatches in sequences within the target region drastically reduce knockdown effect to undetectable levels (unpublished results). The design process involves in silico scanning of the entire human mRNA RefSeq database to avoid any potential sequence-related “off-target effects.” Published data also indicate persistent susceptibility to shRNA-mediated gene knockdown despite recent evidence of reduced Dicer expression in human cancer cells.

The first clinical experience with the bi-shRNA platform involved the ex vivo knockdown of furin, a Ca2+-dependent, nonredundant proprotein convertase that is essential for proteolytic maturational processing of immunosuppressive TGF-β isoforms (β1 and β2). An autologous whole-cell cancer vaccine, FANG™ (furin-knockdown and GMCSF-augmented), was produced based on a dual function immunosensitization principle of augmenting tumor antigen expression, presentation, and processing via granulocyte-macrophage colony-stimulating factor (GMCSF) cytokine transgene expression and attenuating secretory immunosuppressive TGF-β. Harvested, autologous cancer cells are transfected with the GMCSF/bi-shRNAfurin (FANG) expression plasmid via electroporation. A phase I clinical trial (BB-IND 14205) involving 52 cancer patients was recently completed. Results demonstrated better than 90% knockdown of the bi-shRNA target, furin, and better than 90% knockdown of furin-regulated proteins TGF-β1 and TGF-β2, thereby confirming the mechanistic expectation of this novel RNAi platform. Moreover, predicted extensive GMCSF expression verified our ability to successfully construct multi-cassette vectors with good manufacturing practice techniques fulfilling Food and Drug Administration requirements for clinical testing.

Twenty-seven patients received one or more vaccine dose, and 23 patients achieved stable disease as their best response. No toxic effect was identified. Median survival of the FANG™-treated patients from time of procurement was 554 days and has not been reached from time of treatment. Expected survival of similar patients is historically less than 1 year. Sequential enzyme-linked immunosorbent spot (ELISPOT) analysis revealed a dramatic and significant increase in immune response from baseline to month 4 in half of the FANG™-treated patients. Comparison of survival between ELISPOT-positive and ELISPOT-negative patients demonstrated a statistically significant increase in survival from time of procurement (P = .045) and time of treatment (P = .025).

These phase 1 study results demonstrated mechanism, safety, and effectiveness of the bi-shRNA technology and clinical functionality of a multitargeting (dual) DNA expression vector. Further utilization of bi-shRNAi technology is under way clinically (targeting STMN1, a microtubule modulation critical to cancer program) and preclinically targeting PDX1 (an oncogene-like transcription factor for pancreatic embryogenesis using nonviral nanoparticle delivery mechanisms).

Precision Medicine and Surgery

Genes determine our susceptibility to diseases and direct our body’s response to medicine. Because an individual’s genes differ from those of another, the determination of each individual’s genome has the potential to improve the predication, prevention, and treatment of disease. Sequencing of individual genomes holds the key to realize this revolution called precision medicine and surgery. Next-generation sequencing, such as Illumina sequencing and 454 pyrosequencing technology, is promising to reduce the time and cost so that genome sequencing can be affordable within healthcare systems. The goal of precision medicine and surgery is to identify the gene variations in each individual and to target the specific gene variations causing the disease by choosing personalized treatments that effectively work in association with the individual’s genomic profile. The importance of surgeons in this transformational field of
biomedical science is that surgeons have access to the diseased tissues on a daily basis. Surgeons should partner with the genomic scientists to develop genomic biobanks in order to study the genome of the disease tissues and determine how this information can improve the outcomes of surgery, i.e., precision surgery. These discovery studies are rapidly leading to the uncovering of mutations and SNPs that are the underlying cause of an individual’s disease and ultimately lead to targeted therapies. Although precision medicine and surgery holds the potential to revolutionize the practice of modern medicine, there currently exists a gap between our ability to sequence any given individual’s genome and how clinicians can apply this information to guide care. There is a rapidly growing list of single genes that are currently guiding care, and these genes are listed as type 1 precision genes. Examples of these genes are *BRCA1*, *RET* proto-oncogene, and *CHD1* mutation, which guide potential use of mastectomy, thyroidectomy, and gastrectomy, respectively; however, the great challenge before the scientific and medical community this century is to learn to use the entire genome to guide precision care.

**Targeted Genome Editing Using the CRISPR-Cas9 System**

Conventional genetic manipulations have proven their value in biomedical research. Researchers today depend on the manipulation of genetic materials in cells or in animal models in almost every project they work on. These genetic manipulation techniques (see “Genetic Manipulations”), though sufficient for general research purposes, suffer from disadvantages. Transfection of target genes into cells is quick and specific, but nonnative. RNA interference (RNAi) is easy to perform and targets native genes, but RNAi never fully eliminates the target gene, and off-target effects are commonly seen using RNAi. Gene knockout mice provide an ideal platform to study native genes with clean background, but conventional knockout methods are time-consuming and costly.

An entirely new gene-editing method now known as the CRISPR-Cas9 system has emerged since 2013 and has quickly gained popularity among biologists for gene editing. This new method is easy to perform, can work specifically on the desired gene or DNA sequence, and can generate gene knockout, knock-in, point mutation, or the insertion of an epitope tag in almost any cell line or animal models with high efficiencies.

CRISPR stands for Clustered Regularly Interspaced Short Palindromic Repeats. It is a region on the genomic DNA first discovered in the microbes as an adapted immune system against exogenous DNA. A typical CRISPR region contains a cluster of short (21–48 bp) DNA repeats (ranging from 2 to hundreds) interspaced by nonrepetitive sequences called spacers. Within a CRISPR region, while each spacer has its unique sequence, the sequence of the repeats is highly conserved. Several genes, called the CRISPR-associated (Cas) genes, are almost always found directly flanking the CRISPR region.

Extensive studies in the past decade revealed the function of the CRISPR-cas system in DNA-interfering. When bacteria or archaea carrying the CRISPR-cas system are invaded by phage or plasmid DNA, a new spacer can be added to the CRISPR region with its sequence identical to the “proto-spacer,” a fragment of the invading DNA. It was found that the proto-spacer must be followed by a recognition sequence (NGG in the case of Cas9) called proto-spacer associated motif (PAM). With constitutive transcription, the CRISPR region is transcribed as mRNA, and cut by Cas proteins to generate RNA fragments minimally containing one spacer and parts of the repetitive sequence. This fragment associates with target DNA through Watson-Crick base-pairing and directs the cutting of target DNA by Cas proteins with nuclease activity.

Several CRISPR-Cas systems have been characterized, with a variety of Cas proteins found in these systems. Cas9 in the type II system from *Streptococcus pyogenes* is the most commonly used Cas for gene editing. Cas9 contains one RuvC-like nuclease domain near its N-terminal and one HNH-like nuclease domain in the middle of the protein. The RuvC-like domain cuts the proto-spacer strand (the strand with PAM), and the HNH-like domain cuts the strand pairing with the spacer, resulting in a double-strand break (DSB) on the target DNA. It is known that Cas9 cuts a blunt end 3 bp 5’ to the PAM sequence. The specificity provided by the spacer, or so-called “guide RNA,” and the ability of Cas9 to cut double-stranded DNA means that this system can specifically target anywhere in a genome with a known sequence.

**CRISPR-Cas9–Guided Gene Editing.** The CRISPR-Cas9 system was made suitable for gene editing in mammalian cell lines and in animal models through several years of optimization. The key concept for CRISPR-Cas9–mediated gene editing is to introduce DNA strand break and let the cell repair the break. Through the repair process, sequence deletions, insertions, and mutations can be applied to the target gene.

Two DNA repair pathways are utilized in CRISPR-Cas9–mediated gene editing: the nonhomologous end-joining (NHEJ) pathway and the homology-directed repair (HDR). When a homologous repair template is unavailable, the NHEJ pathway joins the ends of the DSB together, usually with random insertion/deletion (indel) mutations. Such mutations within an open reading frame can cause frameshift and/or premature stop codons. When a repair template is available, however, the cell may choose HDR and repair the DSB through pairing with the template. The HDR pathway can introduce longer insertion/deletion mutations than NHEJ and can specify the mutated sequence (by contrast, NHEJ creates random mutations). However, HDR is usually only active in dividing cells and is of a lower efficiency than NHEJ. Its efficiency also depends on gene location, cell type, and the repair template. Therefore, the choice of the pathway depends on the need of the outcome: to simply create a gene knockout, NHEJ is much simpler and highly efficient; to achieve precise gene editing (introduce a specific mutation, add or delete a specific sequence), HDR must be used. The two pathways have similar protocols, but they differ in certain details in experimental designs.

**Gene Editing Through NHEJ** Tools for NHEJ-mediated gene editing can be incorporated onto one simple plasmid, including a single-guide RNA (sgRNA) sequence containing the guide RNA, a U6 promoter driving the expression of the sgRNA, and an expression cassette including codon-optimized Cas9 fused with nuclear localization sequences (NLS) and an optional selection marker (puromycin resistance or a GFP). NHEJ is usually used to knock out a target gene (Fig. 15-24). Generally, a 20 bp sequence preceding a PAM sequence (NGG for Cas9) is selected as the guide RNA sequence. This sequence can be chosen within the first couple of exons through running online search engines against the target genome to minimize off-target probabilities. The guide RNA sequence is then inserted into the
Figure 15-25. CRISPR-Cas9–mediated gene editing through NHEJ or HDR. The Cas9 protein cleaves target DNA through its RuvC-like and HNH-like domains, guided by the sgRNA. The cell repairs the double-strand break through either NHEJ or HDR. In NHEJ, the broken ends are recognized, bound, and tethered by end-binding protein complexes. The ends are then processed and ligated but may result in random insertion/deletion (experimentally, deletions are more commonly seen) mutations at the break site. The repair process in HDR, however, generates no random error due to the presence of a homologous repair template. During the repair, the broken strands find the homologous template and proceed with DNA synthesis using the template. This results in a repaired DNA that has the same sequence as the repair template. Therefore, point mutations, insertions, or deletions carried on the template will be inherited by the repaired DNA and thus achieve precise gene editing.

Designed site within the sgRNA, and the constructed plasmid is transfeeted into target cells. Expressed Cas9 proteins would associate with expressed sgRNAs to mediate DSB directed by the guide RNA sequence. For cell lines, pure mutant clones can be generated by separating single colonies. Knockout mutations can be confirmed through PCR amplification of the target region be generated by separating single colonies. Knockout mutations can be confirmed through PCR amplification of the target region and sequencing. If antibodies against the target protein are available, western blotting can be used as a supplemental method to DNA sequencing results to confirm the knockout of the gene.

Gene Editing Through HDR HDR requires a homologous template to repair the DSB. Therefore, apart from the guide RNA, another sequence homologous to the flanking regions of the DSB needs to be present (Fig. 15-25). The homologous sequence can carry insertions, deletions, and mutations to replace the sequence at the DSB, thus achieving precise gene editing on the target. The homologous sequence can be introduced into the cell either as a template on a plasmid or as a single-stranded oligonucleotide. After the induction of the DSB, the homologous template pairs with the flanking regions of the DSB and serves as the template for the repair of the break site. HDR happens at lower efficiency than NHEJ and therefore is of lower success rate than NHEJ-mediated gene editing. Therefore, HDR is only recommended when precise mutations are desired. Fast screen of positive mutant clones can be achieved by incorporating restriction enzyme cutting sites within the homologous template.

Reducing Off-Target Effects Using Cas9 Nickase The CRISPR-Cas9 system uses a 20 bp guide RNA for sequence recognition. Due to the similar length of recognition sequence to RNAi techniques, CRISPR-Cas9 system also suffers from off-target effects. Because CRISPR-Cas9 requires a PAM site directly following the guide RNA sequence, one way to reduce off-targets is to run through online databases against the genome of the target cell to select a target sequence with the least number of possible off-targets (CRISPR-Cas9 system tolerates no more than three mismatches). On the other hand, a mutant version of the Cas9 called Cas9 nickase can be used to minimize the risk of off-targets.

As mentioned, Cas9 contains one RuvC-like and one HNH-like nuclease domain, each responsible for cutting one strand. The D10A mutant Cas9 (Cas9 nickase) lacks the activity of the RuvC-like domain, leaving the proto-spacer strand intact and a nick on the antisense strand. Using two properly spaced (0–20 bp apart), oppositely oriented guide RNA, the Cas9 nickase will leave two single-strand nicks on both strands in close proximity, creating a double-strand break with 5’ overhangs on both strands. This leads to NHEJ or HDR at this breaking site, while other off-target single-strand nicks will be repaired without impact. This strategy doubles the number of base pairs required for site recognition, reducing off-target possibilities almost to zero (from about 1 off-target in 30 million bp to 1 in 1000 trillion bp).

Application of the CRISPR-Cas9 System in Biomedical Sciences The biggest advantage of the CRISPR-Cas9 system is its ability to edit genes in almost any cell type and any animal model with high efficiency and accuracy. Plus, it is easy to design and easy to use. Within a couple of years, successful gene editing in C. elegans, zebrafish, fruit fly, mouse, dogs, and even nonhuman primates was achieved. CRISPR-Cas9 was also reported to be successful in a variety of cell types including stem cells.

Currently, CRISPR-Cas9 is most used for editing single genes, through gene knockout, gene mutation, or the addition of an epitope tag to a native gene, for functional characterization of the gene of interest. For example, oncogenes or tumor suppressor genes can be knocked out to identify the causative gene for a particular cancer type; point mutations in functional domains
may illustrate the mechanism of action of a protein; for proteins without available antibodies, epitope tags can be inserted onto the native gene for the detection of the native protein.

Apart from single gene editing, CRISPR-Cas9 can be used for large scale loss-of-function gene screen. Multiple lentiviral guide RNA libraries have been established covering the whole human/mouse genome or particular subsets.

Owing to its ability to bind to specific nucleotide sequences, CRISPR-Cas9 has also been used for non–gene-editing purposes. To achieve this, both the RuvC-like and the HNH-like domains are mutated to give a catalytically inactive Cas9 (dCas9). Directed by guide RNA, this dCas9 can bind to particular genes to reversibly suppress or activate gene transcription by the fusion of transcription activators or suppressors with dCas9. Epigenetic modulators (e.g., DNA methylase) can also be fused with dCas9 to achieve controlled epigenetic modulations. Moreover, dCas9 can also be fused with fluorescent markers such as GFP to track a particular DNA region in live cells.

The most exciting potential application of the CRISPR-Cas9 system is perhaps the correction or modification of disease-causing genes in human embryos or in human patients to eradicate disease-causing genes. However, extension of this application may lead to the creation of the so called “perfect human,” hence raising huge ethical concerns and controversies.57 Because of such ethical considerations, gene editing in human embryos should be cautiously conducted. Nonetheless, reports have shown the great potential of this powerful gene-editing technique in correcting gene mutations.58-60

In conclusion, the CRISPR-Cas9 system is the most powerful gene-editing system characterized so far, showing its strong ability to edit genes efficiently and precisely. Its applications greatly benefit biomedical researches and, with a solution to ethical issues, can greatly benefit clinical medicine as well.

REFERENCES

Entries highlighted in bright blue are key references.


This page intentionally left blank
This page intentionally left blank
INTRODUCTION

The skin is a complex organ encompassing the body's surface and is continuous with the mucous membranes. Accounting for approximately 15% of total body weight, it is the largest organ in the human body. Enabled by an array of tissue and cell types, intact skin protects the body from external insults. However, the skin is also the source of a myriad of pathologies that include inflammatory disorders, mechanical and thermal injuries, infectious diseases, and benign and malignant tumors. The intricacies and complexities of this organ and associated pathologies are reasons the skin and subcutaneous tissue remain of great interest and require the attention of various surgical disciplines that include plastic surgery, dermatology, general surgery, and surgical oncology.

ANATOMY AND HISTOLOGY

Background

It is important that surgeons understand completely the cutaneous anatomy and its variability as they play an enormous role in patient health and satisfaction. The skin is made up of tissues derived from both the ectodermal and mesodermal germ cell layers. Three distinct tissue layers comprise the organ, and differ in composition based on location, age, sex, and ethnicity, among other variables. The outermost layer is the epidermis, which is predominantly characterized by a protective, highly keratinized layer of cells. The next layer is the dermis, which is made up of an organized collagen network to support the numerous epidermal appendages, neurovascular structures, and supportive cells within the skin. The fatty layer below the dermis is collectively known as the hypodermis and functions in body processes of thermoregulation and energy storage, among others. These three distinct layers function together harmoniously and participate in numerous activities essential to life.

Epidermis

The epidermis is the outermost layer of the cutaneous tissue, and consists primarily of continually regenerating keratinocytes. The tissue is also stratified, forming four to five histologically distinct layers, depending on the location in the body. These layers are, from deep to superficial, the stratum basale, stratum spinosum, stratum granulosum, stratum lucidum and stratum corneum (Fig. 16-1). The different layers of the epidermis represent layers of keratinocytes at differing stages of their approximately thirty-day life cycle. A minority of other cell types are found in different layers of the epidermis as well. Some of these cells are permanent residents, while others are visitors from other parts of the body. All the epidermal appendages, such as sweat glands and pilosebaceous follicles, are derived from this tissue. The thickness of the epidermis is quite variable with regard to location and age, ranging from 75 to 150 µm in thin skin (eyelids) to 0.4 to 1.5 mm in thick skin (palms and soles). The epidermis lacks any vascular
Key Points

1. The epidermis consists of continually regenerating stratified epithelium, and 90% of cells are ectodermally derived keratinocytes.

2. Pilosebaceous units are lined by the germinal epithelium of the epidermis and thus serve as an important source of epidermal regeneration after partial-thickness injury or split-thickness skin graft.

3. Dermal fibers are predominantly made of type I and III collagen in a 4:1 ratio. They are responsible for the mechanical resistance of skin.

4. The drugs most commonly associated with epidermal necrolysis include aromatic anticonvulsants, sulfonamides, allopurinol, oxicams (nonsteroidal anti-inflammatory drugs), and nevirapine.

5. In wounds being allowed to heal secondarily, negative pressure wound therapy can increase the rate of granulation tissue formation.

6. Staphylococcus aureus is the most common isolate of all skin infections. Impetigo, cellulitis, erysipelas, folliculitis, furuncles, and simple abscesses are examples of uncomplicated infections, whereas deep-tissue infections, extensive beginning of the life cycle of the keratinocytes of the epidermis. This layer is bound to its basement membrane by complexes made of keratin filaments and anchoring structures called hemidesmosomes. They are bound to other keratinocytes by structures called desmosomes. High mitotic activity and thus large nuclei and basophilic staining characterize the stratum basale on light microscopy. This layer also lines the epidermal appendages that reside largely within the substance of the dermis and later serves as a regenerative source of epithelium in the event of partial thickness wounds.

7. Hemangiomas arise from benign proliferation of endothelial cells surrounding blood-filled cavities. They most commonly present after birth, rapidly grow during the first year of life, and gradually involute in most cases.

8. Basal cell carcinoma represents the most common tumor diagnosed in the United States, and the nodular variant is the most common subtype. The natural progression of basal cell carcinoma is one of local invasion rather than distant metastasis.

9. Squamous cell carcinoma is the second most common skin cancer, and typically arises from an actinic keratosis precursor. Primary treatment modalities are surgical excision and Mohs microsurgery. Cautery and ablation, cryotherapy, drug therapy, and radiation therapy are alternative treatments.

10. Tumor thickness, ulceration, and mitotic rate are the most important prognostic indicators of survival in melanoma. Sentinel lymph node biopsy is often used to stage individuals with biopsy-proven high risk melanoma and clinically node-negative disease.

Figure 16-1. Schematic representation of the skin and its appendages. Note that the root of the hair follicle may extend beneath the dermis into the subcutis.
The next layer is the stratum spinosum, or “spiny” layer. This layer is from five to fifteen cells in thickness and is so named due to the spinous appearance of the intercellular desmosomal attachments under light microscopy. The production of keratin in this cell layer is responsible for their eosinophilic appearance on hematoxylin and eosin (H&E) staining.

As the keratinocytes continue to migrate superficially, they begin to flatten and develop basophilic keratohyalin granules. There are also structures called lamellar granules within these cells that contain the lipids and glycolipids that will ultimately undergo exocytosis to produce the lipid layer around the cells. It is in this layer that the keratinocytes manufacture many of the structures that will eventually serve to protect the skin and underlying tissues from environmental insult. At the superficial aspect of this layer, the keratinocytes begin to undergo programmed cell death, losing all cellular structures except for the keratin filaments and their associated proteins. In thick skin, such as that found on the palms and soles, there is a layer of flat, translucent keratinocytes called the stratum lucidum.

The final stage of the keratinocyte life cycle results in the layer of the epidermis known as the stratum corneum, or cornified layer. The protein-rich, flattened keratinocytes are now anucleate and surrounded by a lipid-rich matrix. Together the cells and surrounding matrix of this layer serve to protect the tissue from mechanical, chemical, and bacterial disruption while preventing insensible water losses through the skin.

Langerhans Cells. Of the cells in the epidermis, 3% to 6% are immune cells known as Langerhans cells. Typically found within the stratum spinosum, these mobile, dendritic cells interdigitate between keratinocytes of the epidermis to create a dense network, sampling any antigens that attempt to pass through the cutaneous tissue. Through use of their characteristic rod- or racket-shaped Birbeck granules, they take up antigens for presentation to T-cells. These monocyte-derived cells represent a large part of the skin’s adaptive immunity. Because of the effectiveness of their antigen presentation, Langerhans cells could be utilized as vaccine vehicles in the future. The Langerhans cells are functionally impaired by UV radiation, specifically UVB radiation, and may play a role in the development of cutaneous malignancies after UV radiation exposure.

Melanocytes. Within the stratum basale are melanocytes, the cells responsible for production of the pigment melanin in the skin. These neural crest-derived cells are present in a density of four to ten melanocytes per melanocyte, and about 500 to 2000 melanocytes per mm² of cutaneous tissue. This density varies based on location in the body, but differences in skin pigmentation are based on the activity of individual melanocytes and the number of melanocytes. In darker-skinned ethnicities, melanocytes create and store melanosomes in keratinocytes at a higher rate, but still have a pale-staining cytoplasm on light microscopy. Hemidesmosomes also attach these cells to the basement membrane, but the intercellular desmosomal connections are not present. The melanocytes interact with keratinocytes of the stratum basale and spinosum via long cytoplasmic extensions leading to invaginations in several keratinocytes. Tyrosinase is created and distributed into melanosomes, and these organelles travel along the dendritic processes to eventually become phagocytized by keratinocytes and distributed in a supranuclear orientation. This umbrella-like cap then serves to protect the nuclear material from damage by radiation; this could explain why light-skinned ethnicities are more prone to the development of cutaneous malignancies. Melanocytes express the bcl-2 protein, S100 protein, and vimentin, which are important in the pathology and histologic diagnosis of disorders of melanocytes.

Merkel Cells. Merkel cells are slow-adapting mechanoreceptors of unclear origin essential for light touch sensation. Thus, they typically aggregate among basal keratinocytes of the skin in areas where light tactile sensation is warranted, such as the digits, lips, and bases of some hair follicles. They are joined to keratinocytes in the basal layer by desmosomes and have dense neurosecretory granules containing peptides. These neurosecretory granules allow communication with the CNS via afferent, unmyelinated nerve fibers that contact the basolateral portion of the cell via expanded terminal discs. The clinical significance of Merkel cells arises in the setting of Merkel cell carcinoma, a rare, but difficult-to-treat malignancy.

Lymphocytes. Less than 1% of the cells in the epidermis are lymphocytes, and these are found primarily within the basal layer of keratinocytes. They typically express an effector memory T-cell phenotype.

Toker Cells. Toker cells are found in the epidermis of the nipple in 10% of both males and females and were first described in 1970. While distinct from Paget’s cells, immunohistochemical studies have implicated them as a possible source of Paget’s disease of the nipple.

Epidermal Appendages

Sweat Glands. Sweat glands, like other epidermal appendages, are derived from the embryologic ectoderm, but the bulk of their substance resides within the dermis. Their structure consists of a tubular-shaped exocrine gland and excretory duct. Eccrine sweat glands make up a majority of the sweat glands in the body and are extremely important to the process of thermoregulation. Solutes are released into the gland via exocytosis. They are present in greatest numbers on the palms, soles, axillae, and forehead. Collectively they produce approximately 10 L/d in an adult. These glands are the most effective means of temperature regulation in humans via evaporative heat loss.

A second type of sweat gland, known as the apocrine sweat gland, is found around the axilla, anus, areola, eyelid, and external auditory canal. The cells in this gland undergo an excretion process that involves decapitation of part of the cell. These apocrine glands are typically activated by sex hormones and thus activate around the time of puberty. The secretion from apocrine glands is initially odorless, but bacteria in the region may cause an odor to develop. Pheromone production may have been a function of the apocrine glands, but this may now be vestigial. While eccrine sweat glands are activated by the cholinergic system, apocrine glands are activated by the adrenergic system.

There is also a third type of sweat gland called apocrine. This is similar to an apocrine gland but opens directly to the skin surface and does not present until puberty. Both types of glands are surrounded by a layer of myoepithelial cells that can contract and assist in the excretion of glandular contents to the skin surface.

Pilosebaceous Units. A pilosebaceous unit is a multicomponent unit made up of a hair follicle, sebaceous gland, an erector pili muscle, and a sensory organ. These units are responsible for the production of hair and sebum and are present almost entirely
Throughout the body, sparing the palms, soles, and mucosa. They are lined by the germinal epithelium of the epidermis and thus serve as an important source of epidermal regeneration after partial-thickness injury or split-thickness skin graft. The sebaceous glands secrete sebum into the follicle and skin via a duct. The lipid-secreting glands are largely influenced by androgens and become functionally active during puberty. They are present in greatest numbers on the face and scalp.

**Nails.** The nails are keratinaceous structures overlying the distal phalanges of the fingers and toes. The nail is made of three main parts. The proximal portion of the nail, continuous with the germinal nail matrix, is the nail root. The root is an adherence point for the nail. The nail plate is the portion of the nail that lies on top of the nail bed, the shape of which is determined by the underlying phalanx. The third part of the nail is the free edge, which overlies a thickened portion of epidermis known as the hyponychium. The nail functions to protect the distal digits and augment the function of the pulp of the digits as a source of counter-pressure.

**Dermal Components**

**Architecture.** The dermis is a mesoderm-derived tissue that protects and supports the epidermis while anchoring it to the underlying subcutaneous tissue. It consists primarily of three unique components: a fibrous structure, the ground substance that surrounds those fibers, and the cell population that is supported by the dermis. In addition, the dermis houses the neurovasculature that supports the epidermis and facilitates interaction with the outward environment, as well as the epidermal appendages previously described. The dermis varies in thickness based upon body region, thinnest in the eyelids and reaching a thickness of up to 4 mm on the back, and is composed of two distinct layers, the papillary layer and the reticular layer. The papillary layer is made up of papillae that interdigitate with the rete ridges of the deep portion of the epidermis. This structure increases the surface area between the dermis and epidermis, increasing the resistance to shear forces as well as facilitating greater diffusion of nutrients across the dermal-epidermal junction. The papillary layer is characterized by a greater density of cells, and the reticular layer is almost entirely made up of a coarse network of fibers and the ground substance within which those fibers reside. They are typically spindle- or stellate-shaped and have a well-developed rough endoplasmic reticulum, typical of cells engaged in active protein production. The fibroblasts can also differentiate into myofibroblasts, cell types that harbor myofilaments of smooth muscle, actin, and desmin, which help to decrease the surface area of the wound by contraction. Because of these fundamental functions of fibroblasts, they are the workhorses of wound healing, while macrophages are the orchestrators.

**Fibers and Ground Substance.** Ninety-eight percent of the dry weight of the dermis is made up of collagen, typically 80% to 90% type I collagen and 8% to 12% type III collagen. Collagen types IV and VII are also found in much smaller quantities in the dermo-epidermal junction. The structure of the fibers varies along the depth of the dermis. At the superficial part of the dermis, in the papillary layer, the collagen bundles are arranged more loosely and are primarily made up of type III collagen. Deeper in the reticular layer of the dermis, the collagen fibrils are larger in diameter and organized into interwoven bundles surrounded by elastic fibers all within the hydrated ground substance. In a healthy adult, these dermal fibers are in a constant state of breakdown and production, dictated by the activity of matrix metalloproteases and fibroblasts, respectively. The activity of the MMPs is induced by UV radiation, thus leading to increased degradation and disorganization of the collagen fibers, resulting in wrinkling and weakening of the dermis in sun-exposed areas.

The retractile properties of skin are due in part to elastic fibers found throughout the dermis. These fibers, like the collagen fibers, are thinner and more perpendicularly oriented in the papillary dermis and become thicker and parallel in the reticular dermis. These elastic fibers are also produced by fibroblasts, but they are unique in that they can stretch to twice their original length, and return to their original configuration. The elastic fibers are also in a constant state of turnover that can be negatively impacted by the effects of UV radiation.

The fibrous network of the dermis lies within a hydrated amorphous ground substance made of a variety of proteoglycans and glycosaminoglycans, molecules that can contain up to 1000 times their weight in water. This ground substance facilitates the development of the structure of the dermis and cell migration within the dermis. It also assists in redistributing forces placed on the cutaneous tissues.

**Cells**

**Fibroblasts.** Fibroblasts, like most cells in the dermis, are found in the loose, papillary layer, and are the fundamental cells of the dermis. They are responsible for producing all dermal fibers and the ground substance within which those fibers reside. They are typically spindle- or stellate-shaped and have a well-developed rough endoplasmic reticulum, typical of cells engaging in active protein production. The fibroblasts can also differentiate into myofibroblasts, cell types that harbor myofilaments of smooth muscle, actin, and desmin, which help to decrease the surface area of the wound by contraction. Because of these fundamental functions of fibroblasts, they are the workhorses of wound healing, while macrophages are the orchestrators.

**Dermal Dendrocytes.** Dermal dendrocytes are comprised of a variety of mesenchymal dendritic cells recognizable mainly by immunohistochemistry. They are responsible for antigen uptake and processing for presentation to the immune system, as well as the orchestration of processes involved in wound healing and tissue remodeling. They are typically found in the papillary dermis around vascular structures as well as sweat glands and pilosebaceous units.

**Mast Cells.** Mast cells are effector secretory cells of the immune system that are responsible for immediate type 1 hypersensitivity reactions. When primed with IgE antibodies, encounter with a provoking antigen causes the release of histamine and cytokines, leading to vasodilation and dermatitis commonly seen in allergic reactions.

**Cutaneous Vasculature**

While the epidermis is void of any vasculature structures, the dermis has a rich supply of blood and nutrients supported by paired plexuses connected by a system of arteriovenous shunts. The superficial, subpapillary plexus is located between the papillary and reticular dermis and provides a vascular loop to every papilla of the papillary dermis. The deep dermal plexus is located at the junction of the reticular dermis and hypodermis, and it derives its blood supply from perforating arteries of larger vessels below the cutaneous tissues. The arteriovenous shunts connecting the two horizontal plexuses can divert blood flow to or away from the skin when necessary to conserve or release body heat, or to divert blood flow to vital organs when needed. Associated with the vascular loops of the dermal papillae are the blind-ended beginnings of lymphatic vessels, which serve to transport extravasated fluid and proteins from the soft tissues back into the venous circulatory system.
The skin is a highly specialized tool for interacting with our environment and, as such, carries a rich network of nervous tissue to facilitate this purpose. An afferent component made up of free nerve endings and specialized corpuscular receptors is responsible for conveying to our brain information about the environment, while numerous functions of the cutaneous tissues, such as AV-shunting, piloerection, and sweat secretion are controlled by the myelinated and unmyelinated fibers of an efferent component of the CNS.25

The hypodermis, or subcutaneous tissue, is a richly vascularized loose connective tissue that separates and attaches the dermis to the underlying muscle and fascia. It is made up primarily of pockets of lipid-laden adipocytes separated by septae that contain cellular components similar to the dermis, neurovascular structures supplying the cutaneous tissue, and the deepest parts of sweat glands.26 The hypodermis serves multiple functions—namely insulation, storage of energy, and protection from mechanical forces, allowing the skin to glide over the underlying tissues.

INFLAMMATORY CONDITIONS

Hidradenitis Suppurativa
Hidradenitis suppurativa, also known as acne inversa, is a painful skin condition typically affecting areas of the body bearing apocrine glands—typically the axillae, perineum, and the inframammary and inguinal folds. It is characterized by tender, deep nodules that can expand, coalesce, spontaneously drain, and form persistent sinus tracts in some cases leading to significant scarring and hyperkeratosis. There can be superimposed bacterial infection during episodic flares of the disease as well. In women, flares often occur premenstrually.

Hidradenitis suppurativa typically affects females (female to male ratio of 3:1), most commonly during the third decade of life and has demonstrated associations with smoking and obesity.27 While the etiology of hidradenitis is incompletely understood, it is thought to be the consequence of a genetic predisposition exacerbated by environmental factors. About one-third of affected patients endorse a family history of the disease. A specific gene locus has not been identified, but mutations in the γ-secretase gene have been linked to the disease in some familial cases.28 The histologic progression of the disease is characterized by atrophy of the sebaceous gland, followed by inflammation of the pilosebaceous unit from both the innate and adaptive immune systems, causing hyperkeratosis and eventual granuloma formation.29 Some studies have shown involvement of the IL12-IL23 pathway and TNF-α, supporting the theory that the disease is at least in part caused by an inflammatory disorder.30,31

The diagnosis of hidradenitis is clinical, and the presentation is most commonly categorized by the Hurley classification system, divided into three stages. Single or multiple nodules or abscesses without any sinus tracts or scarring would be classified as stage I disease. As abscesses recur and sinus tracts and scarring form, the disease is classified as Hurley stage 2. Stage 3 is the most advanced stage, with diffuse disease and interconnected sinus tracts and abscesses.

Treatment is typically based on Hurley staging, with topical and systemic antibiotics (typically clindamycin) being used for stage I and II disease,32 while radical excision, laser treatment, and biologic agents are reserved for more advanced stage II and III disease.33-36 Even with complete surgical resection, recurrence rates are still high, reaching up to 50% in the infra mammary and inguino-perineal regions. Because of increased risks of recurrence with primary closure, it is preferable to pursue other methods of wound closure, like split-thickness skin grafting, local or regional flaps, and healing by secondary intention. Topical antimicrobial creams should be used during the healing process.

Pyoderma Gangrenosum
Pyoderma gangrenosum is an uncommon inflammatory condition of the skin characterized by the development of sterile pustules which progress to painful, ulcerating lesions with purple borders. This disease is typically diagnosed between the ages of 40 and 60 years and has a slightly higher prevalence in females. Although the exact etiology is currently unknown, it typically arises in individuals with a hematologic malignancy or inflammatory disorder, such as inflammatory bowel disease or rheumatoid arthritis. The most commonly affected sites are the legs, but lesions can occur anywhere. Extracutaneous manifestations are also possible, and it can affect mucosal tissue and solid organs. While the initial pathology is sterile, it can easily become secondarily infected. The diagnosis of this condition is based upon history and clinical presentation after the exclusion of infectious etiologies. There are five distinct types of pyoderma gangrenosum described: vegetative, pustular, peritoma, ulcerative, and bullous. The pathogenesis of this disease is incompletely understood, but it is thought to be a genetic predisposition that is triggered by an environmental influence. An inciting cutaneous injury can often be identified preceding the ulceration. Histopathologic studies have demonstrated significantly elevated levels of inflammatory cytokines, as well as neutrophils exhibiting aberrant chemotactic signaling.37-39 Treatment of pyoderma gangrenosum generally involves treatment of the underlying disorder (i.e., management of Crohn’s disease) as well as systemic anti-inflammatory medications such as steroids or immunosuppressants like calcineurin inhibitors. Patients with Crohn’s disease and PG treated with infliximab (tumor necrosis factor [TNF]-α inhibitor) and etanercept (TNF-α antagonist) had a marked improvement in their PG.40,41 In cases of peristomal pyoderma gangrenosum, topical calcineurin inhibitors have been shown to be useful.42 Concurrent treatment with systemic and topical antimicrobials, as well as local wound care, including the debridement of purulent exudate and devitalized tissue, is also beneficial. Surgical therapy without proper systemic treatment will generally result in recurrent disease. Final wound closure can be achieved with primary closure or grafts.

Epidermal Necrolysis
Epidermal necrolysis (EN) is a rare mucocutaneous disorder characterized by cutaneous destruction at the dermoepidermal junction. EN is commonly referred to as either Stevens-Johnson syndrome (SJS) or toxic epidermal necrolysis (TEN) depending on the extent of skin involvement present. SJS refers to cases in which <10% of total body surface area is involved, while cases with >30% involvement are considered TEN, with an SJS-TEN overlap syndrome referring to all cases in between. These two disorders are now considered to be the same clinical entity that vary simply on the extent of cutaneous involvement. Erythema multiforme was once considered as part of the clinical subgroup...
encompassing SJS and TEN, but it is now thought to be a separate entity related to herpetic or *Mycoplasma* infections.

The clinical presentation usually occurs within 8 weeks of initiation of a new drug treatment and is characterized by a macular rash beginning in the face and trunk and progressing to the extremities within hours to days. A positive Nikolsky sign is often present, in which lateral pressure on the skin causes separation of the epidermis from the dermis. The macular rashes then begin to blister and coalesce, forming bullae that eventually burst, leaving partial thickness wounds with exposed dermis. Mucous membrane involvement is seen in 90% of cases and can involve the oral, genital, and ocular mucosa, as well as the respiratory and gastrointestinal tracts. The cutaneous manifestations can also be associated with high fever and pain. It is important to distinguish EN from infectious etiologies like staphylococcal scalded skin syndrome due to their similar clinical presentation.

While the etiology is not entirely clear, it is well documented to be a reaction to various drugs. While over 100 drugs have been implicated as the inciting agent of EN, there are a handful of high-risk drugs that account for a majority of the cases. The drugs most commonly associated with EN include aromatic anticonvulsants, sulfonamides, allopurinol, oxycams (nonsteroidal anti-inflammatory drugs), and nevirapine. The pathophysiology is also incompletely understood, but it has generally been accepted that it involves cell-mediated cytotoxicity targeted at keratinocytes and the cytokine-induced expression of “death-receptors” like Fas-L. Recently, studies have demonstrated greatly increased concentrations of granulysin, an apoptotic protein secreted by cytotoxic T cells, within EN lesions, and thus this protein may be implicated in the pathogenesis of EN. A genetic component may also exist, and genetic testing before carbamazepine treatment is recommended in people of Han Chinese ancestry to exclude carriers of HLA-B1502.

The prognosis of EN is generally related to the surface area affected and secondary complications of extensive cutaneous damage, like secondary infections and loss of hemodynamic stability due to increased insensible losses and third spacing of fluid. Modern burn- and ICU-care has decreased mortality significantly. The first principle of management of EN is discontinuation of the offending agent, and in drugs with short half-lives, this can significantly increase chances of survival. Other management principles include maintenance of euvoolemia, early enteral feeding, and measures to reduce risk of infection. This includes surgical debridement of devitalized tissue, the use of topical antibiotics or antimicrobial dressings, nonadherent dressings, or temporary biologic or synthetic grafts until the underlying dermis can reepithelialize. The cornea should regularly be inspected with a Wood’s lamp to evaluate for corneal sloughing.

The use of systemic corticosteroids in the acute setting is controversial as there have been mixed results. Some studies have shown a slowed disease progression when corticosteroid therapy was administered early, while others showed increased rates of sepsis and overall mortality with no effect on disease progression. IVIG has also been used in an effort to inhibit the Fas-L cytotoxic pathway, with some mixed results. A 2007 meta-analysis of nine IVIG trials concluded that high-dose IVIG improves survival, while a large retrospective analysis in 2013 concluded that there was no mortality benefit. Other agents, like cyclosporine A, plasmapheresis and anti-TNF-α have been studied with mixed results. Recent guidelines out of the United Kingdom confirm that there is still no treatment with clearly demonstrated benefit in the management of EN. The cutaneous manifestations of EN generally progress for 7 to 10 days, while reepithelialization generally occurs over 3 weeks.

**INJURIES**

**Radiation-Induced Injuries**

Radiation injuries can result from exposure to electromagnetic radiation from industrial/occupation applications or, more commonly, from environmental exposure and medical treatments. This is especially true in the continually evolving role of radiation therapy in the multidisciplinary approach to oncologic disease and other skin conditions. In addition to treatment for lymphomas, head and neck squamous cell carcinomas, and prostate adenocarcinoma, it is often an adjuvant or neoadjuvant component of the surgical treatment of rectal, breast, esophageal, and cervical cancers. Although the new modalities and principles of radiation therapy have allowed for more precise administration of this therapy, there is still collateral damage in the cutaneous and visceral tissues surrounding the treatment site.

Environmental sources of radiation damage are typically from UV radiation. UVC rays are filtered by the ozone layer, so the only UV rays that humans typically encounter are UVA (320–400 nm) and UVB (290–320 nm). The amount of exposure to UV radiation is dependent on seasonal, temporal, geographic and environmental variables. Ninety-five percent of the UV rays that reach the earth’s surface are UVA rays. This radiation is less energetic (longer wavelength) than UVB rays and affects the cutaneous tissues differently. UVA waves penetrate deeper into the tissues, with 20% to 30% reaching the deep dermis. UVB rays are mostly absorbed in the epidermis, with 70% reaching the stratum corneum, 20% reaching the deep epidermis, and only 10% reaching the papillary dermis. Major chromophores in the cutaneous tissue include nucleic acids, aromatic amino acids, and melanin.

The short-term effects of solar radiation include erythema and pigmentation. The resultant erythema peaks at 6 to 24 hours...
after exposure. The pigmentation occurs differently for UVA and UVB rays. Partial fading of this pigment change occurs within an hour after exposure, but with higher and repeated doses of UVA, stable residual pigmentation is observed. UVB waves induce neomelanization, increasing the total amount of melanin in the epidermal tissues and resulting in an effect that is observable 72 hours after exposure. The increase in melanin as a result of UVB exposure serves as a protective mechanism to defend the nuclei of the basal keratinocytes from further radiation-induced damage by absorbing the high-energy radiation in future exposures. Long-term effects of exposure to UV radiation can lead to chronic skin changes, such as irregular pigmentation, melasma, postinflammatory pigmentation, and actinic lentigines (sun spots). Lysozyme, an enzyme secreted by cells of the immune system, typically inhibits the activity of collagenase and elastase, playing a role in turnover of the elastin and collagen network of the dermis. Long-term exposure to UV radiation increases the activity of lysozyme, thus impairing the natural turnover of these fibers, resulting in a disorganized accumulation of elastin, and an increase in the ratio of type III to type I collagen. This results in loss of firmness and resilience of the skin, leading to wrinkles and an aged appearance.

The other major source of radiation injury that a surgeon will likely encounter is from therapeutic radiation. The various forms of radiation work to destroy the replicative potential of the target cells via damage to the nucleic acid structures in the cell. This is typically used to treat oncologic disease, but it can also be used to treat benign disease like eczema, psoriasis, and keloid scarring at relatively low exposures. While this goal is accomplished, surrounding tissues are also affected and damaged. The most radiosensitive components of the cutaneous tissue are the basal keratinocytes, hair follicle stem cells, and melanocytes. Exposure to this intense radiation results in disorganized, uncontrolled cell death, leading to the release of reactive oxygen species and further damage and inflammation to the surrounding cellular network. Damage to the basal keratinocytes and fibroblasts hinders the replicative capacity of the epidermis and dermis, respectively.

Acute skin changes to these structures manifest within weeks as erythema, edema, and alopecia. Permanent hyperpigmentation, tightening, thickening, and fibrosis of the skin become apparent as the tissue attempts to heal. In severe radiation injury, there can be complete loss of the epidermis, resulting in partial-thickness wounds and fibrinous exudate. Reepithelialization typically occurs 14 days following initial injury, provided other variables affecting wound healing are optimized (bacterial colonization, nutrition.) Long-term effects include compromise of the functional integrity of the skin secondary to thrombosis and necrosis of capillaries, hypovascularity, telangiectasia, ulceration, fibrosis, poor wound healing, and infection. These can present weeks to years after exposure.

Treatment of minor radiation injury includes skin moisturizers and local wound care when appropriate. Severe radiation injury may warrant surgical excision and reconstruction with free-tissue transfer from a part of the body unaffected by radiation.

**Trauma-Induced Injuries**

**Mechanical Injury.** Physical disruption of the skin can occur via numerous mechanisms. Treatment of the wound is dependent on the size of the defect left behind by the insult, any exposed structures that remain in the wound bed, and the presence of contaminating debris or infection. Clean, simple lacerations can be irrigated, debrided, and closed primarily. There is no systematic evidence to guide the optimal timing of closure within 24 hours, but many surgeons will close primarily within 6 hours of injury. Grossly contaminated or infected wounds should be allowed to heal by secondary intention or delayed primary closure. In wounds allowed to heal secondarily, negative pressure wound therapy can increase the rate of granulation tissue formation. Tangential abrasions are treated similarly to burn wounds, with depth of injury dictating management. Partial thickness injuries with preservation of the regenerative pilosebaceous units can be allowed to heal on their own while maintaining a moist, antimicrobial wound environment. Full thickness wounds may require reconstruction with split- or full-thickness skin grafting depending on the size of the defect and the need for future cosmesis and durability. In the setting of devitalization of full thickness tissue, the damaged tissue may be used as a full thickness graft, provided the wound is appropriately cleaned.

**Bite Wounds.** Dog bites alone recently accounted for 4.5 million bites to humans in a single year. Bites from dogs, humans, and other animals can quickly lead to severe deep-tissue infections if not properly recognized and treated. The most common location of bite wounds is the hand. This area is of particular importance, as the anatomy of the hand allows for rapid progression of deep infection long relatively avascular structures and can lead to long term morbidity if not adequately treated. Bite bacteriology is influenced by normal mouth flora, as well as the content of the offending animal’s diet. Early presentation bite wounds yield polymicrobial cultures, while cultures from a late infection will typically exhibit one dominant pathogen. Common aerobic bacteria include *Pasteurella multocida*, *Streptococcus*, *Staphylococcus*, *Neisseria*, and *Corynebacterium*; anaerobic organisms include *Fusobacterium*, *Porphyromonas*, *Prevotella*, *Propionibacterium*, *Bacteroides*, and *Peptostreptococcus*. *Capnocytophaga canimorsus* bacteria after a dog bite are rare, and it appears that immunocompromised patients are most susceptible to this type of infection and its complications. The bacterial load in dog bites is heavily influenced by the last meal of the animal, increasing with wet food and shorter time since the last meal (Fig. 16-3). Cat bite bacteriology is similar, with slightly higher prevalence of *Pasteurella* species. Infections from *Francisella tularensis* (tularemia) and *Yersinia pestis* (human plague) have been reported.

Bacteria colonizing human bites are those present on the skin or in the mouth. These include the gram-positive aerobic organisms *Staphylococcus aureus*, *Staphylococcus epidermidis*, and *Streptococcus* species, and anaerobes including *Peptococcus* species, *Peptostreptococcus* species, *Bacteroides* species, and *Eikenella corrodens* (facultative anaerobe). Human bites are characterized by a higher bacterial load (>10⁶). Antibiotic prophylaxis after a human bite is recommended as it has been shown to significantly decrease the rate of infection. A course of 3 to 7 days of amoxicillin/clavulanate is typically used. Alternatives are doxycycline or clindamycin with ciprofloxacin.

There is controversy over the closure of bite wounds. Typically, in areas of aesthetic importance, the wound is thoroughly irrigated and debrided and primarily closed with a short course of antibiotics and close follow-up to monitor for signs of infection. In areas that are less cosmetically sensitive and bites that look grossly contaminated or infected, the wounds
Figure 16-3. A. Dog bite to the face involving the lip. B. Primary multilayer closure following debridement and irrigation. Closure was performed due to aesthetic and functional considerations. C. Follow up 1 week after injury following suture removal.

are allowed to close secondarily. Special consideration should be paid to puncture wounds in areas like the hands, which have multiple small compartments. Some groups have found that as long as wounds are properly irrigated and cleansed with povidone iodine solution while a short course of antibiotics is prescribed, there is no difference in infection rates in dog bite wounds closed primarily.62

Rabies in domestic animals in the United States is rare, and most cases are contracted from bat bites. In developing countries, dog bites remain the most common source of rabies. Management of this is beyond the scope of this chapter.

Caustic Injury
Chemical burns make up to 10.7% of all burns but account for up to 30% of all burn-related deaths.63 The number of cases of industrial chemical burns is declining while chemical burns in the domestic setting is on the rise. The extent of tissue destruction from a chemical burn is dependent on type of chemical agent, concentration, volume, and time of exposure, among other variables.

Injuries from acidic solutions are typically not as severe as those from basic solutions. This is due to the mechanism of injury of each. Acidic injuries typically result in superficial eschar formation because the coagulative necrosis caused by acids limits tissue penetration. Acids can cause thermal injury in addition to the coagulative necrosis due to exothermic reactions. Without treatment, acid injuries will progress to erythema and ulcers through the subcutaneous tissue. Injuries from basic solutions undergo liquefactive necrosis, unlike acids, and thus have no barrier preventing them from causing deeper tissue injury.
Common examples of agents that often cause alkaline chemical burns are sodium hydroxide (drain decloggers and paint removers) and calcium hydroxide (cement).

Treatment for acidic or alkaline chemical burns is first and foremost centered around dilution of the offending agent, typically using distilled water or saline for 30 minutes for acidic burns and 2 hours for alkaline injuries. Attempting to neutralize the offending agent is typically discouraged, as it does not offer an advantage over dilution and the neutralization reaction could be exothermic, increasing the amount of tissue damage. After removal of the caustic agent, the burn is treated like other burns and is based on the depth of tissue injury. Topical antimicrobials and nonadherent dressings are used for partial-thickness wounds with surgical debridement and reconstruction if needed for full-thickness injuries. Liposuction and saline dilution have been used in cases were injury to deeper structures was suspected. Prophylactic use of antibiotics is generally avoided.

There are several chemical agents that have specific treatments, including the use of calcium gluconate for hydrofluoric acid burns and polyethylene glycol for phenol burns. These types of treatments are specific to the offending agent and outside of the scope of this chapter.

One type of caustic injury that is commonly seen in the hospital is extravasation injury, especially in the setting of chemotherapy administration. Extravasation is estimated to occur in 0.1% to 0.7% of all cytotoxic drug administrations. Like other chemical burns, extravasation injuries depend on properties of the offending agent, time of exposure, concentration, and volume of drug delivered to the tissues. Extravasation injuries typically cause little damage, but they can cause significant morbidity in those with thin skin, fragile veins, and poor tissue perfusion, like neonates and the critically ill. Extravasation injuries usually involve swelling, pain, erythema, and blistering. It may take days or longer for the extent of tissue damage to demarcate. Thorough evaluation to rule out injury to deeper tissues should be conducted. The treatment for extravasation injuries is usually conservative management with limb elevation, but saline aspiration with a liposuction cannula in an effort to dilute and remove the offending agent has been used soon after injury presentation. Infiltration of specific antidotes directed toward the offending agent has been described, but it lacks the support of randomized controlled trials, and no consensus in treatment has been reached. It is best to avoid cold or warm compression because the impaired temperature regulation of the damaged tissue may lead to thermal injury. After the wound demarcates, full-thickness skin death should be surgically debrided and managed like other wounds based on depth of injury.

**Thermal Injury**

Thermal injury involves the damage or destruction of the soft tissue due to extremes of temperature, and the extent of injury is dependent on the degree temperature to which the tissue is exposed and the duration of exposure. The pathophysiology and management are discussed in detail in a separate chapter. Briefly, the management of thermal wounds is initially guided by the concept of three distinct zones of injury. The focus of thermal injury that has already undergone necrosis is known as the zone of coagulation. Well outside the zone of coagulation is the zone of hyperemia, which exhibits signs of inflammation but
will likely remain viable. In between these two zones is a zone of stasis with questionable tissue viability, and it is this area at which proper burn care can salvage viable tissue and decrease the extent of injury\(^6\) (Fig. 16-6).

The mechanisms of injury in hypothermic situation differ. Direct cellular damage can occur as a result of the crystallization of intracellular and extracellular components with resultant dehydration of the cell and disruption of lipid protein

---

**Figure 16-5.** A. Potassium chloride intravenous infiltrate in a critically ill patient on multiple vasopressors. B. Following operative debridement to paratenon layer. C. Temporary coverage with Integra skin substitute.
complexes. During rewarming, further damage occurs because of the shifts of fluid in response to melting ice. Indirect effects of hypothermic injury include microvascular thrombosis and tissue ischemia. This, together with subsequent edema and inflammation upon rewarming, propagates tissue injury even further. Even so, the standard treatment of frostbite injury begins with rapid rewarming to 40°C to 42°C. In addition, further treatment includes debridement of all devitalized tissue, hydrotherapy, elevation, topical antimicrobials, topical antithromboxanes (aloe vera), and systemic antiprostaglandins (aspirin).

**Pressure Injury**

A problem that all surgeons will encounter very early in their careers is pressure necrosis. The development of pressure ulcers is increasingly being regarded as a marker of quality of care, and strategies aimed at prevention have been the source of recent study. Pressure ulcers are known to affect the critically ill (22% to 49% of all critically ill patients are affected), but pressure sources can also affect the chronically bed- or wheelchair-bound, patients undergoing surgical procedures, and those with Foley catheters, artificial airways, or other medical equipment (Fig. 16-7).

Pressure ulcers can present in several ways depending on the stage at presentation. They are typically grouped into 4 stages: stage 1, nonblanching erythema over intact skin; stage 2, partial-thickness injury with blistering or exposed dermis; stage 3, full-thickness injury extending down to, but not including, fascia and without undermining of adjacent tissue; and stage 4, full-thickness skin injury with destruction or necrosis of muscle, bone, tendon, or joint capsule. Tissue destruction occurs most easily at bony prominences due to the inability to redistribute forces along a greater surface area. The average perfusion pressure of the microcirculation is about 30 mmHg, and pressures greater than that cause local tissue ischemia. In animal models, pressure greater than twice the capillary perfusion pressure produces irreversible tissue necrosis in just 2 hours. The most common areas affected are the ischial tuberosity (28%), greater trochanter (19%), sacrum (17%), and heel (9%). Tissue pressures can measure up to 300 mmHg in the ischial region during sitting and 150 mmHg over the sacrum while lying supine. Tissues with a higher metabolic demand are
typically susceptible to insult from tissue hypoperfusion more rapidly than tissues with a lower metabolic demand. Because of this, it is possible to have muscle necrosis beneath cutaneous tissue that has yet to develop signs of irreversible damage.

Management of pressure sores first and foremost involves avoidance of prolonged pressure to at-risk areas. Strategies typically employed are pressure-offloading hospital beds or assist devices, patient repositioning every 2 hours, early mobilization, prophylactic silicone dressings, and nursing education. From a wound healing perspective, patients should be nutritionally optimized and surgically debrided as appropriate. The presence of stage III or IV pressure ulcers is not necessarily an indication for surgery, and fevers in a patient with chronic pressure ulcers are often from a urinary or pulmonary source. Goals of surgical intervention are drainage of fluid collections, wide debridement of devitalized and scarred tissue, excision of pseudobursa, ostectomy of involved bones, hemostasis, and tension-free closure of dead space with well-vascularized tissue (muscle, musculocutaneous, or fasciocutaneous flaps). Stage 2 and 3 ulcers may be left to heal secondarily after debridement. Subatmospheric pressure wound therapy devices (vacuum-assisted closure) play a role in wound management by removing excess interstitial fluid, promoting capillary circulation, decreasing bacterial colonization, increasing vascularity and granulation tissue formation, and contributing to wound size reduction.

**BIOENGINEERED SKIN SUBSTITUTES**

The management of soft tissue defects is more commonly including the use of bioengineered skin substitutes. These products are typically derived from or designed to imitate dermal tissue, providing a regenerative matrix or stimulating autogenous dermal regeneration while protecting the underlying soft tissue and structures. There are generally four types of skin substitutes: (a) autografts, which are taken from the patient and placed over a soft tissue defect (split-thickness and full-thickness skin grafts); (b) allografts, which are taken from human organ donors; (c) xenografts, which are taken from members of other animal species; and (d) synthetic and semisynthetic biomaterials that are constructed de novo and may be combined with biologic materials. Acellular dermal matrices are one type of skin substitute and are used quite often for wound healing and support of soft tissue reconstruction. They are from allogenic or xenogeneic sources and are composed of collagen, elastin, laminin, and glycosaminoglycans. Tissue incorporation generally occurs within 1 to 2 weeks. Dermal matrices have been shown to be an effective bridge to split-thickness skin grafting for wounds that have exposed nerves, vessels, tendons, bones, or cartilage. Bilayered matrices can also be used to promote dermal regeneration in acute or chronic wounds. These products can be temporary, needing to be removed prior to grafting, or permanent, integrating into the host tissue and being grafted directly.

**BACTERIAL INFECTIONS OF THE SKIN AND SUBCUTANEOUS TISSUE**

**Introduction**

In 1998, the Food and Drug Administration (FDA) categorized infections of the skin and skin structures for the purpose of clinical trials. A revision of this categorization in 2010 excluded specific diagnoses such as bite wounds, decubitus ulcers, diabetic foot ulcers, perirectal abscesses, and necrotizing fasciitis. The general division into “uncomplicated” and “complicated” skin infections can be applied to help guide management.

The agent most commonly responsible for skin and soft tissue infections is *S. aureus* and is isolated in 44% of specimens. Less common isolates include other gram-positive bacteria such as *Enterococcus* species (9%), β-hemolytic streptococci (4%), and coagulase-negative staphylococci (3%). *S. aureus* is more commonly responsible for causing abscesses. Patients with an impaired immune system (diabetic, cirrhotic, or neutropenic patients) are at higher risk of infection from gram-negative species like *Pseudomonas aeruginosa* (11%), *Escherichia coli* (7.2%), *Enterobacter* (5%), *Klebsiella* (4%), and *Serratia* (2%), among others.

**Uncomplicated Skin Infections**

Uncomplicated infections involve relatively small surface area (<75 cm²) and bacterial invasion limited to the skin and its appendages. Impetigo, erysipelas, cellulitis, folliculitis, and simple abscess fall into this category. Impetigo is a superficial infection, typically of the face, that occurs most frequently in infants or children, resulting in honey-colored crusting. Erysipelas is a cutaneous infection localized to the upper layers of the dermis, while cellulitis is a deeper infection, affecting the deeper dermis and subcutaneous tissue. Folliculitis describes inflammation of the hair follicle, and a furuncle describes a follicle with swelling and a collection of purulent material. These lesions can sometimes coalesce into a carbuncle, an abscess with multiple different draining sinuses.

It is recommended to culture infectious lesions to help identify the causative agent, but treatment without these studies is reasonable in typical cases. Minor infections can be safely treated with topical antimicrobials like 2% mupirocin to provide coverage for methicillin-resistant *S. aureus* (MRSA). Folliculitis generally resolves with adequate hygiene and warm soaks. Furuncles, carbuncles and other simple abscesses can be incised, drained, and packed, typically without the use of systemic antibiotics. The decision to use systemic antibiotics after incision and drainage of abscess should be made based upon presence or absence of systemic inflammatory response syndrome (SIRS) criteria.

For nonpurulent, uncomplicated cellulitis in which there is no drainable collection, systemic antibiotic coverage for β-hemolytic streptococcus is recommended. If there is no improvement in 48 to 72 hours or worsening of symptoms, antibiotic coverage should be added for MRSA. Systemic therapy for purulent cellulitis, which includes cutaneous abscesses, should cover MRSA, and empiric coverage for streptococcus is likely unnecessary. Antibiotic coverage for streptococcus is generally accomplished with β-lactam antibiotics like penicillins or first-generation cephalosporins. MRSA coverage is accomplished with clindamycin, trimethoprim-sulfamethoxazole, linezolid, and tetracyclines. Clindamycin, trimethoprim-sulfamethoxazole, linezolid, or tetracycline combined with a β-lactam can all be used for dual coverage of streptococcus and MRSA.

**Complicated Skin Infections**

Complicated skin infections include superficial cellulitis encompassing a large surface area (>75 cm²) or deeper infections extending below the dermis. Necrotizing soft tissue infections (NSTIs), including necrotizing fasciitis, can rapidly cause extensive morbidity and mortality, thus their prompt diagnosis and appropriate management is crucial. A thorough history and
exam should be performed to elicit information (e.g., history of trauma, diabetes mellitus, cirrhosis, neutropenia, bites, IV or subcutaneous drug abuse) as well as physical findings such as crepitus (gas-forming organism), fluctuance (abscess), purpura (sepsis in streptococcal infections), bullae (streptococci, *Vibrio vulnificus*), lymphangitis, and signs of a systemic inflammatory response.

Extensive cellulitis is managed in a similar fashion as simple cellulitis. Initial treatment consists of intravenous antibiotics that cover β-hemolytic streptococcus, such as cephalosporins, with the addition of MRSA coverage if there is no improvement in symptoms. Vancomycin is typically the first choice for MRSA coverage, but this drug is inferior to β-lactams for coverage of MSSA. Alternative antibiotics that are typically effective against MRSA are linezolid, daptomycin, tigecycline, and telavancin. Clindamycin is approved for use against MRSA, but resistance rates are increasing, and its use is discouraged if institutional rates of clindamycin resistance are >15%.

Necrotizing soft tissue infections occur 500 to 1500 times a year in the United States and are frequently associated with diabetes mellitus, intravenous drug abuse, obesity, alcohol abuse, immune suppression, and malnutrition. Because NSTIs can often present initially with nonspecific findings, the physician should always have a high index of suspicion when evaluating a patient. The threshold for surgical exploration and debridement should be low, particularly in a weakened host. Occasionally an inciting event or point of entry can be identified, but in 20% to 50% of cases, the exact cause is unknown. These infections are associated with a high mortality, ranging from 25% to 40%, with higher rates in the truncal and perineal cases.

NSTIs are classified based on anatomic site, involved tissues, and the offending organisms. NSTIs commonly originate at the genitalia, perineum (Fournier’s gangrene), and abdominal wall. Subcutaneous tissue, fascia and muscle can all be affected. Necrotizing fasciitis involves infection of the fascia, and the infection can quickly travel along the easily separable, avascular planes. There are three types of NSTIs when classified by the offending agent. The most common is type 1, which is caused by a polymicrobial source including gram-positive cocci, gram-negative rods, and anaerobic bacteria, specifically *Clostridium perfringens* and *C. septicum*. Type 2 is caused by a monomicrobial source of β-hemolytic *Streptococcus* or *Staphylococcus* species, with MRSA contributing to the increasing number of community-acquired NSTIs. A history of trauma is often elicited and can be associated with toxic shock syndrome. Type 3 is a rare but fulminant subset resulting from a *V. vulnificus* infection of traumatized skin exposed to a body of salt-water.

In addition to signs of SIRS, patients can present with skin changes like erythema, bullae, necrosis, pain, and crepitis. They may exhibit signs of hemodynamic instability, and gas within the soft tissues on imaging is pathognomonic. Patients can present with a range of symptoms, from minimal skin change to frank necrosis, and the time of progression to fulminant disease varies in each patient. Laboratory values are nonspecific and resemble values seen in sepsis. There have been attempts at creating scoring systems to assist in the diagnosis of NSTI. One study in 2000 used the criteria of a white blood cell count >15,400 and a serum sodium level <135 mmol/L. This test was found to have a negative predictive value of 99%, but a positive predictive value of only 26%. In 2004, six criteria were used and referred to as the Laboratory Risk Indicator for Necrotizing Fasciitis, or LRINEC, and included C-reactive protein (CRP), white blood cell (WBC) count, hemoglobin, plasma sodium, creatinine, and glucose. A score of 8 or greater
suggested a high probability of NSTI, 6 or 7 an intermediate probability, and <5 a low probability. This test was internally validated and found to have a PPV of 92% and an NPV of 96%. However, some have criticized this study because of its small sample size and over-reliance on CRP, which can be elevated in multiple other conditions. Blood cultures are not always positive, and tissue samples will demonstrate necrosis, white blood cell infiltration, thrombosis, angiitis, and microorganisms. The use of cross-sectional imaging in the diagnosis of NSTI is limited, and it should not delay appropriate surgical treatment.

Three principles form the foundation of the management of NSTIs: (a) source control with wide surgical debridement, (b) broad-spectrum intravenous antibiotics, and (c) supportive care and resuscitation. As soon as the diagnosis is clear or the suspicion is high, the patient should be taken for operative exploration and debridement. Incisions should be made parallel to neurovascular structures and through the fascial plane, removing any purulent or devitalized tissue until viable, bleeding tissue is encountered. On inspection, the tissue will appear necrotic with dead muscle, thrombosed vessels, the classic “dishwater” fluid, and a positive finger test, in which the tissue layers can be easily separated from one another. In Fourmier’s gangrene, one should aim to preserve the anal sphincter as well as the testicles (blood supply is independent of the overlying tissue and is usually not infected). Return to the OR should be planned for the next 24 to 48 hours to verify source control and the extent of damage. Broad spectrum antibiotic therapy should be initiated as soon as possible, with the intent of covering gram positives (including MRSA), gram negatives, and anaerobic organisms. The Infectious Diseases Society of America recommends initiating therapy with intravenous vancomycin and piperacillin/tazobactam, unless a monomicrobial agent is identified, in which case more directed therapy would be appropriate.81 Antibiotic therapy should continue until the patient requires no further debridement, is clinically improving, and has been afebrile for 48 to 72 hours.

Adjuncts to surgery include topical antimicrobial creams, subatmospheric pressure wound dressings, and optimization of nutrition. Controversial topics include the role of hyperbaric oxygen82 (may inhibit infection by creating an oxidative burst, with anecdotally fewer debridements required and improved survival, but limited availability) and IVIG (may modulate the immune response to streptococcal superantigens). Wound closure is performed once bacteriologic, metabolic, and nutritional balances are obtained.

**Actinomycosis**

*Actinomyces* is a genus of gram positive rods that inhabit the oropharynx, gastrointestinal tract, and female genital tract. The most commonly isolated species causing disease in humans is *A. israelii*. The cervicofacial form of *Actinomyces* infection is the most common presentation, representing 55% of cases, and typically presenting as an acute pyogenic infection in the submandibular or paramandibular area. Patients can also exhibit chronic soft tissue swelling, fibrosis, and sinus discharge of sulfur granules.86 Demonstration of gram-positive filamentous organisms and sulfur granules on histological examination is strongly supportive of a diagnosis of actinomycosis.86 These infections are typically treated with high doses of intravenous followed by oral penicillin therapy. Surgical treatment is utilized if there is extensive necrotic tissue, poor response to antibiotics, or the need for tissue biopsy to rule out malignancy.

**VIRAL INFECTIONS WITH SURGICAL IMPLICATIONS**

**Human Papillomavirus Infections**

Human papillomaviruses represent a group of over 100 isolated types of small DNA viruses of the *Papovavirus* family that is highly host-specific to humans.83 These viruses are transmitted via cutaneous contact with individuals who have clinical or subclinical infection and occur more frequently in immunocompromised individuals. The viruses are responsible for the development of verrucae, or warts. These are histologically characterized by nonspecific findings of hyperkeratosis, papillomatosis, and acanthosis, as well as the hallmark koilocytes (clear halo around nucleus). Clinically, these generally arise as slow-growing papules on the skin or mucosal surfaces. Regression of HPV lesions is frequently an immune-mediated, spontaneous event that is exemplified by the persistent and extensive manifestation of this virus in the immune-compromised patient.

The subtypes are generally grouped, based on their presentation, as cutaneous or mucosal. Cutaneous types most commonly affect the hands and fingers. Verruca vulgaris, or common warts, are caused by HPV types 1, 2, and 4, with a prevalence of up to 33% in school children and 3.5% in adults, and a higher prevalence in the immunosuppressed population.81 Plantar and palmar warts (HPV-1 and -4) typically occur at points of pressure and are characterized by a keratotic plug surrounded by a hyperkeratotic ring with black dots (thrombosed capillaries) on the surface. Plane warts occur on the face, dorsum of hands, and shins. They are caused by HPV-3 and -10 and tend to be multiple, flat-topped lesions with a smooth surface and light brown color. Cutaneous warts typically regress spontaneously in the immunocompetent patient. Erythrodysplasia verruciformis is a rare, autosomal recessive inherited genetic skin disorder that confers increased susceptibility to certain types of HPV. This presents with difficult-to-treat and often widespread verrucae that carry a higher risk of malignant transformation (30%-50% risk of squamous cell carcinoma), especially when caused by HPV types 5 and 8.82 A similar clinical picture has been described in human immunodeficiency virus (HIV) and transplant patients.83,84

Mucosal HPV types cause lesions in the mucosal or genital areas and behave like sexually transmitted infections. The most common mucosal types are HPV-6, -11, -16, -18, -31, and -33. These lesions present as condylomata acuminata, genital or veneral warts, papules that occur on the perineum, external genitalia, anus, and can extend into the mucosal surfaces of the vagina, urethra and rectum. These lesions are at risk for malignant transformation, with types 6 and 11 conferring low risk, and types 16, 18, 31 and 33 conferring a high risk. The recently developed quadrivalent HPV vaccine, targeting HPV types -6, -11, -16, and -18, is now available to both males and females age 9 to 26 and is associated with an up to 90% reduction of infections from those HPV types.85

Treatment is aimed at physical destruction of the affected cells. Children often require no treatment as spontaneous regression is common. In cases causing physical or emotional discomfort, or in cases of immunocompromise or risk of transmission, treatment may be indicated. Cryotherapy using liquid nitrogen is an effective treatment for most warts, but care must be taken not to damage underlying structures.86 Topical preparations of salicylic acid, silver nitrate, and glutaraldehyde may also be
used. Treatment of recalcitrant lesions includes a variety of therapeutic options aimed at physically destroying the lesions by electrodessication, cryoablation, and pulsed dye laser therapy. Additional modalities such as H2-antagonists and zinc sulfate may have a role in augmenting the immune response and reducing recurrence rates.

**Cutaneous Manifestations of Human Immunodeficiency Virus**

The HIV-infected patient is significantly more susceptible to infectious and inflammatory skin conditions than the rest of the population. These skin disorders may be due to the HIV infection itself or from opportunistic infections secondary to immunosuppression. During early stages, nonspecific cutaneous manifestations may occur. Acute retroviral syndrome occurs following inoculation in one-half to two-thirds of patients, and 30% to 50% of these patients can present with an acute viral exanthem. This is usually a morbilliform rash affecting the face, trunk, and upper extremities. Other skin changes, as well as common skin disorders with atypical features, can occur, including recurrent varicella zoster, hyperkeratotic warts, and seborrheic dermatitis. Condylomata acuminata and verrucae appear early; however, their frequency and severity do not change with disease progression.

Late-presenting cutaneous manifestations include chronic herpes simplex virus (HSV), cytomegalovirus, and, to a lesser extent, molluscum contagiosum, which is typically treatable with imiquimod. HSV is the most common viral infection in the patient with HIV, and is more likely to display atypical features and less likely to spontaneously resolve in these patients. Mycobacterial infections and mucocutaneous candidiasis also occur. Bacterial infections such as impetigo and folliculitis may be more persistent and widespread.

Malignant lesions such as Kaposi’s sarcoma occur in less than 5% of HIV-infected patients in the United States, although the worldwide prevalence in acquired immunodeficiency syndrome (AIDS) patients exceeds 30%. Kaposi’s sarcoma is a vascular neoplasm that can affect cutaneous and visceral tissues. While the rates of Kaposi’s sarcoma development have sharply declined since the widespread use of antiretroviral therapy, the rates of other cutaneous malignancies have remained stable. The risk of an HIV-infected patient developing a cutaneous malignancy is about 5.7%, with basal cell carcinoma being the most common type encountered.

With regard to general surgical considerations in HIV patients, contributing related morbidities such as malnutrition, decreased CD4 count, and presence of opportunistic infection may result in delayed and attenuated wound healing capacity.

**BENIGN TUMORS**

**Hemangioma**

Hemangiomas are benign vascular tumors that arise from the proliferation of endothelial cells that surround blood-filled cavities. They occur in about 4% of children by 1 year of age. Their natural history is typically presentation shortly after birth, a period of rapid growth during the first year, and then gradual involution over childhood in more than 90% of cases. These hemangiomas are generally managed nonsurgically prior to involution. Occasionally, during the rapid growth phase, the lesions can obstruct the airway, GI tract, vision, and musculoskeletal function. In these cases, surgical resection is indicated prior to the involution phase. Hemangiomas can sometimes consume a large percentage of cardiac output, resulting in high-output heart failure or a consumptive coagulopathy, which may also necessitate resection. These lesions characteristically express the GLUT-1 glucose transporter protein, which is absent in cells of the normal cutaneous vasculature. First-line therapy for these infantile hemangiomas is propranolol, which causes cessation of growth and, in most cases, actual regression of the lesions. Systemic corticosteroids and interferon-α can impede tumor progression, and laser therapy has been used as well. If tumors persist into adolescence leaving a cosmetically undesirable defect, surgical resection may be considered. When surgical resection or debulking is considered, upfront selective embolization can help with planned resection.

**Nevi**

Nevi (singular, nevus) are areas of melanocytic hyperplasia or neoplasia. These collections can be found in the epidermis (junctional), partially in the dermis (compound), or completely within the dermis (dermal). They commonly develop in childhood and young adulthood, and will sometimes spontaneously regress. Exposure to UV radiation is associated with increased density of these lesions. Nevus are typically symmetric and small. Congenital nevi are the result of abnormal development of melanocytes. The events leading to this abnormal development may also affect the surrounding cells, resulting in longer, darker hair. Congenital nevi are found in less than 1% of neonates, and when characterized as giant congenital nevi, they have up to a 5% chance of developing into a malignant melanoma, and may do so even in the first years of childhood. Treatment, therefore, consists of surgical excision of the lesion as early as is feasible. For larger lesions, serial excision and tissue expansion may be required, with the goal of lesion excision being maintenance of function and form while decreasing oncologic risk.

**Cystic Lesions**

Cutaneous cysts are benign lesions that are characterized by overgrowth of epidermis towards the center of the lesion, resulting in keratin accumulation. Epidermoid cysts (often mistakenly referred to as sebaceous cysts) are classically the result of keratin-plugged pilosebaceous units. They commonly affect adult men and women, and present as a dermal or subcutaneous cyst with a single, keratin-plugged punctum at the skin surface, often at or above the upper chest and back. Epidermoid cysts are the most common cutaneous cyst and are histologically characterized by mature epidermis complete with granular layer. Another type of cystic lesion is known as a trichilemmal cyst. These cysts are derived from the outer sheath of hair follicles, and, in contrast to epidermoid cysts, lack a granular layer. They are almost always found on the scalp and more commonly in women. A third type of cutaneous cyst is a dermoid cyst. Dermoid cysts are congenital variants that occur as the result of persistent epithelium within embryonic lines of fusion. They occur most commonly between the forehead and nose tip, and the most frequent site is the eyebrow. They can lie in the subcutaneous tissue or intracranially, and often communicate with the skin surface via a small fistula. These cystic structures contain epithelial tissue, hair, and a variety of epidermal appendages. Treatment for these cystic structures includes surgical excision with care taken to remove the cyst lining to prevent recurrence.
Keratosis

Actinic Keratosis. Actinic keratoses are neoplasms of epidermal keratinocytes that represent a range in a spectrum of disease from sun damage to squamous cell carcinoma. They typically occur in fair-skinned, elderly individuals in primarily sun-exposed areas, and UV radiation exposure is the greatest risk factor. There are multiple variants, and they can present as erythematous and scaly to hypertrophic, keratinized lesions. They can become symptomatic, causing bleeding, pruritus and pain. They can regress spontaneously, persist without change, and transform into invasive squamous cell carcinoma. It is estimated that approximately 10% of actinic keratoses will transform into invasive squamous cell carcinoma, and that progression takes about 2 years on average. About 60% to 65% of squamous cell carcinomas are believed to originate from actinic keratoses. The presence of actinic keratoses also serves as a predictor of development of other squamous cell and basal cell carcinomas. Treatment options are excision, cautery and destruction, and dermabrasion.

Seborrhiec Keratosis. Seborrhiec keratoses are benign lesions of the epidermis that typically present as well-demarcated, “stuck on” appearing papules or plaques over elderly individuals. Clonal expansion of keratinocytes and melanocytes make up the substance of these lesions. They carry no malignant potential and treatment is primarily for cosmetic purposes.

Soft Tissue Tumors

Acrochordons. Acrochordons, also known as skin tags, are benign, pedunculated lesions on the skin made up of epidermal keratinocytes surrounding a collagenous core. Although they can become irritated or necrotic, their removal is generally cosmetic.

Dermatofibromas. Dermatofibromas are benign cutaneous proliferations that appear most commonly on the lower extremities of women. They appear as pink to brown papules that pucker or dimple in the center when the lesion is pinched. It remains unclear whether these lesions have a neoplastic etiology or if they are the result of minor trauma or infection. These lesions are typically asymptomatic, and treatment is only indicated for cosmetic concerns or when a histologic diagnosis is required. Surgical excision is the recommended treatment, although cryotherapy and laser treatment may be used.

Lipomas. Lipomas are the most common subcutaneous neoplasm and have no malignant potential. They present as a painless, slow-growing, mobile mass of the subcutaneous tissue. Usually less than 5 cm in diameter, these neoplasms can reach much larger sizes. Lipomas are largely asymptomatic but may cause pain due to regional nerve deformation. Surgical resection is indicated in cases of local pain, mass effect, or cosmetically sensitive areas. The tumors are usually well circumscribed and amenable to surgical resection. Liposarcoma is a malignant fatty tumor that can mimic a lipoma, but it is often deep-seated, rapidly growing, painful, and invasive. In these cases, cross-sectional imaging is recommended prior to any surgical resection.

Neural Tumors

Neuromas. Neuromas do not represent a true clonal proliferation of neural tissue, but rather disordered growth of Schwann cells and nerve axons, often at the site of previous trauma. They can present within surgical scar lines or at the site of previous trauma as flesh-colored papules or nodules and are typically painful.

Schwannomas. A schwannoma is a benign proliferation of the Schwann cells of the peripheral nerve sheath, and can arise sporadically or in association with type 2 neurofibromatosis. It contains no axons, but may displace the affected nerve and cause pain along the distribution of the nerve.

Neurofibromas. Neurofibromas, in contrast, are benign proliferations that are made up of all nerve elements, and arise as fleshy and nontender, sessile or pedunculated masses on the skin. They can arise sporadically or in association with type 1 neurofibromatosis, and in these cases, are associated with café-au-lait spots and Lisch nodules. They are often asymptomatic, but may be pruritic. The development of pain at the site of a previously asymptomatic neurofibroma may indicate a rare malignant transformation and requires surgical excision and biopsy.

MALIGNANT TUMORS

Basal Cell Carcinoma

Basal cell carcinoma (BCC) is the most common tumor diagnosed in the United States, with an estimated one million new cases occurring each year. It represents 75% of nonmelanoma skin cancers and 25% of all cancers diagnosed each year. BCC is seen slightly more commonly in males and individuals over the age of 60, though the incidence in younger age groups is increasing. The primary risk factor for disease development is sun exposure (UVB rays more than UVA rays), particularly during adolescence. The pathogenesis of BCC stems from mutations of genes involved in tumor suppression, often caused by ionizing radiation. The p53 tumor suppressor gene is defective in approximately 50% of cases. There is a latency period of 20 to 50 years.

BCC tends to occur on sun-exposed areas of the skin, most commonly the nose and other parts of the face. A malignant lesion on the upper lip is almost always BCC, and BCC is the most common malignant eyelid tumor. Because of the photoprotective effect of melanin, dark-skinned individuals are far less commonly affected. Other risk factors for development of BCC include immune suppression, chemical exposure, and ionizing radiation exposure. There are also genetic susceptibilities to development of BCC in conditions such as xeroderma pigmentosa, unilateral basal cell nevus syndrome, and nevoid BCC syndrome.

The natural history of BCC is typically one of local invasion rather than distant metastasis, but untreated BCC can often result in significant morbidity.

There are multiple variants of BCC, and presentation can range from red, flesh-colored, or white macule or papule, to nodules and ulcerated lesions. Growth patterns of these lesions can either be well-circumscribed or diffuse and the most common types of BCC are nodular and micronodular, superficial spreading, and infiltrative.

The most common subtype is the nodular variant, characterized by raised, pearly pink papules with telangiectasias and occasionally a depressed tumor center with raised borders giving the classic “rodent ulcer” appearance. Superficial spreading BCC is confined to the epidermis as a flat, pink, scaling or crusting lesion, often mistaken for eczema, actinic keratosis, fungal infection, or psoriasis. This subtype typically appears on the trunk or extremities and the mean age of diagnosis is 57 years. The infiltrative form appears on the
head and neck in the late 60s, often at embryonic fusion lines, with an opaque yellow-white color that blends with surrounding skin and has no raised edges. The morpheaform subtype represents 2% to 3% of all BCC and is the most aggressive subtype. It usually presents as an indurated macule or papule with the appearance of an enlarging scar. The clinical margins are often indistinct, and the rate of positive margins after excision is high. There is also a pigmented variant of BCC that can be difficult to distinguish from certain melanoma subtypes.

Treatment of BCC varies according to size, location, type, and high- or low-risk. Treatment options include surgical excision, medical, or destructive therapies. Surgical excision should include 4 mm margins for small, primary BCC on cosmetically sensitive areas, and 10 mm margins otherwise. Mohs microsurgical excision is sequential horizontal excision and has been shown to be cost-effective and associated with low recurrence rates for BCC (1%). It is the treatment of choice for morpheaform or other BCC with aggressive features, poorly delineated margins, recurrent tumors, or cosmetically sensitive areas, especially in the midface. A common approach used by dermatologists for very small (<2 mm) and low risk lesions is curettage and destruction, although it should be kept in mind that the local cure rates can be operator and institution dependent. Other destructive techniques include cryosurgery and laser ablation. Radiation therapy can be used as adjuvant therapy following surgery, or as primary therapy in poor surgical candidates with low-risk lesions. The practitioner must be aware of the potential consequences of radiation therapy, including poor cosmetic outcomes and future cancer risk.

Superficial medical therapies are generally reserved for patients in whom surgical and radiation treatment is not an option. Topical imiquimod or 5-fluorouracil have been used for periods of 6 to 16 weeks for small, superficial BCC of the neck, trunk or extremities. Lastly, topical photodynamic therapy has shown some benefit in treatment of premalignant or superficial low-risk lesions as well.

Patients with BCC need to have regular follow-up with full skin examinations every 6 to 12 months. Sixty-six percent of recurrences develop within 3 years, and with a few exceptions occurring decades after initial treatment, the remaining recur within 5 years of initial treatment. A second primary BCC may develop after treatment and, in 40% of cases, presents within the first 3 years after treatment.

Squamous Cell Carcinoma

Squamous cell carcinoma (SCC) is the second most common skin cancer and accounts for approximately 100,000 cases each year. The primary risk factor for the development of SCC is UV radiation exposure; however, other risks include light Fitzpatrick skin type (I or II), environmental factors such as chemical agents, physical agents (ionizing radiation), psoralen, HPV-16 and -18 infections, immunosuppression, smoking, chronic wounds, burn scars, and chronic dermatoses. Heritable risk factors include xeroderma pigmentosum, epidermolysis bullosa, and oculocutaneous albinism.

SCC classically appears as a scaly or ulcerated papule or plaque, and bleeding of the lesion with minimal trauma is not uncommon, but pain is rare. It can exhibit in situ (confined to the epidermis) or invasive subtypes. The most common in situ variant of SCC is actinic keratosis, described previously in this chapter. Invasive squamous cell carcinomas may arise de novo, but more commonly evolve from these precursors. Another in situ variant is known as Bowen disease. This is characterized by full-thickness epidermal dysplasia and clinically appears as a scaly, erythematous patch often with pigmentation and fissuring. When it occurs on the glans penis, it is known as erythroplasia of Queyrat. Ten percent of these cases will eventually become invasive. Outside of these instances, most in situ cases grow slowly and do not progress to invasive disease.

Invasive SCC is characterized by invasion through the basement membrane into the dermis of the skin. It usually arises from an actinic keratosis precursor, but de novo varieties do occur and are higher risk. De novo invasive SCC commonly occurs in organ transplant and immunocompromised patients, and has a metastatic rate as high as 14%. De novo invasive SCC arising in areas of chronic wounds or burn scars are known as Marjolin’s ulcers, and have a higher metastatic potential (Fig. 16-9). Keratoacanthoma is now being accepted as a subtype of SCC that is characterized by a rapidly growing nodule with a central keratin plug. The natural history of invasive disease depends on location and inherent tumor characteristics. Clinical risk factors for recurrence include presentation with neurologic symptoms, immunosuppression, tumor with poorly defined borders, and tumor that arises at a site of prior radiation. Perineural involvement also has a poorer survival with increased local recurrence and lymph node metastasis. Grades of differentiation are based on the ratio of differentiated to undifferentiated cells, with a lower ratio associated with a greater metastatic and recurrent potential. Large (>2 cm) lesions, depth of invasion >4 mm, rapid growth, and location on the ear, lips, nose, scalp, or genitals are all also indicators of worse prognosis.

When feasible, wide surgical excision including subcutaneous fat is the treatment of choice for SCC. Margins of 4 mm are recommended for low-risk lesions and 6 mm for high-risk lesions. Mohs microsurgical excision is indicated for positive margins, recurrent tumors, sites where cosmesis or function preservation is critical, poorly differentiated tumors, invasive lesions, and verrucous tumors. Using this modality often results in lower recurrence rates. It has also found use in nail bed lesions and in those arising in a background of osteomyelitis. The role of lymph node dissection in the setting of SCC continues to evolve. Lymphadenectomy is indicated following fine-needle aspiration or core biopsy for clinically palpable lymph nodes or nodes detected on cross-sectional imaging. Nodes
should also be removed from susceptible regional lymph node basins in patients with SCC in the setting of chronic wounds. Patients with parotid disease benefit from a superficial or total parotidectomy (with facial nerve preservation) and adjuvant radiotherapy. Sentinel lymph node dissection may be used in high-risk cases with clinically negative nodal disease. Radiation therapy is typically reserved as primary therapy for those who are poor surgical candidates, and as adjuvant therapy after surgical resection for large, high-risk tumors. When used as primary therapy, cure rates may approach 90%.^{121}

### Melanoma

**Background.** In 2017, an estimated 87,110 new cases of melanoma were diagnosed, as well as 9730 melanoma-related deaths. The incidence of melanoma is rising faster than most other solid malignancies, and these numbers likely represent an underestimation given the many in situ and thin melanoma cases that are underreported. These tumors primarily arise from melanocytes at the epidermal-dermal junction but may also originate from mucosal surfaces of the oropharynx, nasopharynx, eyes, proximal esophagus, anorectum, and female genitalia. Melanoma characteristically metastasizes quite often, and can travel to most other tissues in the body. This metastasis confers a poor prognosis in patients, with a median life span of 6 to 8 months after diagnosis.^{132}

The most important risk factor for the development of melanoma is exposure to UV radiation. It was recently reported that greater than 10 tanning bed sessions by adolescents and young adults increased their relative risk of developing melanoma twofold,^{133} and there is a positive association with intermittent childhood sunburns and melanoma development.^{134} There is also an association with residence at high altitudes or in close proximity to the equator. Both personal and family history of melanomas increase the risk of primary melanoma development. Individuals with dysplastic nevi have a 6% to 10% overall lifetime risk of melanoma, with tumors arising from preexisting nevi or de novo. Individuals with familial atypical multiple mole melanoma syndrome have numerous melanocytic nevi and a greatly increased risk of cutaneous melanoma. Congenital nevi increase the risk for melanoma proportionally with size, and giant congenital nevi (generally considered >20 cm in diameter) are associated with a 5% to 8% lifetime risk. Melanoma development is strongly associated with the p16/CDK4,6/Rb and p14ARF/HMD2/p53 tumor suppressor pathways and the RAF-MEK-ERK and PI3K-Akt oncogenic pathways.^{135}

**Clinical Presentation.** The presentation of melanoma is commonly used to determine subtype but often starts as a localized, radial growth phase followed by a more aggressive, vertical growth phase. Approximately 30% of melanoma lesions arise from a preexisting melanocytic nevus. The most common subtype of melanoma is superficial spreading (Fig. 16-10). This accounts for 50% to 70% of melanomas and typically arises from a precursor melanocytic nevus. Nodular subtype accounts for 15% to 30% of melanomas, and typically arises de novo, most commonly in men and on the trunk (Figs. 16-11 and 16-12). This subtype is aggressive with an early vertical growth pattern and is often diagnosed at a later stage. Up to 5% of these lesions will lack melanin and can be mistaken for other cutaneous lesions. Lentigo maligna represents 10% of melanoma cases and is a less aggressive subtype of melanoma in situ that typically arises on sun-exposed areas of the head and neck. Acral lentiginous melanoma accounts for 29% to 72% of melanomas in dark-skinned individuals, is occasionally seen in Caucasians, and is found on palmar, plantar, and subungual surfaces. This subtype is not thought to be due to sun exposure.

**Diagnosis and Staging.** Workup should begin with a history and physical exam. The entire skin should be checked for synchronous primaries, satellite lesions, and in-transit metastases, and all nodal basins should be examined for lymphadenopathy. Suspicious lesions should undergo excisional biopsy with 1- to 3-mm margins; however, tumors that are large or are in a cosmetically or anatomically challenging area can be approached by incisional biopsy, including punch biopsy.^{136}

---

**Figure 16-10.** Primary cutaneous melanoma seen in the scalp of a 61-year-old male.

**Figure 16-11.** Nodular melanoma seen in the leg of a 55-year-old male.
Melanoma is characterized according to the American Joint Committee on Cancer (AJCC) as localized disease (stage I and II), regional disease (stage III), or distant metastatic disease (stage IV). The Breslow tumor thickness replaced the Clark’s level as the most important prognostic indicator for melanoma staging.132,140 The Breslow tumor thickness measures the depth of penetration of the lesions from the top of the granular layer of the epidermis into the dermal layer and is directly related to the risk of disease progression. Tumor ulceration, mitotic rate $\geq 1$ per mm$^2$, and metastasis are all also associated with worse prognosis. In the presence of regional node metastasis, the number of nodes affected is the most important prognostic indicator. For stage IV disease, the site of metastasis is strongly associated with prognosis, and elevated lactate dehydrogenase (LDH) is associated with a worse prognosis.141

There is no supportive evidence for chest X-ray or computed tomography (CT) in the staging of patients unless there is positive regional lymph node disease, although it can be used to work up specific signs and symptoms when metastatic disease is suspected.136 In patients with stage III or greater disease, there is a high risk for distant metastasis, and imaging is recommended for baseline staging. These patients should receive additional imaging that includes CT of the chest, abdomen, and pelvis; whole-body positron emission tomography (PET)-CT; or brain magnetic resonance imaging (MRI).136

The sentinel lymph node biopsy (SLNB) technique for melanoma was introduced in 1992 and has become a cornerstone in the management of melanoma, although its role in management continues to be refined. SLNB is a standard staging procedure to evaluate the regional nodes for patients with clinically node-negative malignant melanoma. Detecting subclinical nodal metastasis may benefit from lymphadenectomy or adjuvant therapy. This technique identifies the first draining lymph node from the primary lesion and has shown excellent accuracy and significantly less morbidity compared to complete resection of nodal basins. It is almost always performed at the time of initial wide excision, as SLN mapping after lymphatic violation from surgical excision could decrease the accuracy of the test. Recently, the results of MSLT-1, an international, multicenter, phase III trial were published. This study randomized clinically node negative patients to either SLNB at the time of primary melanoma excision (and completion lymphadenectomy if positive) or nodal basin monitoring (and delayed complete lymphadenectomy for recurrent lymph node disease).142 The results of this study demonstrated that SLNB, with immediate lymphadenectomy if positive, improved disease-free survival by 7% and 10% in patients with intermediate thickness (1.2–3.5 mm) and thick (>3.5 mm) lesions respectively. The use of SLNB in lesions <1.2 mm thick did not affect disease-free survival. SLNB should also be offered to thin lesions with high-risk features (thickness >0.75, ulceration, mitoses $\geq 1$ per mm$^2$).136 The SLNB involves preoperative lymphoscintigraphy with intradermal injections of technetium-sulfur colloid to delineate lymphatic drainage and intraoperative intradermal injection of 1 mL of isosulfan or methylene blue dye near the tumor or biopsy site. (Figs. 16-13 and 16-14). The radioactive tracer-dye combination allows the sentinel node to be identified in 98% of cases. An incision over the lymph node basin of interest allows nodes to be excised and studied with hematoxylin and eosin and immunohistochemistry (S100, HMB45, and MART-1/Melan-A) staining (Fig. 16-15).
Surgical Management of the Primary Tumors and Lymph Nodes. The appropriate excision margin is based on primary tumor thickness. Several retrospective studies suggest that for melanoma in situ, 0.5 to 1 cm margins are sufficient.\textsuperscript{143-145} We believe that 1-cm margins should be obtained in anatomically feasible areas given the possibility of an incidental finding of a small invasive component in permanent sections. Several studies compared 1- to 3-cm margins and 2- to 5-cm margins in melanoma <2 mm thick, and 2- to 4-cm margins in melanoma lesions 1 to 4 mm thick and found no difference.\textsuperscript{146-149} A British trial suggested that there is a limit to how narrow margins can be for melanomas >2 mm thick by showing that 1-cm margins provide worse outcomes compared to 3-cm margins.\textsuperscript{150} Tumors <1 mm thick require 0.5 to 1 cm margins. Tumors 1 to 2 mm thick require 1 to 2 cm margins, and tumors >2 mm thick require 2-cm margins.

Completion lymphadenectomy is commonly performed in cases of sentinel nodes with metastatic disease, but it has been shown that most of these nodal basins do not have additional disease. Thus, many surgeons do not perform routine completion lymphadenectomy for positive nodes, and data from the MSLT-2 may provide guidance. It has been shown that those patients with nonsentinel lymph node positivity found on completion lymph node dissection after a positive SLN have higher rates of recurrence and lower rates of survival. The therapeutic value, however, has not been clearly demonstrated. In patients with clinically positive lymph nodes but absent signs of distant metastasis on PET-CT, therapeutic lymph node dissection is associated with 5-year survival rates of 30% to 50%. In these cases, resection of the primary melanoma lesion and a completion lymphadenectomy should be performed.

Individuals with face, anterior scalp, and ear primaries who have a positive SLNB should undergo a superficial parotidectomy in addition to a modified radical neck dissection.
Figure 16-15. Operation of sentinel lymph node biopsy for cutaneous melanoma. After preoperative injection of radioactive technetium-99m-labeled sulfur colloid tracer and intraoperative injection of Lymphazurin blue dye around the primary melanoma excision site, the nodal basin of interest is identified. An incision is made directly overlying the lymph node basin in the posterior axillary space. The sentinel lymph nodes are identified and excised.

Patients with positive sentinel nodes in the inguino-femoral nodal basin should undergo an inguino-femoral lymphadenectomy that includes removal of Cloquet’s node. If Cloquet’s node is positive or the patient has three or more nodes that contain melanoma metastases the probability of clinically occult positive pelvic nodes is increased. The effect of ileo-obturator lymph node dissection on the survival of these patients is unknown.

Surgery for Regional and Distant Metastasis. Nonmetastatic, in-transit disease should undergo excision to clear margins when feasible. However, disease not amenable to complete excision derives benefit from isolated limb perfusion (ILP) and isolated limb infusion (ILI) (Fig. 16-16). These two modalities are used to treat regional disease, and their purpose is to administer high doses of chemotherapy, commonly melphalan, to an affected limb while avoiding systemic drug toxicity. ILI was shown to provide a 31% response rate in one study, while hyperthermic ILP provided a 63% complete response rate in an independent study.151-154

The most common sites of metastasis of melanoma are the lung and liver. These are followed by the brain, gastrointestinal tract, distant skin, and subcutaneous tissue. A limited subset of patients with small-volume, limited distant metastases to the brain, gastrointestinal tract, or distant skin can be treated with surgical resection or directed radiation. Liver metastases are better dealt without surgical resection unless they arise from an ocular primary. Adjuvant therapy after resection of metastatic lesions is not standard of care. However, there are ongoing clinical trials addressing whether drugs and vaccines will be beneficial in this setting.115 Surgery may provide palliation for patients with gastrointestinal obstruction, gastrointestinal hemorrhage, and nongastrointestinal hemorrhage. Radiotherapy for symptomatic bony or brain metastases provides palliation in diffuse disease.

Adjuvant and Palliative Therapies. Eastern Cooperative Oncology Group (ECOG) Trials 1684, 1690, and 1694 were prospective randomized controlled trials that demonstrated...
disease-free survival advantages in patients with melanoma >4 mm in thickness with or without lymph node involvement if they received adjuvant treatment with high-dose interferon (IFN). A European Organization for Research and Treatment of Cancer (EORTC) trial also showed recurrence-free survival benefit with pegylated IFN. It is important to note that IFN therapy is not well tolerated and the pooled analysis of these trials did not show an improvement in overall survival benefit.

Most patients with melanoma will not be surgical candidates. Although medical options for melanoma have historically been poor, several recent studies have shown promise in drug therapy for metastatic melanoma. BRAF inhibitors (sorafenib), anti-PD1 antibodies, CTLA antibodies (ipilimumab), and high-dose interleukin-2 (IL-2) with and without vaccines have been shown in randomized studies to provide survival benefit in metastatic disease. Despite the excitement of recent drugs, surgery will likely play an adjunct role in treating individuals who develop resistance to these drugs over time.

Special Circumstances. Special circumstances of note are melanoma in pregnant women, melanoma of unknown primaries, and noncutaneous melanomas. The prognosis of pregnant patients is similar to women who are not pregnant. Extrapolation of studies examining the SLNB technique in pregnant women with breast cancer suggests lymphoscintigraphy may be done safely during pregnancy without risk to the fetus (blue dye is contraindicated). General anesthesia should be avoided during the first trimester, and local anesthetics should be used during this time. It has been suggested by some that after excising the primary tumor during pregnancy, the SLNB may be performed after delivery.

Unknown primary melanoma occurs in 2% to 5% of cases and most commonly occurs in the lymph nodes. In these cases, a thorough search for the primary lesion should be sought, including eliciting a history about prior skin lesions, skin procedures (e.g., curettage and electrodesiccation, excision, laser), and review of any prior “benign” pathology. The surgeon should be aware that melanoma is known to spontaneously regress because of an immune response. Melanoma of unknown primary has survival rates comparable to melanoma diagnosed with a known primary of the same stage.

The most common noncutaneous disease site is ocular melanoma, and treatment of this condition includes photoocoagulation, partial resection, radiation, or enucleation. Ocular melanomas exclusively metastasize to the liver and not regional lymph nodes, and some patients benefit from liver resection. Melanoma of the mucous membranes most commonly presents in the oral cavity, oropharynx, nasopharynx, paranasal sinus, anus, rectum, and female genitalia. Patients with this presentation have a worse prognosis (10% 5-year survival) than patients with cutaneous melanomas. Management should be excision to negative margins, and radical resections should be avoided because the role of surgery is loco-regional control, not cure. Generally speaking, lymph node dissection should be avoided because the benefit is unclear.

Merkel Cell Carcinoma
Merkel cell carcinoma (MCC) is an aggressive neuroendocrine tumor of the skin whose incidence has been rapidly increasing. Although it is a much rarer malignancy than melanoma, the prognosis is much worse, with a 5-year survival of 46%. Merkel cells are epidermal appendages involved in the sensation of light touch, and along with Merkel cell carcinoma, are cyto-keratin-20 positive. This stain is now used to confirm the diagnosis. Other risk factors include age >65 years (the median age of diagnosis is 70 years), UV exposure, Merkel cell polyoma virus, and immunosuppression. MCC typically presents as a rapidly growing, flesh-colored to red or purple papule or plaque (Fig. 16-17). Regional nodes are involved in 30% of patients at diagnosis, and 50% will develop systemic disease (skin, lymph nodes, liver, lung, bone, and brain). There are no standardized diagnostic imaging studies for staging, but CT of the chest, abdomen, pelvis and octreotide scans may provide useful information when clinically indicated.

After a thorough skin examination, treatment should begin by evaluating nodal basins. Patients without clinical nodal disease should undergo an SLNB prior to wide local excision because studies suggest a benefit. In patients with sentinel lymph nodes with metastatic disease, completion lymphadenectomy and/or radiation therapy may follow, and in patients with node-negative disease, observation or radiation therapy should be considered. SLNB is important for staging and treatment, and the literature suggests that it predicts recurrence and relapse-free survival. Elective lymph node dissection may decrease regional nodal recurrence and in-transit metastases. Patients with clinically positive nodes should have an FNA to confirm disease. If positive, a metastatic staging workup should follow, and, if negative, treatment of the primary and nodal basin as managed for sentinel lymph node-positive disease should be considered. A negative FNA and open biopsy-negative disease should be managed by treatment of the primary disease alone.

Figure 16-17. Merkel cell carcinoma seen just above the left knee in a 44-year-old female.
Patients with metastatic disease should be managed according to consensus from a multidisciplinary tumor board.

Important surgical principles for excision of the primary lesion are to excise with wide margins down to fascia and complete circumferential and peripheral deep-margin assessment. Recommended management for margins is 1 to 3 cm, but there are no randomized trials defining these margins. Chemotherapy and adjuvant radiation are commonly used, but there are no data to support a specific regimen or that demonstrate a definitive survival benefit.

Recurrence of MCC is common. One study of 95 patients showed a 47% recurrence, with 80% of recurrences occurring within 2 years and 96% occurring within 5 years.\textsuperscript{173,174} Regional lymph node disease is common, and 70% of patients will have nodal spread within 2 years of disease presentation. Five-year overall survival of head and neck disease in surgically treated patients is between 40% and 68%.

**Kaposi’s Sarcoma**

Kaposi’s sarcoma is characterized by the proliferation and inflammation of endothelial-derived spindle cell lesions. There are five major forms of this angioproliferative disorder: classic (Mediterranean), African endemic, HIV-negative men having sex with men (MSM)-associated, and immunosuppression-associated. They are all driven by the human herpesvirus (HHV-8).\textsuperscript{175} Kaposi’s sarcoma is diagnosed after the fifth decade of life and predominantly found on the skin but can occur anywhere in the body. In North America, the Kaposi’s sarcoma herpes virus is transmitted via sexual and nonsexual routes and predominantly affects individuals with compromised immune systems such as those with HIV and transplant recipients on immune-suppressing medications. Clinically, Kaposi’s sarcoma appears as multifocal, rubbery blue-red nodules. Treatment of AIDS-associated Kaposi’s sarcoma is with antiviral therapy, and many patients experience a dramatic treatment response.\textsuperscript{176,177} Those individuals who do not respond and have limited mucocutaneous disease may benefit from cryotherapy, photodynamic therapy, radiation therapy, intralesional injections, and topical therapy. Surgical biopsy is important for disease diagnosis, but given the high local recurrence and the fact that Kaposi’s sarcoma represents more of a systemic rather than local disease, the benefit of surgery is limited and generally should not be pursued except for palliation.

**Dermatofibrosarcoma Protuberans**

This rare, low-grade sarcoma of fibroblast origin commonly afflicts individuals during their third decade of life. It has low distant metastatic potential, but it behaves aggressively locally with finger-like extensions. Tumor depth is the most important prognostic variable. Presentation is characteristically a slow-growing, asymptomatic, violaceous plaque involving the trunk, head, neck, or extremities (Fig. 16-18). Nearly all cases are positive for CD34 and negative for factor XIIIa.\textsuperscript{178,179} Treatment is wide local excision with 3-cm margins down to deep underlying fascia or Mohs microsurgery in cosmetically sensitive areas where maximum tissue preservation will benefit.\textsuperscript{180} No nodal dissection is needed, and both approaches provide similar local control.\textsuperscript{181} Some clinicians have used radiation therapy and biologic agents (imatinib) as adjuvant therapy with some success in patients with advanced disease. Local recurrence occurs in 50% to 75% of cases, usually within 3 years of treatment. Thus, clinical follow-up is important. Recurrent tumors should be resected whenever possible.

**Malignant Fibrous Histiocytoma (Undifferentiated Pleomorphic Sarcoma and Myxofibrosarcoma)**

This uncommon, cutaneous, spindle-cell, soft tissue sarcoma occurs in the extremities, head, and neck of elderly patients. They present as solitary, soft to firm, skin-colored subcutaneous nodules. Complete surgical resection is the treatment of choice, and adjuvant radiation therapy provides local control; patients with positive margins benefit most from this combination. Nevertheless, patients undergoing complete gross resection will experience recurrence in 30% to 35% of cases.\textsuperscript{182} Up to 50% of patients may present with distant metastasis, and this is a contraindication to surgical resection.

**Angiosarcoma**

Angiosarcoma is an uncommon, aggressive cancer that arises from vascular endothelial cells and occurs in four variants, all of which have a poor prognosis.\textsuperscript{183} The 5-year survival estimate is 15%. The head and neck variant presents in individuals older than 40 years as an ill-defined red patch on the face or scalp, often with satellite lesions and distant metastasis, and has a median survival of 18 to 28 months. Lymphedema-associated angiosarcoma (Stewart-Treves) develops on an extremity ipsilateral to an axillary lymphadenectomy. It appears on the upper, medial arm as a violaceous plaque in an individual with nonpitting edema and has a poor survival. Radiation-induced angiosarcoma occurs 4 to 25 years after radiation therapy for benign and malignant conditions. Finally, the epithelioid variant of angiosarcoma involves the lower extremities and also has a poor prognosis. Surgical excision with wide margins is the treatment...
of choice for localized disease, but the rate of recurrence is high. Adjuvant radiation therapy can be considered in a multidisciplinary fashion. Cases of extremity disease can be considered for amputation. For widely metastatic disease, chemotherapy and radiation may provide palliation, but these modalities do not prolong overall survival.\textsuperscript{115}

**Extramammary Paget’s Disease**

This rare adenocarcinoma of apocrine glands arises in axillary, perianal, and genital regions of men and women.\textsuperscript{184} Clinical presentation is that of erythematous or nonpigmented plaques with an eczema-like appearance that often persist after failed treatment from other therapies. An important characteristic and one that the surgeon must be acutely aware of is the high incidence of concomitant other malignancies with this cutaneous disease. Forty percent of cases are associated with primary gastrointestinal and genitourinary malignancies, and a diligent search should be made after a diagnosis of extramammary Paget’s disease is made. Treatment is surgical resection with negative microscopic margins, and adjuvant radiation may provide additional locoregional control.

**CONCLUSION**

The skin is the largest organ in the human body and is composed of three organized layers that are the source of numerous pathologies. Recognition and management of cutaneous and subcutaneous diseases require an astute clinician to optimize clinical outcomes. Improvements in drugs, therapies, and healthcare practices have helped recovery from skin injuries. Skin and subcutaneous diseases are often managed medically, although surgery frequently complements treatment. Benign tumors are surgical diseases, while malignant tumors are primarily treated surgically, and additional modalities including chemotherapy and radiation therapy are sometimes required. The management of melanoma is at an exciting phase, requiring the coordinated multidisciplinary care of medical oncologists, surgical oncologists, radiation oncologists, dermatopathologists, and plastic and reconstructive surgeons. The advent of new drug therapies will redefine the role of surgery in this disease in the coming years.

**REFERENCES**

Entries highlighted in bright blue are key references.


Breast cancer has captured the attention of surgeons throughout the ages. The Smith Surgical Papyrus (3000–2500 B.C.) is the earliest known document to refer to breast cancer. The cancer was in a man, but the description encompassed most of the common clinical features. In reference to this cancer, the author concluded, “There is no treatment.” There were few other historical references to breast cancer until the first century. In De Medicina, Celsus commented on the value of operations for early breast cancer: “None of these may be removed but the cacoethes (early cancer), the rest are irritated by every method of cure. The more violent the operations are, the more angry they grow.” In the second century, Galen inscribed his classical clinical observation: “We have often seen in the breast a tumor exactly resembling the animal the crab. Just as the crab has legs on both sides of his body, so in this disease the veins extending out from the unnatural growth take the shape of a crab’s legs. We have often cured this disease in its early stages, but after it has reached a large size, no one has cured it. In all operations we attempt to excise the tumor in a circle where it borders on the healthy tissue.”

The Galenic system of medicine ascribed cancers to an excess of black bile and concluded that excision of a local bodily outbreak could not cure the systemic imbalance. Theories espoused by Galen dominated medicine until the Renaissance. In 1652, Tulp introduced the idea that cancer was contagious when he reported an elderly woman and her housemaid who both developed breast cancer (N. Tulp, Observationes medicæ 1652). This single incidence was accepted as conclusive.

---

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Brief History of Breast Cancer Therapy</td>
<td>541</td>
</tr>
<tr>
<td>Embryology and Functional Anatomy of the Breast</td>
<td>543</td>
</tr>
<tr>
<td>Physiology of the Breast</td>
<td>547</td>
</tr>
<tr>
<td>Infectious and Inflammatory Disorders of the Breast</td>
<td>550</td>
</tr>
<tr>
<td>Common Benign Disorders and Diseases of the Breast</td>
<td>551</td>
</tr>
<tr>
<td>Aberrations of Normal Development and Involvement</td>
<td>551</td>
</tr>
<tr>
<td>Pathology of Nonproliferative Disorders</td>
<td>552</td>
</tr>
<tr>
<td>Pathology of Proliferative Disorders Without Atypia</td>
<td>553</td>
</tr>
<tr>
<td>Pathology of Atypical Proliferative Diseases</td>
<td>553</td>
</tr>
<tr>
<td>Treatment of Selected Benign Breast Disorders and Diseases</td>
<td>554</td>
</tr>
<tr>
<td>Risk Factors for Breast Cancer</td>
<td>555</td>
</tr>
<tr>
<td>Hormonal and Nonhormonal Risk Factors</td>
<td>555</td>
</tr>
<tr>
<td>Risk Assessment Models</td>
<td>555</td>
</tr>
<tr>
<td>Risk Management</td>
<td>556</td>
</tr>
<tr>
<td>BRCA Mutations</td>
<td>558</td>
</tr>
<tr>
<td>Epidemiology and Natural History of Breast Cancer</td>
<td>561</td>
</tr>
<tr>
<td>Epidemiology</td>
<td>561</td>
</tr>
<tr>
<td>Natural History</td>
<td>562</td>
</tr>
<tr>
<td>Histopathology of Breast Cancer</td>
<td>563</td>
</tr>
<tr>
<td>Carcinoma In Situ</td>
<td>563</td>
</tr>
<tr>
<td>Invasive Breast Carcinoma</td>
<td>565</td>
</tr>
<tr>
<td>Diagnosis of Breast Cancer</td>
<td>567</td>
</tr>
<tr>
<td>Examination</td>
<td>567</td>
</tr>
<tr>
<td>Imaging Techniques</td>
<td>567</td>
</tr>
<tr>
<td>Breast Biopsy</td>
<td>574</td>
</tr>
<tr>
<td>Breast Cancer Staging and Biomarkers</td>
<td>575</td>
</tr>
<tr>
<td>Breast Cancer Staging</td>
<td>575</td>
</tr>
<tr>
<td>Biomarkers</td>
<td>575</td>
</tr>
<tr>
<td>Overview of Breast Cancer Therapy</td>
<td>580</td>
</tr>
<tr>
<td>In Situ Breast Cancer (Stage 0)</td>
<td>580</td>
</tr>
<tr>
<td>Early Invasive Breast Cancer (Stage I, IIA, or IIB)</td>
<td>582</td>
</tr>
<tr>
<td>Advanced Local-Regional Breast Cancer (Stage IIIA or IIIB)</td>
<td>585</td>
</tr>
<tr>
<td>Internal Mammary Lymph Nodes</td>
<td>587</td>
</tr>
<tr>
<td>Distant Metastases (Stage IV)</td>
<td>587</td>
</tr>
<tr>
<td>Local-Regional Recurrence</td>
<td>587</td>
</tr>
<tr>
<td>Breast Cancer Prognosis</td>
<td>587</td>
</tr>
<tr>
<td>Surgical Techniques in Breast Cancer Therapy</td>
<td>588</td>
</tr>
<tr>
<td>Excisional Biopsy With Needle Localization</td>
<td>588</td>
</tr>
<tr>
<td>Sentinel Lymph Node Dissection</td>
<td>590</td>
</tr>
<tr>
<td>Breast Conservation</td>
<td>591</td>
</tr>
<tr>
<td>Mastectomy and Axillary Dissection</td>
<td>591</td>
</tr>
<tr>
<td>Modified Radical Mastectomy</td>
<td>592</td>
</tr>
<tr>
<td>Reconstruction of the Breast and Chest Wall</td>
<td>593</td>
</tr>
<tr>
<td>Nonsurgical Breast Cancer Therapies</td>
<td>594</td>
</tr>
<tr>
<td>Radiation Therapy</td>
<td>594</td>
</tr>
<tr>
<td>Chemotherapy Adjuvant</td>
<td>594</td>
</tr>
<tr>
<td>Antiestrogen Therapy</td>
<td>597</td>
</tr>
<tr>
<td>Ablative Endocrine Therapy</td>
<td>598</td>
</tr>
<tr>
<td>Anti-HER2 Therapy</td>
<td>598</td>
</tr>
<tr>
<td>Special Clinical Situations</td>
<td>599</td>
</tr>
<tr>
<td>Nipple Discharge</td>
<td>599</td>
</tr>
<tr>
<td>Axillary Lymph Node Metastases in the Setting of an Unknown Primary Cancer</td>
<td>600</td>
</tr>
<tr>
<td>Breast Cancer During Pregnancy</td>
<td>600</td>
</tr>
<tr>
<td>Male Breast Cancer</td>
<td>600</td>
</tr>
<tr>
<td>Phyllodes Tumors</td>
<td>600</td>
</tr>
<tr>
<td>Inflammatory Breast Carcinoma</td>
<td>601</td>
</tr>
<tr>
<td>Rare Breast Cancers</td>
<td>602</td>
</tr>
</tbody>
</table>
The breast receives its principal blood supply from perforating branches of the internal mammary artery, lateral branches of the posterior intercostal arteries, and branches from the axillary artery, including the highest thoracic, lateral thoracic, and pectoral branches of the thoracoacromial artery.

The axillary lymph nodes usually receive >75% of the lymph drainage from the breast, and the rest flows through the lymph vessels that accompany the perforating branches of the internal mammary artery and enters the parasternal (internal mammary) group of lymph nodes.

Breast development and function are initiated by a variety of hormonal stimuli, with the major trophic effects being modulated by estrogen, progesterone, and prolactin.

Benign breast disorders and diseases are related to the normal processes of reproductive life and to involution, and there is a spectrum of breast conditions that ranges from normal to disorder to disease (aberrations of normal development and involution classification).

To calculate breast cancer risk using the Gail model, a woman’s risk factors are translated into an overall risk score by multiplying her relative risks from several categories. This risk score is then compared with an adjusted population risk of breast cancer to determine the woman’s individual risk. This model is not appropriate for use in women with a known BRCA1 or BRCA2 mutation or women with lobular or ductal carcinoma in situ.

Routine use of screening mammography in women ≥50 years of age reduces mortality from breast cancer by 25%. Magnetic resonance imaging (MRI) screening is recommended in women with BRCA mutations and may be considered in women with a greater than 20% to 25% lifetime risk of developing breast cancer.

Core-needle biopsy is the preferred method for diagnosis of palpable or nonpalpable breast abnormalities.

When a diagnosis of breast cancer is made, the surgeon should determine the clinical stage, histologic characteristics, and appropriate biomarker levels before initiating local therapy.

Sentinel node dissection is the preferred method for staging of the regional lymph nodes in women with clinically node-negative invasive breast cancer. Axillary dissection may be avoided in women with one to two positive sentinel nodes who are treated with breast conserving surgery, whole breast radiation, and systemic therapy.

Local-regional and systemic therapy decisions for an individual patient with breast cancer are best made using a multidisciplinary treatment approach. The sequencing of therapies is dependent on patient and tumor related factors including breast cancer subtype.
Adjuvant Breast and Bowel Project (NSABP) conducted a randomized trial in the early 1970s to determine the impact of local and regional treatments on survival in operable breast cancer. In the B-04 trial, 1665 women were enrolled and stratified by clinical assessment of the axillary lymph nodes. The clinically node-negative women were randomized into three treatment groups: (a) Halsted radical mastectomy; (b) total mastectomy plus radiation therapy; and (c) total mastectomy alone. Clinically node-positive women were randomized to Halsted radical mastectomy or total mastectomy plus radiation therapy. This trial accrued patients between 1971 and 1974, an era that predated widespread availability of effective systemic therapy for breast cancer and therefore reflect survival associated with local-regional therapy alone. There were no differences in survival between the three groups of node-negative women or between the two groups of node-positive women. These overall survival equivalence patterns persisted at 25 years of follow-up.9

The next major advance in the surgical management of breast cancer was the development of breast conserving surgery. Breast conserving surgery and radiation treatment was first reported by Geoffrey Keynes of St Bartholomew’s Hospital, London in the British Medical Journal in 1937.10 Several decades later, the NSABP launched the B-06 trial, a phase 3 study that randomized 1851 patients to total mastectomy, lumpectomy alone, or lumpectomy with breast irradiation. The results showed no difference in disease-free, distant disease-free, and overall survival among the three groups; however, the omission of radiation therapy resulted in significantly higher rates of ipsilateral breast tumor recurrence in those who received lumpectomy alone.11 The B-06 trial excluded patients who had palpable axillary lymph nodes, and those patients randomized to breast conserving surgery had frozen sections performed. If on frozen section the margins were involved, the surgeon proceeded to perform a mastectomy, but the patient was included in the analysis as having had a breast conserving operation. Furthermore, in B-06, local in-breast recurrences were regarded as “nonevents” in terms of disease-free survival. Both the NSABP B-04 and B-06 trials were taken to refute the Halstedian concept that cancer spread throughout a region of the breast to lymphatics and then on to distant sites. Bernard Fisher proposed the “alternative hypothesis” that breast cancer was a systemic disease at diagnosis and that tumor cells had access to both the blood and lymphatic systems and that regional lymph nodes were a marker of systemic disease and not a barrier to the dissemination of cancer cells. He proposed that host factors were important in the development of metastasis and that variations in the local-regional approach to breast cancer treatment were not likely to substantially impact survival. This idea was dominant for a number of years but has been challenged by the Early Breast Cancer Trialists’ Collaborative Group (EBCTCG) overview analysis, which reported that “the avoidance of recurrence in a conserved breast . . . avoids about one breast cancer death over the next 15 years for every four such recurrences avoided,”12 indicating that not all breast cancer is a systemic disease at presentation.

During the 1970s, clinical trials were initiated to determine the value of systemic therapy in the postoperative setting as an adjuvant to surgery. The EBCTCG was established in 1985 to coordinate the meta-analysis of data from randomized clinical trials in order to examine the impact of adjuvant treatments for breast cancer on recurrence and mortality. The EBCTCG overview has demonstrated that anthracycline containing regimens are superior to cyclophosphamide, methotrexate, and 5-fluorouracil (CMF), and more recently, that the addition of a taxane to an anthracycline-based regimen reduces cancer mortality by one-third.13 The overview has also demonstrated that tamoxifen is of benefit only in patients with estrogen receptor (ER) positive breast cancer and that tamoxifen may decrease mortality from breast cancer by as much as 30%.14 Importantly, the EBCTCG data have shown that proportional reduction in risk was not significantly affected by standard clinical and pathologic factors such as tumor size, ER status, and nodal status.14 This underscores the importance of stratification of risk in determining adjuvant therapy decisions in order to minimize the toxicities of therapies in those unlikely to benefit, yet realize the substantial benefits gained in local-regional control and survival in those at higher risk.

Many early randomized clinical trials considered all patients similarly in terms of treatment viewing breast cancer as more of a homogeneous disease. Breast cancer has traditionally been defined by pathologic determinants using conventional light microscopy and basic histologic techniques. In the 1980s, immunohistochemistry allowed assessment of the expression of individual tumor markers (most commonly proteins) while DNA was initially assessed in terms of its ploidy status. Subsequently, breast cancer specimens have been interrogated at the level of the DNA by labeling genes of interest and allowing fluorescent dyes to quantify the abundance of a particular gene and comparing a large number of genes simultaneously in a single breast cancer specimen. Gene expression arrays have shown that breast cancers cluster according to their intrinsic gene expression patterns into at least five intrinsic subtypes and these intrinsic subtypes correlate with breast cancer outcomes.15 Breast cancers are now classified by molecular subtypes and these are being used for risk stratification and decision making in terms of local-regional and systemic therapies.

Currently, 50% of American women will consult a surgeon regarding breast disease, 25% will undergo breast biopsy for diagnosis of an abnormality, and 12% will develop some variant of breast cancer. Considerable progress has been made in the integration of surgery, radiation therapy, and systemic therapy to control local-regional disease, enhance survival, and improve the quality of life of breast cancer survivors. Surgeons are traditionally the first physician consulted for breast care, and it is critical for them to be well trained in all aspects of the breast from embryologic development, to growth and development, to benign and malignant disease processes. This will allow the greatest opportunity to achieve optimal outcomes for patients and their families.

**EMBRYOLOGY AND FUNCTIONAL ANATOMY OF THE BREAST**

**Embryology**

At the fifth or sixth week of fetal development, two ventral bands of thickened ectoderm (mammary ridges, milk lines) are evident in the embryo.16 In most mammals, paired breasts develop along these ridges, which extend from the base of the forelimb (future axilla) to the region of the hind limb (inguinal area). These ridges are not prominent in the human embryo and disappear after a short time, except for small portions that may persist in the pectoral region. Accessory breasts (polymastia) or accessory nipples (polythelia) may
occur along the milk line (Fig. 17-1) when normal regression fails. Each breast develops when an ingrowth of ectoderm forms a primary tissue bud in the mesenchyme. The primary bud, in turn, initiates the development of 15 to 20 secondary buds. Epithelial cords develop from the secondary buds and extend into the surrounding mesenchyme. Major (lactiferous) ducts develop, which open into a shallow mammary pit. During infancy, a proliferation of mesenchyme transforms the mammary pit into a nipple. If there is failure of a pit to elevate above skin level, an inverted nipple results. This congenital malformation occurs in 4% of infants. At birth, the breasts are identical in males and females, demonstrating only the presence of major ducts. Enlargement of the breast may be evident, and a secretion, historically referred to as witch’s milk, may be produced. These transitory events occur in response to maternal hormones that cross the placenta.

The breast remains undeveloped in the female until puberty, when it enlarges in response to ovarian estrogen and progesterone, which initiate proliferation of the epithelial and connective tissue elements. However, the breasts remain incompletely developed until pregnancy occurs. Absence of the breast (amastia) is rare and results from an arrest in mammary ridge development that occurs during the sixth fetal week. Poland’s syndrome consists of hypoplasia or complete absence of the breast, costal cartilage and rib defects, hypoplasia of the subcutaneous tissues of the chest wall, and brachysyndactyly. Breast hypoplasia also may be iatrogenically induced before puberty by trauma, infection, or radiation therapy. Symmastia is a rare anomaly recognized as webbing between the breasts across the midline. Accessory nipples (polythelia) occur in <1% of infants and may be associated with abnormalities of the urinary and cardiovascular systems. Supernumerary breasts may occur in any configuration along the mammary milk line but most frequently occur between the normal nipple location and the symphysis pubis. Turner’s syndrome (ovarian agenesis and dysgenesis) and Fleischer’s syndrome (displacement of the nipples and bilateral renal hypoplasia) may have polymastia as a component. Accessory axillary breast tissue is uncommon and usually is bilateral.

**Functional Anatomy**

The breast is composed of 15 to 20 lobes (Fig. 17-2), which are each composed of several lobules.17 Fibrous bands of connective tissue travel through the breast (Cooper’s suspensory ligaments), insert perpendicularly into the dermis, and provide structural support. The mature female breast extends from the level of the second or third rib to the inframammary fold at the sixth or seventh rib. It extends transversely from the lateral border of the sternum to the anterior axillary line. The deep or posterior surface of the breast rests on the fascia of the pectoralis major, serratus anterior, and external oblique abdominal muscles, and the upper extent of the rectus sheath. The retro-mammary bursa may be identified on the posterior aspect of the breast between the investing fascia of the breast and the fascia of the pectoralis major muscles. The axillary tail of Spence extends laterally across the anterior axillary fold. The upper outer quadrant of the breast contains a greater volume of tissue than do the other quadrants. The breast has a protuberant conical form. The base of the cone is roughly circular, measuring 10 to 12 cm in diameter. Considerable variations in the size, contour, and density of the breast are evident among individuals. The nulliparous breast has a hemispheric configuration with distinct flattening above the nipple. With the hormonal stimulation that accompanies pregnancy and lactation, the breast becomes larger and increases in volume and density, whereas with senescence, it assumes a flattened, flaccid, and more pendulous configuration with decreased volume.

**Nipple-Areola Complex.** The epidermis of the nipple-areola complex is pigmented and is variably corrugated. During puberty, the pigment becomes darker and the nipple assumes an elevated configuration. Throughout pregnancy, the areola
enlarges and pigmentation is further enhanced. The areola contains sebaceous glands, sweat glands, and accessory glands, which produce small elevations on the surface of the areola (Montgomery’s tubercles). Smooth muscle bundle fibers, which lie circumferentially in the dense connective tissue and longitudinally along the major ducts, extend upward into the nipple, where they are responsible for the nipple erection that occurs with various sensory stimuli. The dermal papilla at the tip of the nipple contains numerous sensory nerve endings and Meissner’s corpuscles. This rich sensory innervation is of functional importance because the sucking of the infant initiates a chain of neurohumoral events that results in milk letdown.

**Inactive and Active Breast.** Each lobe of the breast terminates in a major (lactiferous) duct (2–4 mm in diameter), which opens through a constricted orifice (0.4–0.7 mm in diameter) into the ampulla of the nipple (see Fig. 17-2). Immediately below the nipple-areola complex, each major duct has a dilated portion (lactiferous sinus), which is lined with stratified squamous epithelium. Major ducts are lined with two layers of cuboidal cells, whereas minor ducts are lined with a single layer of columnar or cuboidal cells. Myoepithelial cells of ectodermal origin reside between the epithelial cells in the basal lamina and contain myofibrils. In the inactive breast, the epithelium is sparse and consists primarily of ductal epithelium (Fig. 17-3). In the early phase of the menstrual cycle, minor ducts are cord-like with small lumina. With estrogen stimulation at the time of ovulation, alveolar epithelium increases in height, duct lumina become more prominent, and some secretions accumulate. When the hormonal stimulation decreases, the alveolar epithelium regresses.

With pregnancy, the breast undergoes proliferative and developmental maturation. As the breast enlarges in response to hormonal stimulation, lymphocytes, plasma cells, and eosinophils accumulate within the connective tissues. The minor ducts branch and alveoli develop. Development of the alveoli is asymmetric, and variations in the degree of development may occur within a single lobule (Fig. 17-4). With parturition, enlargement of the breasts occurs via hypertrophy of alveolar epithelium and accumulation of secretory products in the lumina of the minor ducts. Alveolar epithelium contains abundant endoplasmic reticulum, large mitochondria, Golgi complexes, and dense lysosomes. Two distinct substances are produced by the alveolar epithelium: (a) the protein component of milk, which is synthesized in the endoplasmic reticulum (merocrine secretion); and (b) the lipid component of milk (apocrine secretion), which forms as free lipid droplets in the cytoplasm. Milk released in the first few days after parturition is called *colostrum* and has low lipid content but contains considerable quantities of antibodies. The lymphocytes and plasma cells that accumulate within the connective tissues of the breast are the source of the antibody component. With subsequent reduction in the number of these cells, the production of colostrum decreases and lipid-rich milk is released.

**Blood Supply, Innervation, and Lymphatics.** The breast receives its principal blood supply from: (a) perforating branches of the internal mammary artery; (b) lateral branches of the posterior intercostal arteries; and (c) branches from the axillary artery, including the highest thoracic, lateral thoracic, and pectoral branches of the thoracoacromial artery (Fig. 17-5). The second, third, and fourth anterior intercostal perforators and branches of the internal mammary artery arborize in the breast as the medial mammary arteries. The lateral thoracic artery gives off branches to the serratus anterior, pectoralis major and pectoralis minor, and subscapularis muscles. It also gives rise to lateral mammary branches. The veins of the breast and chest wall follow the course of the arteries, with venous drainage being toward the axilla. The three principal groups of veins are: (a) perforating branches of the internal thoracic vein, (b) perforating branches of the posterior intercostal veins, and (c) tributaries of the axillary vein. Batson’s vertebral venous plexus, which invests the vertebral columns and extends from the base of the skull to the sacrum, may provide a route for breast cancer metastases to the vertebrae, skull, pelvic bones, and central nervous system. Lymph vessels generally parallel the course of blood vessels.
Lateral cutaneous branches of the third through sixth intercostal nerves provide sensory innervation of the breast (lateral mammary branches) and of the anterolateral chest wall. These branches exit the intercostal spaces between slips of the serratus anterior muscle. Cutaneous branches that arise from the cervical plexus, specifically the anterior branches of the supraclavicular nerve, supply a limited area of skin over the upper portion of the breast. The intercostobrachial nerve is the lateral cutaneous branch of the second intercostal nerve and may be visualized during surgical dissection of the axilla. Resection of the intercostobrachial nerve causes loss of sensation over the medial aspect of the upper arm.

The boundaries for lymph drainage of the axilla are not well demarcated, and there is considerable variation in the position of the axillary lymph nodes. The six axillary lymph node groups recognized by surgeons (Figs. 17-6 and 17-7) are: (a) the axillary vein group (lateral), which consists of four to six lymph nodes that lie medial or posterior to the vein and receive most of the lymph drainage from the upper extremity; (b) the external mammary group (anterior or pectoral group), which consists of five to six lymph nodes that lie along the lower border of the pectoralis minor muscle contiguous with the lateral thoracic vessels and receive most of the lymph drainage from the lateral aspect of the breast; (c) the scapular group (posterior or subscapular), which consists of five to seven lymph nodes that lie along the posterior wall of the axilla at the lateral border of the scapula contiguous with the subscapular vessels and receive lymph drainage principally from the lower posterior neck, the posterior trunk, and the posterior shoulder; (d) the central group, which consists of three or four sets of lymph nodes that are embedded in the fat of the axilla lying immediately posterior to the pectoralis minor muscle and receive lymph drainage both from the axillary vein, external mammary, and scapular groups of lymph nodes, and directly from the breast; (e) the subclavicular group (apical), which consists of six to twelve sets of lymph nodes that lie posterior and superior to the upper border of the pectoralis minor muscle and receive lymph drainage from all of the other groups of axillary lymph nodes; and (f) the interpectoral group (Rotter’s lymph nodes), which consists of one to four lymph nodes that are interposed between the pectoralis major and pectoralis minor muscles and receive lymph drainage directly from the breast. The lymph fluid that passes
through the interpectoral group of lymph nodes passes directly into the central and subclavicular groups.

As indicated in Fig. 17-7, the lymph node groups are assigned levels according to their anatomic relationship to the pectoralis minor muscle. Lymph nodes located lateral to or below the lower border of the pectoralis minor muscle are referred to as level I lymph nodes, which include the axillary vein, external mammary, and scapular groups. Lymph nodes located superficial or deep to the pectoralis minor muscle are referred to as level II lymph nodes, which include the central and interpectoral groups. Lymph nodes located medial to or above the upper border of the pectoralis minor muscle are referred to as level III lymph nodes, which consist of the subclavicular group. The plexus of lymph vessels in the breast arises in the interlobular connective tissue and in the walls of the lactiferous ducts and communicates with the subareolar plexus of lymph vessels. Efferent lymph vessels from the breast pass around the lateral edge of the pectoralis major muscle and pierce the clavipectoral fascia, ending in the external mammary (anterior, pectoral) group of lymph nodes. Some lymph vessels may travel directly to the subscapular (posterior, scapular) group of lymph nodes. From the upper part of the breast, a few lymph vessels pass directly to the subclavicular (apical) group of lymph nodes. The axillary lymph nodes usually receive >75% of the lymph drainage from the breast. The rest is derived primarily from the medial aspect of the breast, flows through the lymph vessels that accompany the perforating branches of the internal mammary artery, and enters the parasternal (internal mammary) group of lymph nodes.

PHYSIOLOGY OF THE BREAST

Breast Development and Function

Breast development and function are initiated by a variety of hormonal stimuli, including estrogen, progesterone, prolactin, oxytocin, thyroid hormone, cortisol, and growth hormone. Estrogen, progesterone, and prolactin especially have profound trophic effects that are essential to normal breast development and function. Estrogen initiates ductal development, whereas progesterone is responsible for differentiation of epithelium and for lobular development. Prolactin is the primary hormonal stimulus for lactogenesis in late pregnancy and the postpartum period. It upregulates hormone receptors and stimulates epithelial development. Fig. 17-8 depicts the secretion of neurotrrophic hormones from the hypothalamus, which is responsible for regulation of the secretion of the hormones that affect the breast tissues. The gonadotropins luteinizing hormone (LH) and follicle-stimulating hormone (FSH) regulate the release of estrogen and progesterone from the ovaries. In turn, the release of LH and FSH from the basophilic cells of the anterior pituitary is regulated by the secretion of gonadotropin-releasing hormone (GnRH) from the hypothalamus. Positive and negative feedback effects of circulating estrogen and progesterone regulate the secretion of LH, FSH, and GnRH. These hormones are responsible for the development, function, and maintenance of breast tissues (Fig. 17-9A). In the female neonate, circulating estrogen and progesterone levels decrease after birth and remain low throughout childhood because of the sensitivity of

Figure 17-8. Overview of the neuroendocrine control of breast development and function. ADH = antidiuretic hormone; CRF = corticotropin-releasing factor; GRF = growth hormone releasing factor; LH-RH = luteinizing hormone–releasing hormone; Oxy = oxytocin; TRH = thyrotropin-releasing hormone. (Reproduced with permission from Bland KI, Copeland EMI: The Breast: Comprehensive Management of Benign and Malignant Diseases, 4th ed. Philadelphia, PA: Elsevier/Saunders; 2009.)
the hypothalamic-pituitary axis to negative feedback from these hormones. With the onset of puberty, there is a decrease in the sensitivity of the hypothalamic-pituitary axis to negative feedback and an increase in its sensitivity to positive feedback from estrogen. These physiologic events initiate an increase in GnRH, FSH, and LH secretion and ultimately an increase in estrogen and progesterone secretion by the ovaries, leading to establishment of the menstrual cycle. At the beginning of the menstrual cycle, there is an increase in the size and density of the breasts, which is followed by engorgement of the breast tissues and epithelial proliferation. With the onset of menstruation, the breast engorgement subsides and epithelial proliferation decreases.

**Pregnancy, Lactation, and Senescence**

A dramatic increase in circulating ovarian and placental estrogens and progestins is evident during pregnancy, which initiates striking alterations in the form and substance of the breast (see Fig. 17-9B).\(^{17-19}\) The breast enlarges as the ductal and lobular epithelium proliferates, the areolar skin darkens, and the accessory areolar glands (Montgomery’s glands) become prominent. In the first and second trimesters, the minor ducts branch and develop. During the third trimester, fat droplets accumulate in the alveolar epithelium, and colostrum fills the alveolar and ductal spaces. In late pregnancy, prolactin stimulates the synthesis of milk fats and proteins.

After delivery of the placenta, circulating progesterone and estrogen levels decrease, permitting full expression of the lactogenic action of prolactin. Milk production and release are controlled by neural reflex arcs that originate in nerve endings of the nipple-areola complex. Maintenance of lactation requires regular stimulation of these neural reflexes, which results in prolactin secretion and milk letdown. Oxytocin release results from the auditory, visual, and olfactory stimuli associated with nursing. Oxytocin initiates contraction of the myoepithelial cells, which results in compression of alveoli and expulsion of milk into the lactiferous sinuses. After weaning of the infant, prolactin and oxytocin release decreases. Dormant milk causes increased pressure within the ducts and alveoli, which results in atrophy of the epithelium (Fig. 17-9C). With menopause, there is a decrease in the secretion of estrogen and progesterone by

![Figure 17-9. The breast at different physiologic stages. The central column contains three-dimensional depictions of microscopic structures. A. Adolescence. B. Pregnancy. C. Lactation. D. Senescence.](image-url)
Gynecomastia

Gynecomastia refers to an enlarged breast in the male. Physiologic gynecomastia usually occurs during three phases of life: the neonatal period, adolescence, and senescence. Common to each of these phases is an excess of circulating estrogens in relation to circulating testosterone. Neonatal gynecomastia is caused by the action of placental estrogens on neonatal breast tissues, whereas in adolescence, there is an excess of estradiol relative to testosterone, and with senescence, the circulating testosterone level falls, which results in relative hyperestrinism. In gynecomastia, the ductal structures of the male breast enlarge, elongate, and branch with a concomitant increase in epithelium. During puberty, the condition often is unilateral and typically occurs between ages 12 and 15 years. In contrast, senescent gynecomastia is usually bilateral. In the nonobese male, breast tissue measuring at least 2 cm in diameter must be present before a diagnosis of gynecomastia may be made. Mammography and ultrasonography are used to differentiate breast tissues. Dominant masses or areas of firmness, irregularity, and asymmetry suggest the possibility of a breast cancer, particularly in the older male. Gynecomastia generally does not predispose the male breast to cancer. However, the hypoandrogeic state of Klinefelter’s syndrome (XXY), in which gynecomastia is usually evident, is associated with an increased risk of breast cancer. Gynecomastia is graded based on the degree of breast enlargement, the position of the nipple with reference to the inframammary fold, and the degree of breast ptosis and skin redundancy: Grade I—mild breast enlargement without skin redundancy; Grade II—moderate breast enlargement without skin redundancy; Grade IIb—moderate breast enlargement with skin redundancy; and Grade III—marked breast enlargement with skin redundancy and ptosis.

Table 17-1 identifies the pathophysiologic mechanisms that may initiate gynecomastia: estrogen excess states; androgen deficiency states; pharmacologic causes; and idiopathic causes. Estrogen excess results from an increase in the secretion of estradiol by the testicles or by nontesticular tumors, nutritional alterations such as protein and fat deprivation, endocrine disorders (hyperthyroidism, hypothyroidism), and hepatic disease (nonalcoholic and alcoholic cirrhosis). Refeeding gynecomastia is related to the resumption of pituitary gonadotropin secretion after pituitary shutdown. Androgen deficiency may initiate gynecomastia. Concurrently occurring with decreased circulating testosterone levels is an elevated level of circulating testosterone-binding globulin, which results in a reduction of free testosterone. This senescent gynecomastia usually occurs in men age 50 to 70 years. Hypoandrogenic states can be from primary testicular failure or secondary testicular failure. Klinefelter’s syndrome (XXY) is an example of primary testicular failure that is manifested by gynecomastia, hypogonadotropic hypogonadism, and azoospermia. Secondary testicular failure may result from trauma, orchitis, and cryptorchidism. Renal failure, regardless of cause, also may initiate gynecomastia.

Pharmacologic causes of gynecomastia include drugs with estrogenic activity (digitalis, estrogens, anabolic steroids, marijuana) or drugs that enhance estrogen synthesis (human chorionic gonadotropin). Drugs that inhibit the action or synthesis of testosterone (cimetidine, ketoconazole, phenytoin, spironolactone, antineoplastic agents, diazepam) also have been implicated. Drugs such as reserpine, theophylline, verapamil, tricyclic antidepressants, and furosemide induce gynecomastia through idiopathic mechanisms.

When gynecomastia is caused by androgen deficiency, then testosterone administration may cause regression. When it is caused by medications, then these are discontinued if possible. When endocrine defects are responsible, then these receive specific therapy. As soon as gynecomastia is progressive and does not respond to other treatments, surgical therapy is considered. Techniques include local excision, liposuction or subcutaneous mastectomy. Attempts to reverse gynecomastia with danazol have been successful, but the androgenic side effects of the drug are considerable.
Infections in the postpartum period remain proportionately the most common time for breast infections to occur. Infections of the breast unrelated to lactation are proportionately less common, however, are still a relatively common presentation to breast specialists. The latter are classified as intrinsic (secondary to abnormalities in the breast) or extrinsic (secondary to an infection in an adjacent structure, e.g., skin, thoracic cavity) the most common being probably periductal mastitis and infected sebaceous cysts, respectively.

**Bacterial Infection**

*Staphylococcus aureus* and *Streptococcus* species are the organisms most frequently recovered from nipple discharge from an infected breast. Typically breast abscesses are seen in staphylococcal infections and present with point tenderness, erythema, and hyperthermia. When these abscesses are related to lactation they usually occur within the first few weeks of breastfeeding. If there is progression of a staphylococcal infection, this may result in subcutaneous, subareolar, interlobular (periductal), and retromammary abscesses (unicentric or multicentric). Previously almost all breast abscesses were treated by operative incision and drainage, but now the initial approach is antibiotics and repeated aspiration of the abscess, usually ultrasound-guided aspiration. Operative drainage is now reserved for those cases that do not resolve with repeated aspiration and antibiotic therapy or cases in which there is some other indication for incision and drainage (e.g., thinning or necrosis of the overlying skin). Preoperative ultrasonography is effective in delineating the required extent of the drainage procedure. While staphylococcal infections tend to be more localized and may be situated deep in the breast tissues, streptococcal infections usually present with diffuse superficial involvement. They are treated with local wound care, including application of warm compresses, and the administration of IV antibiotics (penicillins or cephalosporins). Breast infections may be chronic, possibly with recurrent abscess formation. In this situation, cultures are performed to identify acid-fast bacilli, anaerobic and aerobic bacteria, and fungi. Uncommon organisms may be encountered, and long-term antibiotic therapy may be required.

Biopsy of the abscess cavity wall should be considered at the time of incision and drainage to rule out underlying breast cancer in patients where antibiotics and drainage have been ineffective.

Nowadays hospital-acquired puerperal infections of the breast are much less common, but nursing women who present with milk stasis or noninfectious inflammation may still develop this problem. Epidemic puerperal mastitis is initiated by highly virulent strains of methicillin-resistant *S. aureus* that are transmitted via the suckling neonate and may result in substantial morbidity and occasional mortality. Purulent fluid may be expressed from the nipple. In this circumstance, breastfeeding is stopped, antibiotics are started, and surgical therapy is initiated. *Nonepidemic* (sporadic) puerperal mastitis refers to involvement of the interlobular connective tissue of the breast by an infectious process. The patient develops nipple fissuring and milk stasis, which initiates a retrograde bacterial infection. Emptying of the breast using breast suction pumps shortens the duration of symptoms and reduces the incidence of recurrences. The addition of antibiotic therapy results in a satisfactory outcome in >95% of cases.

**Mycotic Infections**

Fungal infections of the breast are rare and usually involve blastomycosis or sporotrichosis. Intraoral fungi that are inoculated into the breast tissue by the suckling infant initiate these infections, which present as mammary abscesses in close proximity to the nipple-areola complex. Pus mixed with blood may be expressed from skin tracts. Antifungal agents can be administered for the treatment of systemic (noncutaneous) infections. This therapy generally eliminates the necessity of surgical intervention, but occasionally drainage of an abscess, or even partial mastectomy, may be necessary to eradicate a persistent fungal infection. *Candida albicans* affecting the skin of the breast presents as erythematous, scaly lesions of the inframammary or axillary folds. Scrapings from the lesions demonstrate fungal elements (filaments and binding cells). Therapy involves the removal of predisposing factors such as maceration and the topical application of nystatin.

**Hidradenitis Suppurativa**

Hidradenitis suppurativa of the nipple-areola complex or axilla is a chronic inflammatory condition that originates within the accessory areolar glands of Montgomery or within the axillary sebaceous glands. Women with chronic acne are predisposed to developing hidradenitis. When located in and about the nipple-areola complex, this disease may mimic other chronic inflammatory states, Paget’s disease of the nipple, or invasive breast cancer. Involvement of the axillary skin is often multifocal and contiguous. Antibiotic therapy with incision and drainage of fluctuant areas is appropriate treatment. Excision of the involved areas may be required. Large areas of skin loss may necessitate coverage with advancement flaps or split-thickness skin grafts.

**Mondor’s Disease**

Mondor’s disease is a variant of thrombophlebitis that involves the superficial veins of the anterior chest wall and breast. In 1939, Mondor described the condition as “string phlebitis,” a thrombosed vein presenting as a tender, cord-like structure. Frequently involved veins include the lateral thoracic vein, the thoracoepigastric vein, and, less commonly, the superficial epigastric vein. Typically, a woman presents with acute pain in the lateral aspect of the breast or the anterior chest wall. A tender, firm cord is found to follow the distribution of one of the major superficial veins. Rarely, the presentation is bilateral, and most women have no evidence of thrombophlebitis in other anatomic sites. This benign, self-limited disorder is not indicative of a cancer. When the diagnosis is uncertain, or when a mass is present near the tender cord, biopsy is indicated. Therapy for Mondor’s disease includes the liberal use of anti-inflammatory medications and application of warm compresses along the symptomatic vein. The process usually resolves within 4 to 6 weeks. When symptoms persist or are refractory to therapy, excision of the involved vein segment may be considered.
COMMON BENIGN DISORDERS AND DISEASES OF THE BREAST

Benign breast disorders and diseases encompass a wide range of clinical and pathologic entities. Surgeons require an in-depth understanding of benign breast disorders and diseases so that clear explanations may be given to affected women, appropriate treatment is instituted, and unnecessary long-term follow up is avoided.

Aberrations of Normal Development and Involution

The basic principles underlying the aberrations of normal development and involution (ANDI) classification of benign breast conditions are the following: (a) benign breast disorders and diseases are related to the normal processes of reproductive life and to involution; (b) there is a spectrum of breast conditions that ranges from normal to disorder to disease; and (c) the ANDI classification encompasses all aspects of the breast condition, including pathogenesis and the degree of abnormality. The horizontal component of Table 17-2 defines ANDI along a spectrum from normal, to mild abnormality (disorder), to severe abnormality (disease). The vertical component indicates the period during which the condition develops.

Early Reproductive Years. Fibroadenomas are seen and present symptomatically predominantly in younger women age 15 to 25 years (Fig. 17-10). Fibroadenomas usually grow to 1 or 2 cm in diameter and then are stable but may grow to a larger size. Small fibroadenomas (≤1 cm in size) are considered normal, whereas larger fibroadenomas (≤3 cm) are disorders, and giant fibroadenomas (>3 cm) are disease. Similarly, multiple fibroadenomas (more than five lesions in one breast) are very uncommon and are considered disease. It is noted that with the introduction of mammographic screening, asymptomatic fibroadenomas are sometimes found in an older screened population. The precise etiology of adolescent breast hypertrophy is unknown. A spectrum of changes from limited to massive stromal hyperplasia (gigantomastia) is seen. Nipple inversion is a disorder of development of the major ducts, which prevents normal protrusion of the nipple. Mammary duct fistulas arise when nipple inversion predisposes to major duct obstruction, leading to recurrent subareolar abscess and mammary duct fistula.

<table>
<thead>
<tr>
<th>Table 17-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANDI classification of benign breast disorders</td>
</tr>
<tr>
<td><strong>NORMAL</strong></td>
</tr>
<tr>
<td><strong>Early reproductive years</strong> (age 15–25 y)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Later reproductive years</strong> (age 25–40 y)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Involution</strong> (age 35–55 y)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

ANDI = aberrations of normal development and involution.


Figure 17-10. Fibroadenoma (40x). These benign tumors are typically well circumscribed and comprised of both stromal and glandular elements. (Used with permission from Dr. Sindhu Menon, Consultant Histopathologist and Dr. Rahul Deb, Consultant Histopathologist and Lead Breast Pathologist, Royal Derby Hospital, Derby, UK.)
Later Reproductive Years. Cyclic mastalgia and nodularity usually are associated with premenstrual enlargement of the breast and are regarded as normal. Cyclic pronounced mastalgia and severe painful nodularity are viewed differently than are physiologic discomfort and lumpiness. Painful nodularity that persists for >1 week of the menstrual cycle is considered a disorder. In epithelial hyperplasia of pregnancy, papillary projections sometimes give rise to bilateral bloody nipple discharge.

Involutions. Involutions of lobular epithelium is dependent on the specialized stroma around it. However, an integrated involutions of breast stroma and epithelium is not always seen, and disorders of the process are common. When the stroma involutes too quickly, alveoli remain and form microcysts, which are precursors of macrocysts. The macrocysts are common, often subclinical, and do not require specific treatment. Sclerosing adenosis is considered a disorder of both the proliferative and the involutions phases of the breast cycle. Duct ectasia (dilated ducts) and periductal mastitis are other important components of the ANDI classification. Periductal fibrosis is a sequela of periductal mastitis and may result in nipple retraction. About 60% of women ≥70 years of age exhibit some degree of epithelial hyperplasia (Fig. 17-11). Atypical proliferative diseases include ductal and lobular hyperplasia, both of which display some features of carcinoma in situ. Women with atypical ductal or lobular hyperplasia have a fourfold increase in breast cancer risk (Table 17-3).

Pathology of Nonproliferative Disorders

Of paramount importance for the optimal management of benign breast disorders and diseases is the histologic differentiation of benign, atypical, and malignant changes. Determining the clinical significance of these changes is a problem that is compounded by inconsistent nomenclature. The classification system originally developed by Page separates the various types of benign breast disorders and diseases into three clinically relevant groups: nonproliferative disorders, proliferative disorders without atypia, and proliferative disorders with atypia (Table 17-4). Nonproliferative disorders of the breast account for 70% of benign breast conditions and carry no increased risk.

### Table 17-3

<table>
<thead>
<tr>
<th>ABNORMALITY</th>
<th>RELATIVE RISK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonproliferative lesions of the breast</td>
<td>No increased risk</td>
</tr>
<tr>
<td>Sclerosing adenosis</td>
<td>No increased risk</td>
</tr>
<tr>
<td>Intraductal papilloma</td>
<td>No increased risk</td>
</tr>
<tr>
<td>Florid hyperplasia</td>
<td>1.5 to 2-fold</td>
</tr>
<tr>
<td>Atypical lobular hyperplasia</td>
<td>4-fold</td>
</tr>
<tr>
<td>Atypical ductal hyperplasia</td>
<td>4-fold</td>
</tr>
<tr>
<td>Ductal involvement by cells of atypical ductal hyperplasia</td>
<td>7-fold</td>
</tr>
<tr>
<td>Lobular carcinoma in situ</td>
<td>10-fold</td>
</tr>
<tr>
<td>Ductal carcinoma in situ</td>
<td>10-fold</td>
</tr>
</tbody>
</table>


### Table 17-4

| Classification of benign breast disorders |
|------------------------------------------|-------------------------------|
| Nonproliferative disorders of the breast | Cysts and apocrine metaplasia |
|                                          | Duct ectasia                 |
|                                          | Mild ductal epithelial hyperplasia |
|                                          | Calcifications               |
|                                          | Fibroadenoma and related lesions |
| Proliferative breast disorders without atypia | Sclerosing adenosis |
|                                          | Radial and complex sclerosing lesions |
|                                          | Ductal epithelial hyperplasia |
|                                          | Intraductal papillomas       |
| Atypical proliferative lesions            | Atypical lobular hyperplasia |
|                                          | Atypical ductal hyperplasia  |

for the development of breast cancer. This category includes cysts, duct ectasia, periductal mastitis, calcifications, fibroadenomas, and related disorders.

Breast macrocysts are an involutional disorder, have a high frequency of occurrence, and are often multiple. Duct ectasia is a clinical syndrome characterized by dilated subareolar ducts that are palpable and often associated with thick nipple discharge. Haagensen regarded duct ectasia as a primary event that led to stagnation of secretions, epithelial ulceration, and leakage of duct secretions (containing chemically irritating fatty acids) into periductal tissue. This sequence was thought to produce a local inflammatory process with periductal fibrosis and subsequent nipple retraction. An alternative theory considers periductal mastitis to be the primary process, which leads to weakening of the ducts and secondary dilatation. It is possible that both processes occur and together explain the wide spectrum of problems seen, which include nipple discharge, nipple retraction, inflammatory masses, and abscesses.

Calcium deposits are frequently encountered in the breast. Most are benign and are caused by cellular secretions and debris or by trauma and inflammation. Calcifications that are associated with cancer include microcalcifications, which vary in shape and density and are <0.5 mm in size, and fine, linear calcifications, which may show branching. Fibroadenomas have abundant stroma with histologically normal cellular elements. They show hormonal dependence similar to that of normal breast lobules in that they lactate during pregnancy and involute in the postmenopausal period. Adenomas of the breast are well circumscribed and are composed of benign epithelium with sparse stroma, which is the histologic feature that differentiates them from fibroadenomas. They may be divided into tubular adenomas and lactating adenomas. Tubular adenomas are seen in young nonpregnant women, whereas lactating adenomas are seen during pregnancy or during the postpartum period. Hamartomas are discrete breast tumors that are usually 2 to 4 cm in diameter, firm, and sharply circumscribed. Adenolipomas consist of sharply circumscribed nodules of fatty tissue that contain normal breast lobules and ducts.

**Fibrocystic Disease.** The term fibrocystic disease is nonspecific. Too frequently, it is used as a diagnostic term to describe symptoms, to rationalize the need for breast biopsy, and to explain biopsy results. Synonyms include fibrocystic changes, cystic mastopathy, chronic cystic disease, chronic cystic mastitis, Schimmelbusch’s disease, mazoplasia, Cooper’s disease, Reclus’ disease, and fibroadenomatosis. Fibrocystic disease refers to a spectrum of histopathologic changes that are best diagnosed and treated specifically.

**Pathology of Proliferative Disorders Without Atypia**

Proliferative breast disorders without atypia include sclerosing adenosis, radial scars, complex sclerosing lesions, ductal epithelial hyperplasia, and intraductal papillomas. Sclerosing adenosis is prevalent during the childbearing and perimenopausal years and has no malignant potential. Histologic changes are both proliferative (ductal proliferation) and involutorial (stromal fibrosis, epithelial regression). Sclerosing adenosis is characterized by distorted breast lobules and usually occurs in the context of multiple microcysts, but occasionally presents as a palpable mass. Benign calcifications are often associated with this disorder. Sclerosing adenosis can be managed by observation as long as the imaging features and pathologic findings are concordant. Central sclerosis and various degrees of epithelial proliferation, apocrine metaplasia, and papilloma formation characterize radial scars and complex sclerosing lesions of the breast. Lesions up to 1 cm in diameter are called radial scars, whereas larger lesions are called complex sclerosing lesions. Radial scars originate at sites of terminal duct branching where the characteristic histologic changes radiate from a central area of fibrosis. All of the histologic features of a radial scar are seen in the larger complex sclerosing lesions, but there is a greater disturbance of structure with papilloma formation, apocrine metaplasia, and occasionally sclerosing adenosis. Distinguishing between a radial scar and invasive breast carcinoma can be challenging based on core-needle biopsy sampling. Often the imaging features of a radial scar (which can be quite similar to an invasive cancer) will dictate the need for either a vacuum-assisted biopsy or surgical excision in order to exclude the possibility of carcinoma.

Mild ductal hyperplasia is characterized by the presence of three or four cell layers above the basement membrane. Moderate ductal hyperplasia is characterized by the presence of five or more cell layers above the basement membrane. Florid ductal epithelial hyperplasia occupies at least 70% of a minor duct lumen. It is found in >20% of breast tissue specimens, is either solid or papillary, and is associated with an increased cancer risk (see Table 17-3). Intraductal papillomas arise in the major ducts, usually in premenopausal women. They generally are <0.5 cm in diameter but may be as large as 5 cm. A common presenting symptom is nipple discharge, which may be serous or bloody. Grossly, intraductal papillomas are pinkish tan, friable, and usually attached to the wall of the involved duct by a stalk. They rarely undergo malignant transformation, and their presence does not increase a woman’s risk of developing breast cancer (unless accompanied by atypia). However, multiple intraductal papillomas, which occur in younger women and are less frequently associated with nipple discharge, are susceptible to malignant transformation.

**Pathology of Atypical Proliferative Diseases**

The atypical proliferative diseases have some of the features of carcinoma in situ but either lack a major defining feature of carcinoma in situ or have the features in less than fully developed form. Atypical ductal hyperplasia (ADH) appears similar to low grade ductal carcinoma in situ (DCIS) histologically and is composed of monotonous round, cuboidal, or polygonal cells enclosed by basement membrane with rare mitoses. A lesion will be considered to be ADH if it is up to 2 or 3 mm in size but would be called DCIS if it is larger than 3 mm. The diagnosis can be difficult to establish with core-needle biopsy specimen alone and many cases will require excisional biopsy specimen for classification. Individuals with a diagnosis of ADH are at increased risk for development of breast cancer and should be counseled appropriately regarding risk reduction strategies.

In 1978, Haagensen et al described lobular neoplasia, a spectrum of disorders ranging from atypical lobular hyperplasia to lobular carcinoma in situ (LCIS). Atypical lobular hyperplasia (ALH) results in minimal distention of lobular units with cells that are similar to those seen in LCIS. The diagnosis of LCIS is made when small monomorphic cells that distend the terminal ductal lobular unit are noted. In cases of LCIS, the acini are full and distended while the overall lobular architecture is maintained (Fig. 17-12). Classic LCIS is not associated with a specific mammographic or palpable abnormality but is
SPECIFIC CONSIDERATIONS

PART II

Figure 17-12. Lobular carcinoma in situ (100x). There are small monomorphous cells that distend the terminal duct lobular unit, without necrosis or mitoses. (Used with permission from Dr. Sindhu Menon, Consultant Histopathologist and Dr. Rahul Deb, Consultant Histopathologist and Lead Breast Pathologist, Royal Derby Hospital, Derby, UK.)

an incident finding noted on breast biopsy. There is a variant of LCIS that has been termed pleomorphic LCIS. In the case of pleomorphic LCIS, there can be calcifications or other suspicious mammographic changes that dictate the need for biopsy. Classic LCIS is not treated with excision as the patient is at risk for developing invasive breast cancer in either breast and therefore the patient is counseled regarding appropriate risk reduction strategies. Pleomorphic LCIS can be difficult to distinguish from high-grade DCIS and there are some proponents who have suggested that patients with pleomorphic LCIS be managed similar to those with DCIS with attention to margins and consideration for radiation therapy in the setting of breast conserving treatment. The use of immunohistochemical staining for E-cadherin can help to discriminate between LCIS and DCIS. In lobular neoplasias, such as ALH and LCIS, there is a lack of E-cadherin expression, whereas the majority of ductal lesions will demonstrate E-cadherin reactivity.

Treatment of Selected Benign Breast Disorders and Diseases

Cysts. Because needle biopsy of breast masses may produce artifacts that make mammography assessment more difficult, many multidisciplinary teams prefer to image breast masses before performing either fine-needle aspiration or core-needle biopsy.36,37 In practice, however, the first investigation of palpable breast masses may be a needle biopsy, which allows for the early diagnosis of cysts. A 21-gauge needle attached to a 10-mL syringe is placed directly into the mass, which is fixed by fingers of the nondominant hand. The volume of a typical cyst is 5 to 10 mL, but it may be 75 mL or more. If the fluid that is aspirated is not bloodstained, then the cyst is aspirated to dryness, the needle is removed, and the fluid is discarded because cytologic examination of such fluid is not cost effective. After aspiration, the breast is carefully palpated to exclude a residual mass. In most cases, however, imaging has been performed prior to a needle being introduced into the breast, and indeed the majority of cysts are now aspirated under ultrasound guidance. If a mass was noted on initial ultrasound or there is a residual mass post aspiration, then a tissue specimen is obtained, usually by core biopsy. When cystic fluid is bloodstained, fluid can be sent for cytologic examination. A simple cyst is rarely of concern, but a complex cyst may be the result of an underlying malignancy. A pneumocystogram can be obtained by injecting air into the cyst and then obtaining a repeat mammogram. When this technique is used, the wall of the cyst cavity can be more carefully assessed for any irregularities.

Fibroadenomas. Most fibroadenomas are self-limiting and many go undiagnosed, so a more conservative approach is reasonable. Careful ultrasound examination with core-needle biopsy will provide for an accurate diagnosis. Ultrasonography may reveal specific features that are pathognomonic for fibroadenoma, and in a young woman (e.g., under 25 years) where the risk of breast cancer is already very low a core-needle biopsy may not be necessary. In patients where biopsy is performed, the patient is counseled concerning the ultrasound and biopsy results, and surgical excision of the fibroadenoma may be avoided. Cryoaublation and ultrasound-guided vacuum-assisted biopsy are approved treatments for fibroadenomas of the breast, especially lesions <3 cm. Larger lesions are often still best treated by excision. With short-term follow-up, a significant percentage of fibroadenomas will decrease in size and will no longer be palpable.38 However, many will remain palpable, especially those larger than 2 cm.38 Therefore, women should be counseled that the options for treatment include surgical removal, cryoablation, vacuum assisted biopsy, or observation.

Sclerosing Disorders. The clinical significance of sclerosing adenosis lies in its imitation of cancer. On physical examination, it may be confused with cancer, by mammography, and at gross pathologic examination. Excisional biopsy and histologic examination are frequently necessary to exclude the diagnosis of cancer. The diagnostic work-up for radial scars and complex sclerosing lesions frequently involves stereotactic biopsy. It usually is not possible to differentiate these lesions with certainty from cancer by mammographic features, so a larger tissue biopsy is recommended either by way of vacuum-assisted biopsy or an open surgical excisional biopsy. The mammographic appearance of a radial scar or sclerosing adenosis (mass density with spiculated margins) will usually lead to an assessment that the results of a core-needle biopsy specimen showing benign disease are discordant with the radiographic findings.

Periductal Mastitis. Painful and tender masses behind the nipple-areola complex are aspirated with a 21-gauge needle attached to a 10-mL syringe. Any fluid obtained is submitted for culture using a transport medium appropriate for the detection of anaerobic organisms. In the absence of pus, women are started on a combination of antibiotics to cover polymicrobial infections while awaiting the results of culture. Antibiotics are then continued based on sensitivity tests. Many cases respond satisfactorily to antibiotics alone, but when considerable purulent material is present, repeated ultrasound guided aspiration is performed, and ultimately in a proportion of cases surgical treatment is required. Unlike puerperal abscesses, a subareolar abscess is usually unilocular and often is associated with a single duct system. Ultrasound will accurately delineate its extent. In those cases that come to surgery, the surgeon may either undertake simple drainage with a view toward formal
Table 17-5

<table>
<thead>
<tr>
<th>Treatment of recurrent subareolar sepsis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SUITABLE FOR FISTULECTOMY</strong></td>
</tr>
<tr>
<td>Small abscess localized to one segment</td>
</tr>
<tr>
<td>Recurrence involving the same segment</td>
</tr>
<tr>
<td>Mild or no nipple inversion</td>
</tr>
<tr>
<td>Patient unconcerned about nipple inversion</td>
</tr>
<tr>
<td>Younger patient</td>
</tr>
<tr>
<td>No discharge from other ducts</td>
</tr>
<tr>
<td>No prior fistulectomy</td>
</tr>
</tbody>
</table>


**RISK FACTORS FOR BREAST CANCER**

**Hormonal and Nonhormonal Risk Factors**

Increased exposure to estrogen is associated with an increased risk for developing breast cancer, whereas reducing exposure is thought to be protective. Correspondingly, factors that increase the number of menstrual cycles, such as early menarche, nulliparity, and late menopause, are associated with increased risk. Moderate levels of exercise and a longer lactation period, factors that decrease the total number of menstrual cycles, are protective. The terminal differentiation of breast epithelium associated with a full-term pregnancy is also protective, so older age at first live birth is associated with an increased risk of breast cancer. Finally, there is an association between obesity and increased breast cancer risk. Because the major source of estrogen in postmenopausal women is the conversion of androstenedione to estrone by adipose tissue, obesity is associated with a long-term increase in estrogen exposure.

Nonhormonal risk factors include radiation exposure. Young women who receive mantle radiation therapy for Hodgkin’s lymphoma have a breast cancer risk that is 75 times greater than that of age-matched control subjects. Survivors of the atomic bomb blasts in Japan during World War II have a very high incidence of breast cancer, likely because of somatic mutations induced by the radiation exposure. In both circumstances, radiation exposure during adolescence, a period of active breast development, magnifies the deleterious effect. Studies also suggest that the risk of breast cancer increases as the amount of alcohol a woman consumes increases. Alcohol consumption is known to increase serum levels of estradiol. Finally, evidence suggests that long-term consumption of foods with a high fat content contributes to an increased risk of breast cancer by increasing serum estrogen levels.

**Risk Assessment Models**

The average lifetime risk of breast cancer for newborn U.S. women is 12%. The longer a woman lives without cancer, the lower her risk of developing breast cancer. Thus, a woman age 50 years has an 11% lifetime risk of developing breast cancer, and a woman age 70 years has a 7% lifetime risk of developing breast cancer. Because risk factors for breast cancer interact, evaluating the risk conferred by combinations of risk factors is difficult. There are several risk assessment models available to predict the risk of breast cancer. From the Breast Cancer Detection Demonstration Project, a mammography screening program conducted in the 1970s, Gail et al developed the model most frequently used in the United States, which incorporates age, age at menarche, age at first live birth, the number of breast biopsy specimens, any history of atypical hyperplasia, and number of first-degree relatives with breast cancer. It predicts the cumulative risk of breast cancer according to decade of life. To calculate breast cancer risk using the Gail model, a woman’s risk factors are translated into an overall risk score by multiplying her relative risks from several categories (Table 17-6). This risk score is then compared to an adjusted population risk of breast cancer to determine a woman’s individual or absolute risk. The output is a 5-year risk and a lifetime risk of developing breast cancer. A software program incorporating the Gail model is available from the National Cancer Institute at [http://bcra.nci.nih.gov/brc](http://bcra.nci.nih.gov/brc). This model was recently modified to more accurately assess risk in African American women. There have also been modifications that project individualized absolute
Table 17-6
Relative risk estimates for the Gail model

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>RELATIVE RISK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at menarche (years)</td>
<td></td>
</tr>
<tr>
<td>≥14</td>
<td>1.00</td>
</tr>
<tr>
<td>12–13</td>
<td>1.10</td>
</tr>
<tr>
<td>&lt;12</td>
<td>1.21</td>
</tr>
<tr>
<td>Number of biopsy specimens/history of benign breast disease, age &lt;50 y</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.00</td>
</tr>
<tr>
<td>1</td>
<td>1.70</td>
</tr>
<tr>
<td>≥2</td>
<td>2.88</td>
</tr>
<tr>
<td>Number of biopsy specimens/history of benign breast disease, age ≥50 y</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.02</td>
</tr>
<tr>
<td>1</td>
<td>1.27</td>
</tr>
<tr>
<td>≥2</td>
<td>1.62</td>
</tr>
<tr>
<td>Age at first live birth (years)</td>
<td></td>
</tr>
<tr>
<td>&lt;20 y</td>
<td></td>
</tr>
<tr>
<td>Number of first-degree relatives with history of breast cancer</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.00</td>
</tr>
<tr>
<td>1</td>
<td>2.61</td>
</tr>
<tr>
<td>≥2</td>
<td>6.80</td>
</tr>
<tr>
<td>20–24 y</td>
<td></td>
</tr>
<tr>
<td>Number of first-degree relatives with history of breast cancer</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.24</td>
</tr>
<tr>
<td>1</td>
<td>2.68</td>
</tr>
<tr>
<td>≥2</td>
<td>5.78</td>
</tr>
<tr>
<td>25–29 y</td>
<td></td>
</tr>
<tr>
<td>Number of first-degree relatives with history of breast cancer</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.55</td>
</tr>
<tr>
<td>1</td>
<td>2.76</td>
</tr>
<tr>
<td>≥2</td>
<td>4.91</td>
</tr>
<tr>
<td>≥30 y</td>
<td></td>
</tr>
<tr>
<td>Number of first-degree relatives with history of breast cancer</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.93</td>
</tr>
<tr>
<td>1</td>
<td>2.83</td>
</tr>
<tr>
<td>≥2</td>
<td>4.17</td>
</tr>
</tbody>
</table>


invasive breast cancer risk for Asian and Pacific Island American women. The Gail model is the most widely used model in the United States. Gail and colleagues have also described a revised model that includes body weight and mammographic density but excludes age at menarche.54

Claus et al, using data from the Cancer and Steroid Hormone Study, a case-control study of breast cancer, developed the other frequently used risk assessment model, which is based on assumptions about the prevalence of high-penetration breast cancer susceptibility genes.55 Compared with the Gail model, the Claus model incorporates more information about family history but excludes other risk factors. The Claus model provides individual estimates of breast cancer risk according to decades of life based on presence of first- and second-degree relatives with breast cancer and their age at diagnosis. Risk factors that are less consistently associated with breast cancer (diet, use of oral contraceptives, lactation) or are rare in the general population (radiation exposure) are not included in either the Gail or Claus risk assessment model. Other models have been proposed that account for mammographic breast density in assessing breast cancer risk.54,56

Neither the Gail model nor the Claus model accounts for the risk associated with mutations in the breast cancer susceptibility genes BRCA1 and BRCA2 (described in detail in the following section). The BRCAPRO model is a Mendelian model that calculates the probability that an individual is a carrier of a mutation in one of the breast cancer susceptibility genes based on their family history of breast and ovarian cancer.57 The probability that an individual will develop breast or ovarian cancer is derived from this mutation probability based on age-specific incidence curves for both mutation carriers and noncarriers.58 Use of the BRCAPRO model in the clinic is challenging since it requires input of all family history information regarding breast and ovarian cancer. The Tyrer-Cuzick model attempts to utilize both family history information and individual risk information. It uses the family history to calculate the probability that an individual carries a mutation in one of the breast cancer susceptibility genes, and then the risk is adjusted based on personal risk factors, including age at menarche, parity, age at first live birth, age at menopause, history of atypical hyperplasia or LCIS, height, and body mass index.59 Once a risk model has been utilized to assess breast cancer risk, this must be communicated to the individual and put into context with competing risk and medical comorbidities. This information can then be used to discuss options that are available to the individual for managing risk.

Risk Management

Several important medical decisions may be affected by a woman’s underlying risk of developing breast cancer.60-68 These decisions include when to use postmenopausal hormone replacement therapy, at what age to begin mammography screening or incorporate magnetic resonance imaging (MRI) screening, when to use tamoxifen to prevent breast cancer, and when to perform prophylactic mastectomy to prevent breast cancer. Postmenopausal hormone replacement therapy was widely prescribed in the 1980s and 1990s because of its effectiveness in controlling the symptoms of estrogen deficiency, namely vasomotor symptoms such as hot flashes, night sweats and their associated sleep deprivation, osteoporosis, and cognitive changes. Furthermore, these hormone supplements were thought to reduce coronary artery disease as well. Use of combined estrogen and progesterone became standard for women who had not undergone hysterectomy because unopposed estrogen increases the risk of uterine cancer. Concerns of prolonging a woman’s lifetime exposure to estrogen, coupled with conflicting data regarding the impact of these hormones on cardiovascular health, motivated the implementation of large-scale phase 3 clinical trials to definitively evaluate the risks vs. benefits of postmenopausal hormone replacement therapy. The Women’s Health Initiative (WHI) was therefore designed by the National Institutes of Health as a series of clinical trials to study the effects of diet, nutritional supplements, and hormones on the risk of cancer, cardiovascular disease, and bone health in postmenopausal women. Findings from primary studies of postmenopausal hormone replacement therapy were released in 2002, demonstrating conclusively that
breast cancer risk is threefold to fourfold higher after >4 years of use and there is no significant reduction in coronary artery or cerebrovascular risks. The Collaborative Group on Hormonal Factors in Breast Cancer combined and reanalyzed data from a number of studies totaling 52,705 women with breast cancer and 108,411 women without breast cancer. They found an increased risk of breast cancer with every use of estrogen replacement therapy. They also reported increased risk among current users but not past users and risk increased with increasing duration of use of hormone replacement therapy. Chebblawi et al also reported from the WHI study that estrogen + progesterone increased the incidence of breast cancer. This was confirmed by the Million Women study, which also showed that the increased risk was substantially greater for the combined estrogen + progesterone replacement therapy than other types of hormone replacement therapy.

**Breast Cancer Screening.** Routine use of screening mammography in women ≥50 years of age has been reported to reduce mortality from breast cancer by 25%. This reduction comes at an acceptable economic cost. More recently, there has been debate over the potential harms associated with breast screening. Controversy over the age to initiate screening mammography is evident in the current recommendations. The U.S. Preventive Services Task Force (USPSTF), the American Cancer Society (ACS), and the National Comprehensive Cancer Network (NCCN) are three organizations with differing recommendations for screening mammography in average risk women. The guidelines, however, similarly define high-risk women as those with personal history of breast cancer, history of chest radiation at young age, and confirmed or suspected genetic mutation known to increase risk for developing breast cancer. The USPSTF recommends biennial screening mammography for women age 50 to 74 years. The USPSTF applies these guidelines to asymptomatic women age ≥40 years who do not have a preexisting breast cancer or who were not previously diagnosed with a high-risk breast lesion, and who are not at high risk for breast cancer because of a known underlying genetic mutation or history of chest radiation at a young age. In October 2015, the ACS released updated guidelines stating average-risk women should start annual screening mammography at 45 years of age. Women age 45 to 54 years should be screened annually, and those 55 years and older should transition to biennial screening or have the opportunity to continue annual screening. Women should have the opportunity to begin annual screening between the ages of 40 and 44 years and should continue screening as long as their overall health is good and have a life expectancy of 10 years or longer. The ACS recommends that average-risk women begin annual screening mammograms at ≥40 years of age, along with annual clinical breast exams and breast awareness.

The United Kingdom recently established an independent expert panel to review the published literature and estimate the benefits and harms associated with screening women >50 years of age in its national screening program. The expert panel estimated that an invitation to breast screening delivers about a 20% reduction in breast cancer mortality. At the same time, however, the panel estimated that in women invited to the screening, about 11% of the cancers diagnosed in their lifetime constitute overdagnosis. Despite the overdagnosis, the panel concluded that breast screening confers significant benefit and should continue. The use of screening mammography in women <50 years of age is more controversial for several reasons: (a) breast density is greater, and screening mammography is less likely to detect early breast cancer (i.e., reduced sensitivity); (b) screening mammography results in more false-positive test findings (i.e., reduced specificity), which results in unnecessary biopsy specimens; and (c) younger women are less likely to have breast cancer (i.e., lower incidence), so fewer young women will benefit from screening. In the United States, on a population basis, however, the benefits of screening mammography in women between the ages of 40 and 49 years is still felt to outweigh the risks; although targeting mammography to women at higher risk of breast cancer improves the balance of risks and benefits and is the approach some health care systems have taken. In one study of women age 40 to 49 years, an abnormal mammography finding was three times more likely to be cancer in a woman with a family history of breast cancer than in a woman without such a history. Furthermore, as noted previously in the section *Risk Assessment Models*, mounting data regarding mammographic breast density demonstrate an independent correlation with breast cancer risk. Incorporation of breast density measurements into breast cancer risk assessment models appears to be a promising strategy for increasing the accuracy of these tools. Unfortunately, widespread application of these modified models is hampered by inconsistencies in the reporting of mammographic density. Ultrasonography can also be used for breast cancer screening in women with dense breasts, but there is no data available that the additional cancers detected with this modality reduce mortality from breast cancer.

Current recommendations by the United States Preventive Services Task Force are that women undergo biennial mammographic screening between the ages of 50 and 74 years. The use of MRI for breast cancer screening is recommended by the ACS for women with a 20% to 25% or greater lifetime risk using risk assessment tools based mainly on family history. *BRCA* mutation carriers, those individuals who have a family member with a *BRCA* mutation who have not been tested themselves, individuals who received radiation to the chest between the ages of 10 and 30 years, and those individuals with a history of Li-Fraumeni syndrome, Cowden syndrome, or Bannayan-Riley-Ruvalcaba syndrome or those who have a first-degree relative with one of these syndromes. MRI is an extremely sensitive screening tool that is not limited by the density of the breast tissue as mammography is; however, its specificity is moderate, leading to more false-positive events and the increased need for biopsy.

**Chemoprevention.** Tamoxifen, a selective estrogen receptor modulator, was the first drug shown to reduce the incidence of breast cancer in healthy women. There have been four prospective studies published evaluating tamoxifen vs. placebo for reducing the incidence of invasive breast cancer for women at increased risk. The largest trial was the Breast Cancer Prevention Trial (NSABP P-01), which randomly assigned >13,000 women with a 5-year Gail relative risk of breast cancer of 1.66% or higher or LCIS to receive tamoxifen or placebo. After a mean follow-up period of 4 years, the incidence of breast cancer was reduced by 49% in the group receiving tamoxifen. The decrease was evident only in ER-positive breast cancers with no significant change in ER-negative tumors. The Royal Marsden Hospital Tamoxifen Chemoprevention Trial, the Italian Tamoxifen Prevention Trial, and the International Breast Cancer Intervention Study I (IBIS-I) trial all showed a reduction in...
ER-positive breast cancers with the use of tamoxifen compared with placebo. There was no effect on mortality; however, the trials were not powered to assess either breast cancer mortality or all-cause mortality events. The adverse events were similar in all four randomized trials, including an increased risk of endometrial cancer, thromboembolic events, cataract formation, and vasomotor disturbances in individuals receiving tamoxifen.

Tamoxifen therapy currently is recommended only for women who have a Gail relative risk of 1.66% or higher, who are age 35 to 59, women over the age of 60, or women with a diagnosis of LCIS or atypical ductal or lobular hyperplasia. In addition, deep vein thrombosis occurs 1.6 times as often, pulmonary emboli 3.0 times as often, and endometrial cancer 2.5 times as often in women taking tamoxifen. The increased risk for endometrial cancer is restricted to early stage cancers in postmenopausal women. Cataract surgery is required almost twice as often among women taking tamoxifen. Gail et al subsequently developed a model that accounts for underlying risk of breast cancer as well as comorbidities to determine the net risk-benefit ratio of tamoxifen use for chemoprevention.84

The NSABP completed a second chemoprevention trial, designed to compare tamoxifen and raloxifene for breast cancer risk reduction in high-risk postmenopausal women. Raloxifene, another selective estrogen receptor modulator, was selected for the experimental arm in this follow-up prevention trial because its use in managing postmenopausal osteoporosis suggested that it might be even more effective at breast cancer risk reduction, but without the adverse effects of tamoxifen on the uterus. The P-2 trial, the Study of Tamoxifen and Raloxifene (known as the STAR trial), randomly assigned 19,747 postmenopausal women at high-risk for breast cancer to receive either tamoxifen or raloxifene. The initial report of the P-2 trial showed the two agents were nearly identical in their ability to reduce breast cancer risk, but raloxifene was associated with a more favorable adverse event profile.85 An updated analysis revealed that raloxifene maintained 76% of the efficacy of tamoxifen in prevention of invasive breast cancer with a more favorable side effect profile. The risk of developing endometrial cancer was significantly higher with tamoxifen use at longer follow-up.86 Although tamoxifen has been shown to reduce the incidence of LCIS and DCIS, raloxifene did not have an effect on the frequency of these diagnoses.

Aromatase inhibitors (AIs) have been shown to be more effective than tamoxifen in reducing the incidence of contralateral breast cancers in postmenopausal women receiving AIs for adjuvant treatment of invasive breast cancer. The MAP3 trial was the first study to evaluate an AI as a chemopreventive agent in postmenopausal women at high risk for breast cancer. The trial randomized 4560 women to exemestane 25 mg daily vs. placebo for 5 years. After a median follow-up of 35 months, exemestane was shown to reduce invasive breast cancer incidence by 65%. Side effect profiles demonstrated more grade II or higher arthritis and hot flashes in patients taking exemestane.87 The IBIS II trial on the other hand, randomized 3864 postmenopausal women to either anastrozole, a nonsteroidal aromatase inhibitor, vs. placebo with a further randomization to bisphosphate or not based on bone density.88,89 After a median follow-up of 5 years, anastrozole reduced the incidence of invasive breast cancer by about 50%. The trial also had an initial sub-study that looked at the effect of the aromatase inhibitor on cognitive function and reported no adverse effects.90 The American Society of Clinical Oncology recommends tamoxifen for chemoprevention in premenopausal or postmenopausal women and consideration for raloxifene or exemestane in postmenopausal women who are noted to be at increased risk of breast cancer.91,92 The discussion with an individual patient should include risk assessment and potential risks and benefits with each agent.

Risk-Reducing Surgery. A retrospective study of women at high risk for breast cancer found that prophylactic mastectomy reduced their risk by >90%.93 However, the effects of prophylactic mastectomy on the long-term quality of life are poorly quantified. A study involving women who were carriers of a breast cancer susceptibility gene (BRCA1) mutation found that the benefit of prophylactic mastectomy differed substantially according to the breast cancer risk conferred by the mutations. For women with an estimated lifetime risk of 40%, prophylactic mastectomy added almost 3 years of life, whereas for women with an estimated lifetime risk of 85%, prophylactic mastectomy added >5 years of life.94 Domchek et al evaluated a cohort of BRCA1 and 2 mutation carriers who were followed prospectively and reported on outcomes with risk-reducing surgery.95 They found that risk-reducing mastectomy was highly effective at preventing breast cancer in both BRCA1 and 2 mutation carriers. Risk-reducing salpingo-oophorectomy was highly effective at reducing the incidence of ovarian cancer and breast cancer in BRCA1 mutation carriers and was associated with a reduction in breast cancer-specific mortality, ovarian cancer-specific mortality, and all-cause mortality. While studies of bilateral prophylactic or risk-reducing mastectomy have reported dramatic reductions in breast cancer incidence among those without known BRCA1 mutations, there is little data to support a survival benefit. Another consideration is that while most patients are satisfied with their decision to pursue risk-reducing surgery, some are dissatisfied with the cosmetic outcomes mostly due to reconstructive issues.

BRCA Mutations

BRCA1. Up to 5% of breast cancers are caused by inheritance of germline mutations such as BRCA1 and BRCA2, which are inherited in an autosomal dominant fashion with varying degrees of penetrance (Table 17–7).94–100 BRCA1 is located on chromosome arm 17q, spans a genomic region of approximately 100 kilobases (kb) of DNA, and contains 22 coding exons for 1863 amino acids. Both BRCA1 and BRCA2 function as tumor-suppressor genes, and for each gene, loss of both alleles is required for the initiation of cancer. Data accumulated since the isolation of the BRCA1 gene suggest a role in transcription, cell-cycle control, and DNA damage repair pathways. More than 500 sequence variations in BRCA1 have been identified. It now is known that germline mutations in BRCA1 represent a predisposing genetic factor in as many as 45% of hereditary breast cancers and in at least 80% of hereditary ovarian cancers. Female mutation carriers have been reported to have up to an 85% lifetime risk (for some families) for developing breast cancer and up to a 40% lifetime risk for developing ovarian cancer. The initial families reported had high penetrance and subsequently the average lifetime risk has been reported to lie between 60% and 70%. Breast cancer susceptibility in these families appears as an autosomal dominant trait with high penetrance. Approximately 50% of children of carriers inherit the trait. In general, BRCA1-associated breast cancers are invasive ductal carcinomas, are poorly differentiated, are in the majority
hormone receptor negative, and have a triple receptor negative (immunohistochemical profile: ER-negative, PR-negative, and HER2-negative) or basal phenotype (based on gene expression profiling). **BRCA1**-associated breast cancers have a number of distinguishing clinical features, such as an early age of onset compared with sporadic cases; a higher prevalence of bilateral breast cancer; and the presence of associated cancers in some affected individuals, specifically ovarian cancer and possibly colon and prostate cancers.

Several founder mutations have been identified in **BRCA1**. The two most common mutations are 185delAG and 5382insC, which account for 10% of all the mutations seen in **BRCA1**. These two mutations occur at a 10-fold higher frequency in the Ashkenazi Jewish population than in non-Jewish Caucasians. The carrier frequency of the 185delAG mutation in the Ashkenazi Jewish population is 1% and, along with the 5382insC mutation, accounts for almost all **BRCA1** mutations in this population. Analysis of germline mutations in Jewish and non-Jewish women with early-onset breast cancer indicates that 20% of Jewish women who develop breast cancer before age 40 carry the 185delAG mutation. There are founder **BRCA1** mutations in other populations including, among others, Dutch, Polish, Finnish, and Russian populations.

**BRCA2**. **BRCA2** is located on chromosome arm 13q and spans a genomic region of approximately 70 kb of DNA. The 11.2 kb coding region contains 26 coding exons. It encodes a protein of 3418 amino acids. The **BRCA2** gene bears no homology to any previously described gene, and the protein contains no previously defined functional domains. The biologic function of **BRCA2** is not well defined, but like **BRCA1**, it is postulated to play a role in DNA damage response pathways. **BRCA2** messenger RNA also is expressed at high levels in the late G1 and S phases of the cell cycle. The kinetics of **BRCA2** protein regulation in the cell cycle is similar to that of **BRCA1** protein, which suggests that these genes are coregulated. The mutational spectrum of **BRCA2** is not as well established as that of **BRCA1**. To date, >250 mutations have been found. The breast cancer risk for **BRCA2** mutation carriers is close to 85%, and the lifetime ovarian cancer risk, while lower than for **BRCA1**, is still estimated to be close to 20%. Breast cancer susceptibility in **BRCA2** families is an autosomal dominant trait and has a high penetrance. Approximately 50% of children of carriers inherit the trait. Unlike male carriers of **BRCA1** mutations, men with germline mutations in **BRCA2** have an estimated breast cancer risk of 6%, which represents a 100-fold increase over the risk in the general male population. **BRCA2**-associated breast cancers are invasive ductal carcinomas, which are more likely to be well differentiated and to express hormone receptors than are **BRCA1**-associated breast cancers. **BRCA2**-associated breast cancer has a number of distinguishing clinical features, such as an early age of onset compared with sporadic cases, a higher prevalence of bilateral breast cancer, and the presence of associated cancers in some affected individuals, specifically ovarian, colon, prostate, pancreatic, gallbladder, bile duct, and stomach cancers, as well as melanoma. A number of founder mutations have been identified in **BRCA2**. The 6174delT mutation is found in Ashkenazi Jews with a prevalence of 1.2% and accounts for 60% of ovarian cancer and 30% of early-onset breast cancer patients among Ashkenazi women. Another **BRCA2** founder mutation, 999del5, is observed in Icelandic and Finnish populations, while more recently 3036delA has been observed in a number of Spanish families.

### Table 17-7

<table>
<thead>
<tr>
<th>Type</th>
<th>Incidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sporadic breast cancer</td>
<td>65%–75%</td>
</tr>
<tr>
<td>Familial breast cancer</td>
<td>20%–30%</td>
</tr>
<tr>
<td>Hereditary breast cancer</td>
<td>5%–10%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Incidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRCA1</td>
<td>45%</td>
</tr>
<tr>
<td>BRCA2</td>
<td>35%</td>
</tr>
<tr>
<td>p53^(a) (Li-Fraumeni syndrome)</td>
<td>1%</td>
</tr>
<tr>
<td>STK11/LKB1^(a) (Peutz-Jeghers syndrome)</td>
<td>&lt;1%</td>
</tr>
<tr>
<td>PTEN^(a) (Cowden disease)</td>
<td>&lt;1%</td>
</tr>
<tr>
<td>MSH2/MLH1^(a) (Muir-Torre syndrome)</td>
<td>&lt;1%</td>
</tr>
<tr>
<td>ATM^(a) (Ataxia-telangiectasia)</td>
<td>&lt;1%</td>
</tr>
<tr>
<td>Unknown</td>
<td>20%</td>
</tr>
</tbody>
</table>

^(a) Affected gene.


**Identification of BRCA Mutation Carriers**. Identifying hereditary risk for breast cancer is a four-step process that includes: (a) obtaining a complete, multigenerational family history, (b) assessing the appropriateness of genetic testing for a particular patient, (c) counseling the patient, and (d) interpreting the results of testing. Genetic testing should not be offered in isolation, but only in conjunction with patient education and counseling, including referral to a genetic counselor. Initial determinations include whether the individual is an appropriate candidate for genetic testing and whether genetic testing will be informative for personal and clinical decision-making. A thorough and accurate family history is essential to this process, and the maternal and paternal sides of the family are both assessed because 50% of the women with a **BRCA** mutation have inherited the mutation from their fathers. To help clinicians advise women about genetic testing, statistically based models that determine the probability that an individual carries a **BRCA** mutation have been developed. A method for calculating carrier probability that has been demonstrated to have acceptable performance (i.e., both in terms of calibration and discrimination) such as the Manchester scoring system and BODICEA should be used to offer referral to a specialist genetic clinic. A hereditary risk of breast cancer is considered if a family includes Ashkenazi Jewish heritage; a first-degree relative with breast cancer before age 50; a history of ovarian cancer at any age in the patient or first- or second-degree relative with ovarian cancer; breast and ovarian cancer in the same individual; two or more first- or second-degree relatives with breast cancer at any age; patient or relative with bilateral breast cancer; and male breast cancer in a relative at any age. The threshold for genetic testing is lower in individuals who are members of ethnic groups in whom the mutation prevalence is increased.

**BRCA Mutation Testing**. Appropriate counseling for the individual being tested for a **BRCA** mutation is strongly recommended, and documentation of informed consent is required. The test that is clinically available for analyzing **BRCA** mutations is gene sequence analysis. In a family with a history suggestive of hereditary breast cancer and no previously
tested member, the most informative strategy is first to test an affected family member. This person undergoes complete sequence analysis of both the *BRCA1* and *BRCA2* genes. If a mutation is identified, relatives are usually tested only for that specific mutation. An individual of Ashkenazi Jewish ancestry is tested initially for the three specific mutations that account for hereditary breast and ovarian cancer in that population. If results of that test are negative, it may then be appropriate to fully analyze the *BRCA1* and *BRCA2* genes.

A positive test result is one that discloses the presence of a *BRCA* mutation that interferes with translation or function of the *BRCA* protein. A woman who carries a deleterious mutation has a breast cancer risk of up to 85% (in some families) as well as a greatly increased risk of ovarian cancer. A negative test result is interpreted according to the individual’s personal and family history, especially whether a mutation has been previously identified in the family, in which case the woman is generally tested only for that specific mutation. If the mutation is not present, the woman’s risk of breast or ovarian cancer may be no greater than that of the general population. In addition, no *BRCA* mutation can be passed on to the woman’s children. In the absence of a previously identified mutation, a negative test result in an affected individual generally indicates that a *BRCA* mutation is not responsible for the familial cancer. However, the possibility remains of an unusual abnormality in one of these genes that cannot yet be identified through clinical testing. It also is possible that the familial cancer is indeed caused by an identifiable *BRCA* mutation but that the individual tested had sporadic cancer, a situation known as *phenocopy*. This is especially possible if the individual tested developed breast cancer close to the age of onset of the general population (age 60 years or older) rather than before age 50 years, as is characteristic of *BRCA* mutation carriers. Overall, the false-negative rate for *BRCA* mutation testing is <5%. Some test results, especially when a single base-pair change (missense mutation) is identified, may be difficult to interpret. This is because single base-pair changes do not always result in a nonfunctional protein. Thus, missense mutations not located within critical functional domains, or those that cause only minimal changes in protein structure, may not be disease associated and are usually reported as indeterminate results. In communicating indeterminate results to women, care must be taken to relay the uncertain cancer risk associated with this type of mutation and to emphasize that ongoing research might clarify its meaning. In addition, testing other family members with breast cancer to determine if a genetic variant tracks with their breast cancer may provide clarification as to its significance. Indeterminate genetic variance currently accounts for 12% of the test results.

Concern has been expressed that the identification of hereditary risk for breast cancer may interfere with access to affordable health insurance. This concern refers to discrimination directed against an individual or family based solely on an apparent or perceived genetic variation from the normal human genotype. The Health Insurance Portability and Accountability Act of 1996 (HIPAA) made it illegal in the United States for group health plans to consider genetic information as a preexisting condition or to use it to deny or limit coverage. Most states also have passed laws that prevent genetic discrimination in the provision of health insurance. In addition, individuals applying for health insurance are not required to report whether relatives have undergone genetic testing for cancer risk, only whether those relatives have actually been diagnosed with cancer. Currently, there is little documented evidence of genetic discrimination resulting from findings of available genetic tests.

**Cancer Prevention for *BRCA* Mutation Carriers.** Risk management strategies for *BRCA1* and *BRCA2* mutation carriers include the following:

1. Risk-reducing mastectomy and reconstruction
2. Risk-reducing salpingo-oophorectomy
3. Intensive surveillance for breast and ovarian cancer
4. Chemoprevention

Although removal of breast tissue reduces the likelihood that *BRCA1* and *BRCA2* mutation carriers will develop breast cancer, mastectomy does not remove all breast tissue, and women continue to be at risk because a germline mutation is present in any remaining breast tissue. For postmenopausal *BRCA1* and *BRCA2* mutation carriers who have not had a mastectomy, it may be advisable to avoid hormone replacement therapy because no data exist regarding the effect of the therapy on the penetrance of breast cancer susceptibility genes. Because breast cancers in *BRCA* mutation carriers have the same mammographic appearance as breast cancers in noncarriers, a screening mammogram is likely to be effective in *BRCA* mutation carriers, provided it is performed and interpreted by an experienced radiologist with a high level of suspicion. Present screening recommendations for *BRCA* mutation carriers who do not undergo risk-reducing mastectomy include clinical breast examination every 6 months and mammography every 12 months beginning at age 25 years because the risk of breast cancer in *BRCA* mutation carriers increases after age 30 years. Recent attention has been focused on the use of MRI for breast cancer screening in high-risk individuals and known *BRCA* mutation carriers. MRI appears to be more sensitive at detecting breast cancer in younger women with dense breasts. However, as noted previously, MRI does lead to the detection of benign breast lesions that cannot easily be distinguished from malignancy, and these false-positive events can result in more interventions, including biopsy specimens. The current recommendations from the American Cancer Society are for annual MRI in women with a 20% to 25% or greater lifetime risk of developing breast cancer (mainly based on family history), women with a known *BRCA1* or *BRCA2* mutation, those who have a first-degree relative with a *BRCA1* or *BRCA2* mutation and have not had genetic testing themselves, women who were treated with radiation therapy to the chest between the ages of 10 and 30 years, and those who have Li-Fraumeni syndrome, Cowden syndrome, or Bannayan-Riley-Ruvalcaba syndrome, or a first-degree relative with one of these syndromes. Despite a 49% reduction in the overall incidence of breast cancer and a 69% reduction in the incidence of estrogen receptor positive tumors in high-risk women taking tamoxifen reported in the NSABP P1 trial, there is insufficient evidence to recommend the use of tamoxifen uniformly for *BRCA1* mutation carriers. Cancers arising in *BRCA1* mutation carriers are usually high grade and are most often hormone receptor negative. Approximately 66% of *BRCA1*-associated DCIS lesions are estrogen receptor negative, which suggests early acquisition of the hormone-independent phenotype. In the NSABP P1 trial there was a 62% reduction in the incidence of breast cancer in *BRCA2* carriers, similar to the overall reduction seen in the P1 trial. In contrast, there was no reduction seen in breast cancer incidence in *BRCA1* carriers who started tamoxifen in P1 age 35 years or...
older. Tamoxifen appears to be more effective at preventing estrogen receptor-positive breast cancers.

The risk of ovarian cancer in BRCA1 and BRCA2 mutation carriers ranges from 20% to 40%, which is 10 times higher than that in the general population. Risk-reducing salpingo-oophorectomy is a reasonable prevention option in mutation carriers. In women with a documented BRCA1 or BRCA2 mutation, consideration for bilateral risk-reducing salpingo-oophorectomy should be between the ages of 35 and 40 years at the completion of childbearing. Removing the ovaries reduces the risk of ovarian cancer and breast cancer when performed in premenopausal BRCA mutation carriers. Hormone replacement therapy is discussed with the patient at the time of oophorectomy. The Cancer Genetics Studies Consortium recommends yearly transvaginal ultrasound timed to avoid ovulation and annual measurement of serum cancer antigen 125 levels beginning at age 25 years as the best screening modalities for ovarian carcinoma in BRCA mutation carriers who have opted to defer risk-reducing surgery.

PALB2 (partner and localizer of BRCA2) has recently been characterized as a potential high-risk gene for breast cancer. PALB2 allows nuclear localization of BRCA2 and provides a scaffold for the BRCA1–PALB2–BRCA2 complex. Analysis by Antoniou et al has suggested that the risk of breast cancer for PALB2 mutation carriers is as high as that of BRCA2 mutation carriers. The absolute risk of breast cancer for PALB2 female mutation carriers by 70 years of age ranged from 33% (95% CI, 25–44) for those with no family history of breast cancer to 58% (95% CI, 50–66) for those with two or more first-degree relatives with breast cancer at 30 years of age. The risk of breast cancer for female PALB2 mutation carriers, depending on the age, was about five to nine times as high compared with the general population. While screening with mammogram along with MRI has been suggested for PALB2 mutation carriers starting at age 30 with consideration of risk-reducing mastectomy, there is currently insufficient evidence regarding the risk of ovarian cancer and its management.

Other hereditary syndromes associated with an increased risk of breast cancer include Cowden disease (PTEN mutations, in which cancers of the thyroid, GI tract, and benign skin and subcutaneous nodules are also seen), Li-Fraumeni syndrome (TP53 mutations, also associated with sarcomas, lymphomas, and adrenocortical tumors), hereditary diffuse gastric cancer syndrome (CDH1 mutations, associated with diffuse gastric cancer and lobular breast cancers), and syndromes of breast and melanoma. With the discovery of additional genes related to breast cancer susceptibility, panel testing is available for a number of genes in addition to BRCA1 and BRCA2. The interpretation of results is complex and is best done with a genetic counselor.

**Epidemiology and Natural History of Breast Cancer**

**Epidemiology**

Breast cancer is the most common site-specific cancer in women and is the leading cause of death from cancer for women age 20 to 59 years. Based on Surveillance, Epidemiology, and End Results registries (SEER) data, 266,120 new cases were estimated in 2018 with 40,920 estimated deaths attributed to breast cancers. It accounts for 30% of all newly diagnosed cancers in women and is responsible for 14% of the cancer-related deaths in women.

Breast cancer was the leading cause of cancer-related mortality in women until 1987, when it was surpassed by lung cancer. In the 1970s, the probability that a woman in the United States would develop breast cancer at some point in her lifetime was estimated at 1 in 13; in 1980 it was 1 in 11; and in 2004 it was 1 in 8. Cancer registries in Connecticut and upper New York State document that the age-adjusted incidence of new breast cancer cases had steadily increased since the mid-1940s. The incidence in the United States, based on data from nine SEER registries, has been decreasing by 23% per year since 2000. The increase had been approximately 1% per year from 1973 to 1980, and there was an additional increase in incidence of 4% between 1980 and 1987, which was characterized by frequent detection of small primary cancers. The increase in breast cancer incidence occurred primarily in women age ≥55 years and paralleled a marked increase in the percentage of older women who had mammograms taken. At the same time, incidence rates for regional metastatic disease dropped and breast cancer mortality declined. From 1960 to 1963, 5-year overall survival rates for breast cancer were 63% and 46% in white and African American women, respectively, whereas the rates for 1981 to 1983 were 78% and 64%, respectively. For 2002 to 2008 rates were 92% and 78%, respectively.

There is a 10-fold variation in breast cancer incidence among different countries worldwide. Cyprus and Malta have the highest age-adjusted mortality for breast cancer (29.6 per 100,000 population), whereas Haiti has the lowest (2.0 deaths per 100,000 population). The United States has an age-adjusted mortality for breast cancer of 19.0 cases per 100,000 population. Women living in less industrialized nations tend to have a lower incidence of breast cancer than women living in industrialized countries, although Japan is an exception. In the United States, Mormons, Seventh Day Adventists, American Indians, Alaska natives, Hispanic/Latina Americans, and Japanese and Filipino women living in Hawaii have a below-average incidence of breast cancer, whereas nuns (due to nulliparity) and Ashkenazi Jewish women have an above-average incidence.

The incidence rates of breast cancer increased in most countries through the 1990s. Since the estimates for 1990, there was an overall increase in incidence rates of approximately 0.5% annually. It was predicted that there would be approximately 1.4 million new cases in 2010. The cancer registries in China have noted annual increases in incidence of up to 3% to 4%, and in eastern Asia, increases are similar.

Data from the SEER program reveal declines in breast cancer incidence over the past decade, and this is widely attributed to decreased use of hormone replacement therapy as a consequence of the Women’s Health Initiative reports.

Breast cancer burden has well-defined variations by geography, regional lifestyle, and racial or ethnic background. In general, both breast cancer incidence and mortality are relatively lower among the female populations of Asia and Africa, relatively underdeveloped nations, and nations that have not adopted Westernized reproductive and dietary patterns. In contrast, European and North American women and women from heavily industrialized or Westernized countries have a substantially higher breast cancer burden. These international patterns are mirrored in breast cancer incidence and mortality rates observed for the racially, ethnically, and culturally diverse population of the United States.
Although often related, the factors that influence breast cancer incidence may differ from those that affect mortality. Incidence rates are lower among populations that are heavily weighted with women who begin childbearing at young ages and who have multiple full-term pregnancies followed by prolonged lactation. These are features that characterize many underdeveloped nations and also many eastern nations. Breast cancer mortality rates should be lower in populations that have a lower incidence, but the mortality burden will simultaneously be adversely affected by the absence of effective mammographic screening programs for early detection and diminished access to multidisciplinary cancer treatment programs. These features are likely to account for much of the disproportionate mortality risks that are seen in underdeveloped nations. Similar factors probably account for differences in breast cancer burden observed among the various racial and ethnic groups within the United States. Interestingly, breast cancer incidence and mortality rates rise among second- and third-generation Asian Americans as they adopt Western lifestyles.

Disparities in breast cancer survival among subsets of the American population are generating increased publicity because they are closely linked to disparities in socioeconomic status. Poverty rates and proportions of the population that lack health care insurance are two to three times higher among minority racial and ethnic groups such as African Americans and Hispanic/Latino Americans. These socioeconomic disadvantages create barriers to effective breast cancer screening and result in delayed breast cancer diagnosis, advanced stage distribution, inadequacies in comprehensive treatment, and, ultimately, increased mortality rates. Furthermore, the rapid growth in the Hispanic population is accompanied by increasing problems in health education because of linguistic barriers between physicians and recently immigrated, non–English-speaking patients. Recent studies also are documenting inequities in the treatments delivered to minority breast cancer patients, such as increased rates of failure to provide systemic therapy, use of sentinel lymph node dissection, and breast reconstruction. Some of the treatment delivery disparities are related to inadequately controlled comorbidities (such as hypertension and diabetes), which are more prevalent in minority populations. However, some studies that adjust for these factors report persistent and unexplained unevenness in treatment recommendations. It is clear that breast cancer disparities associated with racial or ethnic background have a multifactorial cause, and improvements in outcome will require correction of many public health problems at both the patient and provider levels.

Advances in the ability to characterize breast cancer subtypes and the genetics of the disease are now provoking speculation regarding possible hereditary influences on breast cancer risk that are related to racial or ethnic ancestry. These questions become particularly compelling when one looks at disparities in breast cancer burden between African Americans and Caucasians. Lifetime risk of breast cancer is lower for African Americans, yet a paradoxically increased breast cancer mortality risk also is seen. African Americans also have a younger age distribution for breast cancer; among women <45 years of age, breast cancer incidence is highest among African Americans compared to other subsets of the American population. Lastly and most provocatively, African American women of all ages have notably higher incidence rates for estrogen receptor-negative tumors. These same patterns of disease are seen in contemporary female populations of western, sub-Saharan Africa, who are likely to share ancestry with African American women as a consequence of the Colonial-era slave trade. Interestingly, male breast cancer also is seen with increased frequency among both African Americans and Africans.

**Natural History**

Bloom and colleagues described the natural history of breast cancer based on the records of 250 women with untreated breast cancers who were cared for on charity wards in the Middlesex Hospital, London, between 1805 and 1933. The median survival of this population was 2.7 years after initial diagnosis (Fig. 17-13). The 5- and 10-year survival rates for these women were 18.0% and 3.6%, respectively. Only 0.8% survived for 15 years or longer. Autopsy data confirmed that 95% of these women died of breast cancer, whereas the remaining 5% died of other causes. Almost 75% of the women developed ulceration of the breast during the course of the disease. The longest surviving patient died in the 19th year after diagnosis.

**Primary Breast Cancer.** More than 80% of breast cancers show productive fibrosis that involves the epithelial and stromal tissues. With growth of the cancer and invasion of the surrounding breast tissues, the accompanying desmoplastic response entraps and shortens Cooper’s suspensory ligaments to produce a characteristic skin retraction. Localized edema (peau d’orange) develops when drainage of lymph fluid from the skin is disrupted. With continued growth, cancer cells invade the skin, and eventually ulceration occurs. As new areas of skin are invaded, small satellite nodules appear near the primary ulceration. The size of the primary breast cancer correlates with disease-free and overall survival, but there is a close association between cancer size and axillary lymph node involvement (Fig. 17-14). In general, up to 20% of breast cancer recurrences are local-regional, >60% are distant, and 20% are both local-regional and distant.
nodes adhere to each other and form a conglomerate mass. Cancer cells may grow through the lymph node capsule and fix to contiguous structures in the axilla, including the chest wall. Typically, axillary lymph nodes are involved sequentially from the low (level I) to the central (level II) to the apical (level III) lymph node groups. Approximately 95% of the women who die of breast cancer have distant metastases, and traditionally the most important prognostic correlate of disease-free and overall survival was axillary lymph node status (see Fig. 17-14A). Women with node-negative disease had less than a 30% risk of recurrence, compared with as much as a 75% risk for women with node-positive disease.

**Distant Metastases.** At approximately the 20th cell doubling, breast cancers acquire their own blood supply (neovascularization). Thereafter, cancer cells may be shed directly into the systemic venous blood to seed the pulmonary circulation via the axillary and intercostal veins or the vertebral column via Batson’s plexus of veins, which courses the length of the vertebral column. These cells are scavenged by natural killer lymphocytes and macrophages. Successful implantation of metastatic foci from breast cancer predictably occurs after the primary cancer exceeds 0.5 cm in diameter, which corresponds to the 27th cell doubling. For 10 years after initial treatment, distant metastases are the most common cause of death in breast cancer patients. For this reason, conclusive results cannot be derived from breast cancer trials until at least 5 to 10 years have elapsed. Although 60% of the women who develop distant metastases will do so within 60 months of treatment, metastases may become evident as late as 20 to 30 years after treatment of the primary cancer. Patients with estrogen receptor negative breast cancers are proportionately more likely to develop recurrence in the first 3 to 5 years, whereas those with estrogen receptor positive tumors have a risk of developing recurrence, which drops off more slowly beyond 5 years than is seen with ER-negative tumors. Recently, a report showed that tumor size and nodal status remain powerful predictors of late recurrences compared to more recently developed tools such as the immunohistochemical score (IHC4) and two gene expression profile tests (Recurrence Score and PAM50). Common sites of involvement, in order of frequency, are bone, lung, pleura, soft tissues, and liver. Brain metastases are less frequent overall, although with the advent of adjuvant systemic therapies it has been reported that CNS disease may be seen earlier. There are also reports of factors that are associated with the risk of developing brain metastases. For example, they are more likely to be seen in patients with triple receptor negative breast cancer (ER-negative, PR-negative, and HER2-negative) or patients with HER2-positive breast cancer who have received chemotherapy and HER2-directed therapies.

**HISTOPATHOLOGY OF BREAST CANCER**

**Carcinoma In Situ**

Cancer cells are in situ or invasive depending on whether or not they invade through the basement membrane. Broders’s original description of in situ breast cancer stressed the absence of invasion of cells into the surrounding stroma and their confinement within natural ductal and alveolar boundaries. Because areas of invasion may be minute, the accurate diagnosis of in situ cancer necessitates the analysis of multiple microscopic sections to exclude invasion. In 1941, Foote and Stewart published...
a landmark description of LCIS, which distinguished it from DCIS. 130 In the late 1960s, Gallagher and Martin published their study of whole-breast sections and described a stepwise progression from benign breast tissue to in situ cancer and subsequently to invasive cancer. Before the widespread use of mammography, diagnosis of breast cancer was by physical examination. At that time, in situ cancers constituted <6% of all breast cancers, and LCIS was more frequently diagnosed than DCIS by a ratio of >2:1. However, when screening mammography became popular, a 14-fold increase in the incidence of in situ cancer (45%) was demonstrated, and DCIS was more frequently diagnosed than LCIS by a ratio of >2:1. Table 17-8 lists the clinical and pathologic characteristics of DCIS and LCIS. Multicentricity refers to the occurrence of a second breast cancer outside the breast quadrant of the primary cancer (or at least 4 cm away), whereas multifocality refers to the occurrence of a second cancer within the same breast quadrant as the primary cancer (or within 4 cm of it). Multicentricity occurs in 60% to 90% of women with LCIS, whereas the rate of multicentricity for DCIS is reported to be 40% to 80%. LCIS occurs bilaterally in 50% to 70% of cases, whereas DCIS occurs bilaterally in 10% to 20% of cases.

**Lobular Carcinoma In Situ.** LCIS originates from the terminal duct lobular units and develops only in the female breast. It is characterized by distention and distortion of the terminal duct lobular units by cells that are large but maintain a normal normal to cytoplasmic ratio. Cytoplasmic mucoid globules are a distinctive cellular feature. LCIS may be observed in breast tissues that contain microcalcifications, but the calcifications associated with LCIS typically occur in adjacent tissues. This neighborhood calcification is a feature that is unique to LCIS and contributes to its diagnosis. The frequency of LCIS in the general population cannot be reliably determined because it usually presents as an incidental finding. The average age at diagnosis is 45 years, which is approximately 15 to 25 years younger than the age at diagnosis for invasive breast cancer. LCIS has a distinct racial predilection, occurring 12 times more frequently in white women than in African-American women. Invasive breast cancer develops in 25% to 35% of women with LCIS. Invasive cancer may develop in either breast, regardless of which breast harbored the initial focus of LCIS, and is detected synchronously with LCIS in 5% of cases. In women with a history of LCIS, up to 65% of subsequent invasive cancers are ductal, not lobular, in origin. For these reasons, LCIS is regarded as a marker of increased risk for invasive breast cancer rather than as an anatomic precursor. Individuals should be counseled regarding their risk of developing breast cancer and appropriate risk reduction strategies, including observation with screening, chemoprevention, and risk-reducing bilateral mastectomy.

**Ductal Carcinoma In Situ.** Although DCIS is predominantly seen in the female breast, it accounts for 5% of male breast cancers. Published series suggest a detection frequency of 7% in all biopsy tissue specimens. The term *intraductal carcinoma* is frequently applied to DCIS, which carries a high risk for progression to an invasive cancer. Histologically, DCIS is characterized by a proliferation of the epithelium that lines the minor ducts, resulting in papillary growths within the duct lumina. Early in their development, the cancer cells do not show pleomorphism, mitoses, or atypia, which leads to difficulty distinguishing early DCIS from benign hyperplasia. The papillary growths (papillary growth pattern) eventually coalesce and fill the duct lumina so that only scattered, rounded spaces remain between the clumps of atypical cancer cells, which show hyperchromasia and loss of polarity (cribriform growth pattern). Eventually pleomorphic cancer cells with frequent mitotic figures obliterate the lumina and distend the ducts (solid growth pattern). With continued growth, these cells outstrip their blood supply and become necrotic (comedo growth pattern). Calcium deposition occurs in the areas of necrosis and is a common feature seen on mammography. DCIS is now frequently classified based on nuclear grade and the presence of necrosis (Table 17-9). Based

### Table 17-8

| **Salient characteristics of in situ ductal (DCIS) and lobular (LCIS) carcinoma of the breast** |
|---------------------------------|------|------|
| **Age (years)** | LCIS | DCIS |
| 44–47 | 54–58 |
| **Incidence** | 2%–5% | 5%–10% |
| **Clinical signs** | None | Mass, pain, nipple discharge |
| **Mammographic signs** | None | Microcalcifications |
| **Premenopausal** | 2/3 | 1/3 |
| **Incidence of synchronous invasive carcinoma** | 5% | 2%–46% |
| **Multicentricity** | 60%–90% | 40%–80% |
| **Bilaterality** | 50%–70% | 10%–20% |
| **Axillary metastasis** | 1% | 1%–2% |
| **Subsequent carcinomas:** | | |
| **Incidence** | 25%–35% | 25%–70% |
| **Laterality** | Bilateral | Ipsilateral |
| **Interval to diagnosis** | 15–20 y | 5–10 y |
| **Histologic type** | Ductal | Ductal |

*In biopsy specimens of mammographically detected breast lesions.

### Table 17-9

| **Classification of breast ductal carcinoma in situ (DCIS)** |
|-------------------|------|------|------|
| **HISTOLOGIC SUBTYPE** | **DETERMINING CHARACTERISTICS** | **NUCLEAR GRADE** | **NECROSIS** | **DCIS GRADE** |
| Comedo | High | Extensive | High |
| Intermediate | Intermediate | Focal or absent | Intermediate |
| Noncomedo | Low | Absent | Low |

*Often a mixture of noncomedo patterns.
*Solid, cribriform, papillary, or focal micropapillary.
on multiple consensus meetings, grading of DCIS has been recommended. Although there is no universal agreement on classification, most systems endorse the use of cytologic grade and presence or absence of necrosis. The risk for invasive breast cancer is increased nearly fivefold in women with DCIS. The invasive cancers are observed in the ipsilateral breast, usually in the same quadrant as the DCIS that was originally detected, which suggests that DCIS is an anatomic precursor of invasive ductal carcinoma (Fig. 17-15A and B).

**Invasive Breast Carcinoma**

Invasive breast cancers have been described as lobular or ductal in origin. Early classifications used the term lobular to describe invasive cancers that were associated with LCIS, whereas all other invasive cancers were referred to as ductal.

**Figure 17-15.** Ductal carcinoma in situ (DCIS). A. Craniocaudal mammographic view shows a poorly defined mass containing microcalcifications. (Used with permission from Dr. Anne Turnbull, Consultant Radiologist/Director of Breast Screening, Royal Derby Hospital, Derby, UK.) B. Histopathologic preparation of the surgical specimen confirms DCIS with necrosis (100x). (Used with permission from Dr. Sindhu Menon, Consultant Histopathologist and Dr. Rahul Deb, Consultant Histopathologist and Lead Breast Pathologist, Royal Derby Hospital, Derby, UK.)

Current histologic classifications recognize special types of breast cancers (10% of total cases), which are defined by specific histologic features. To qualify as a special-type cancer, at least 90% of the cancer must contain the defining histologic features. About 80% of invasive breast cancers are described as invasive ductal carcinoma of no special type (NST). These cancers generally have a worse prognosis than special-type cancers. Foote and Stewart originally proposed the following classification for invasive breast cancer:

1. Paget’s disease of the nipple
2. Invasive ductal carcinoma—Adenocarcinoma with productive fibrosis (scirrhous, simplex, NST), 80%
3. Medullary carcinoma, 4%
4. Mucinous (colloid) carcinoma, 2%
5. Papillary carcinoma, 2%
6. Tubular carcinoma, 2%
7. Invasive lobular carcinoma, 10%
8. Rare cancers (adenoid cystic, squamous cell, apocrine)

Paget’s disease of the nipple was described in 1874. It frequently presents as a chronic, eczematous eruption of the nipple, which may be subtle but may progress to an ulcerated, weeping lesion. Paget’s disease usually is associated with extensive DCIS and may be associated with an invasive cancer. A palpable mass may or may not be present. A nipple biopsy specimen will show a population of cells that are identical to the underlying DCIS cells (pagetoid features or pagetoid change). Pathognomonic of this cancer is the presence of large, pale, vacuolated cells (Paget cells) in the rete pegs of the epithelium. Paget’s disease may be confused with superficial spreading melanoma. Differentiation from pagetoid intraepithelial melanoma is based on the presence of S-100 antigen immunostaining in melanoma and carcinoembryonic antigen immunostaining in Paget’s disease. Surgical therapy for Paget’s disease may involve lumpectomy or mastectomy, depending on the extent of involvement of the nipple-areolar complex and the presence of DCIS or invasive cancer in the underlying breast parenchyma.

Invasive ductal carcinoma of the breast with productive fibrosis (scirrhous, simplex, NST) accounts for 80% of breast cancers and presents with macroscopic or microscopic axillary lymph node metastases in up to 25% of screen-detected cases and up to 60% of symptomatic cases. This cancer occurs most frequently in perimenopausal or postmenopausal women in the fifth to sixth decades of life as a solitary, firm mass. It has poorly defined margins, and its cut surfaces show a central stellate configuration with chalky white or yellow streaks extending into surrounding breast tissues. The cancer cells often are arranged in small clusters, and there is a broad spectrum of histologic types with variable cellular and nuclear grades (Fig. 17-16A and B). In a large patient series from the SEER database, 75% of ductal cancers showed estrogen receptor expression.

Medullary carcinoma is a special-type breast cancer; it accounts for 4% of all invasive breast cancers and is a frequent phenotype of BRCA1 hereditary breast cancer. Grossly, the cancer is soft and hemorrhagic. A rapid increase in size may occur secondary to necrosis and hemorrhage. On physical examination, it is bulky and often positioned deep within the breast. Bilaterality is reported in 20% of cases. Medullary carcinoma is characterized microscopically by: (a) a dense lympho- phoreticular infiltrate composed predominantly of lymphocytes and plasma cells; (b) large pleomorphic nuclei that are poorly
differentiated and show active mitosis; and (c) a sheet-like growth pattern with minimal or absent ductal or alveolar differentiation. Approximately 50% of these cancers are associated with DCIS, which characteristically is present at the periphery of the cancer, and <10% demonstrate hormone receptors. In rare circumstances, mesenchymal metaplasia or anaplasia is noted. Because of the intense lymphocyte response associated with the cancer, benign or hyperplastic enlargement of the lymph nodes of the axilla may contribute to erroneous clinical staging.

Women with this cancer have a better 5-year survival rate than those with NST or invasive lobular carcinoma.

Mucinous carcinoma (colloid carcinoma), another special-type breast cancer, accounts for 2% of all invasive breast cancers and typically presents in the older population as a bulky tumor. This cancer is defined by extracellular pools of mucin, which surround aggregates of low-grade cancer cells. The cut surface of this cancer is glistening and gelatinous in quality. Fibrosis is variable, and when abundant it imparts a firm consistency to the cancer. Over 90% of mucinous carcinomas display hormone receptors. Lymph node metastases occur in 33% of cases, and 5- and 10-year survival rates are 73% and 59%, respectively. Because of the mucinous component, cancer cells may not be evident in all microscopic sections, and analysis of multiple sections is essential to confirm the diagnosis of a mucinous carcinoma.

Papillary carcinoma is a special-type cancer of the breast that accounts for 2% of all invasive breast cancers. It generally presents in the seventh decade of life and occurs in a disproportionate number of nonwhite women. Typically, papillary carcinomas are small and rarely attain a size of 3 cm in diameter. These cancers are defined by papillae with fibrovascular stalks and multilayered epithelium. In a large series from the SEER database 87% of papillary cancers have been reported to express estrogen receptor. McDivitt and colleagues noted that these tumors showed a low frequency of axillary lymph node metastases and had 5- and 10-year survival rates similar to those for mucinous and tubular carcinoma.

Tubular carcinoma is another special-type breast cancer and accounts for 2% of all invasive breast cancers. It is reported in as many as 20% of women whose cancers are diagnosed by mammographic screening and usually is diagnosed in the perimenopausal or early menopausal periods. Under low-power magnification, a haphazard array of small, randomly arranged tubular elements is seen. In a large SEER database 94% of tubular cancers were reported to express estrogen receptor. Approximately 10% of women with tubular carcinoma or with invasive cribriform carcinoma, a special-type cancer closely related to tubular carcinoma, will develop axillary lymph node metastases. However, the presence of metastatic disease in one or two axillary lymph nodes does not adversely affect survival. Distant metastases are rare in tubular carcinoma and invasive cribriform carcinoma. Long-term survival approaches 100%.

Invasive lobular carcinoma accounts for 10% of breast cancers. The histopathologic features of this cancer include small cells with rounded nuclei, inconspicuous nucleoli, and scant cytoplasm (Fig. 17-17). Special stains may confirm the
presence of intracytoplasmic mucin, which may displace the nucleus (signet-ring cell carcinoma). At presentation, invasive lobular carcinoma varies from clinically inapparent carcinomas to those that replace the entire breast with a poorly defined mass. It is frequently multifocal, multicentric, and bilateral. Because of its insidious growth pattern and subtle mammographic features, invasive lobular carcinoma may be difficult to detect. Over 90% of lobular cancers express estrogen receptor.¹³³

**DIAGNOSIS OF BREAST CANCER**

In ∼30% of cases, the woman discovers a lump in her breast. Other less frequent presenting signs and symptoms of breast cancer include: (a) breast enlargement or asymmetry; (b) nipple changes, retraction, or discharge; (c) ulceration or erythema of the skin of the breast; (d) an axillary mass; and (e) musculoskeletal discomfort. However, up to 50% of women presenting with breast complaints have no physical signs of breast pathology. Breast pain usually is associated with benign disease.

Misdiagnosed breast cancer accounts for the greatest number of malpractice claims for errors in diagnosis and for the largest number of paid claims. Litigation often involves younger women, whose physical examination and mammogram may be misleading. If a young woman (≤45 years) presents with a palpable breast mass and equivocal mammographic findings, ultrasound examination and biopsy are used to avoid a delay in diagnosis.

**Examination**

**Inspection.** The clinician inspects the woman’s breast with her arms by her side (Fig. 17-18A), with her arms straight up in the air (Fig. 17-18B), and with her hands on her hips (with and without pectoral muscle contraction).¹³⁵,¹³⁶ Symmetry, size, and shape of the breast are recorded, as well as any evidence of edema (peau d’orange), nipple or skin retraction, or erythema. With the arms extended forward and in a sitting position, the woman leans forward to accentuate any skin retraction.

**Figure 17-18.** Examination of the breast. **A.** Inspection of the breast with arms at sides. **B.** Inspection of the breast with arms raised. **C.** Palpation of the breast with the patient supine. **D.** Palpation of the axilla.

**Palpation.** As part of the physical examination, the breast is carefully palpated. With the patient in the supine position (see Fig. 17-18C) the clinician gently palpates the breasts, making certain to examine all quadrants of the breast from the sternum laterally to the latissimus dorsi muscle and from the clavicle inferiorly to the upper rectus sheath. The examination is performed with the palmar aspects of the fingers, avoiding a grasping or pinching motion. The breast may be cupped or molded in the examiner’s hands to check for retraction. A systematic search for lymphadenopathy then is performed. Figure 17-18D shows the position of the patient for examination of the axilla. By supporting the upper arm and elbow, the examiner stabilizes the shoulder girdle. Using gentle palpation, the clinician assesses all three levels of possible axillary lymphadenopathy. Careful palpation of supraclavicular and parasternal sites also is performed. A diagram of the chest and contiguous lymph node sites is useful for recording location, size, consistency, shape, mobility, fixation, and other characteristics of any palpable breast mass or lymphadenopathy (Fig. 17-19).

**Imaging Techniques**

**Mammography.** Mammography has been used in North America since the 1960s, and the techniques used continue to be modified and improved to enhance image quality.¹³⁷-¹⁴⁰ Conventional mammography delivers a radiation dose of 0.1 cGy per study. By comparison, chest radiography delivers 25% of this dose. However, there is no increased breast cancer risk associated with the radiation dose delivered with screening mammography. Screening mammography is used to detect unexpected breast cancer in asymptomatic women. In this regard, it supplements history taking and physical examination. With screening mammography, two views of the breast are obtained: the craniocaudal (CC) view (Fig. 17-20A,B) and the mediolateral oblique (MLO) view (Fig. 17-20C,D). The MLO view images the greatest volume of breast tissue, including the upper outer quadrant and the axillary tail of Spence. Compared with the MLO view, the CC view provides better visualization of the medial aspect of the breast and permits greater breast compression. Diagnostic mammography is used to evaluate women with abnormal
Figure 17-20. A-D. Mammogram of a premenopausal breast with a dense fibroglandular pattern. E-H. Mammogram of a postmenopausal breast with a sparse fibroglandular pattern. (Used with permission from Dr. Anne Turnbull, Consultant Radiologist/Director of Breast Screening, Royal Derby Hospital, Derby, UK.)
findings such as a breast mass or nipple discharge. In addition to the MLO and CC views, a diagnostic examination may use views that better define the nature of any abnormalities, such as the 90° lateral and spot compression views. The 90° lateral view is used along with the CC view to triangulate the exact location of an abnormality. Spot compression may be done in any projection by using a small compression device, which is placed directly over a mammographic abnormality that is obscured by overlying tissues (Fig. 17-21C). The compression device minimizes motion artifact, improves definition, separates overlying tissues, and decreases the radiation dose needed to penetrate the breast. Magnification techniques (×1.5) often are combined with spot compression to better resolve calcifications and the margins of masses. Mammography also is used to guide interventional procedures, including needle localization and needle biopsy.
Specific mammographic features that suggest a diagnosis of breast cancer include a solid mass with or without stellate features, asymmetric thickening of breast tissues, and clustered microcalcifications. The presence of fine, stippled calcium in and around a suspicious lesion is suggestive of breast cancer and occurs in as many as 50% of nonpalpable cancers. These microcalcifications are an especially important sign of cancer in younger women, in whom it may be the only mammographic abnormality. The clinical impetus for screening mammography came from the Health Insurance Plan study and the Breast Cancer Detection Demonstration Project, which demonstrated a 33% reduction in mortality for women after screening mammography. Mammography was more accurate than clinical examination for the detection of early breast cancers, providing a true-positive rate of 90%. Only 20% of women with nonpalpable cancers had axillary lymph node metastases, compared with 50% of women with palpable cancers. Current guidelines of the National Comprehensive Cancer Network suggest that normal-risk women ≥20 years of age should have a breast examination at least every 3 years. Starting at age 40 years, breast examinations should be performed yearly, and a yearly mammogram should be taken. Screening mammography in women ≥50 years of age is more controversial for previously noted reasons: (a) reduced sensitivity, (b) reduced specificity, and (c) lower incidence of breast cancer. Because of the combination of these three reasons, targeting mammography screening to women <50 years of age, who are at higher risk of breast cancer, improves the balance of risks and benefits and is the approach some health care systems have taken. There are now a number of risk assessment models—as described earlier in this chapter—that can be used to estimate a younger woman’s risk of developing breast cancer and that help assess the risks and benefits of regular screening.

Screen film mammography has replaced xeromammography because it requires a lower dose of radiation and provides similar image quality. Digital mammography was developed to allow the observer to manipulate the degree of contrast in the image. This is especially useful in women with dense breasts and women <50 years of age. Recently, investigators directly compared digital vs. screen film mammography in a prospective (DMIST) trial that enrolled over 42,000 women. The investigators found that digital and screen film mammography had similar accuracy; however, digital mammography was more accurate in women <50 years of age, women with mammographically dense breasts, and premenopausal or perimenopausal women. The use of digital breast tomosynthesis with 3D images has been introduced as an alternative to standard 2D mammography imaging that is limited by superimposition of breast parenchyma and breast density. The STORM trial reported that in 7,292 women screened, 3D mammography had a higher cancer detection rate and fewer false-positive recalls than the standard 2D imaging. Randomized controlled trials are planned to further study tomosynthesis and its role in breast cancer screening. Standard two-dimensional mammography has limitations,
such as the parenchymal density or superimposition of breast tissue, which obscures cancers or causes normal structures to appear suspicious reducing the sensitivity of mammography and increasing the false-positive rates. Digital breast tomosynthesis is a technology developed to assist with overcoming these limitations. In digital breast tomosynthesis, multiple projection images are reconstructed to allow visual review of thin breast sections, each reconstructed slice as thin as 0.5 mm, which provides better characterization of noncalcified lesions. These multiple projection exposures are obtained by a digital detector from a mammography X-ray source that moves through a limited arc angle while the breast is compressed. Then these projection image data sets are reconstructed using specific algorithms, which provide the clinical reader a series of images through the entire breast.148

In 2011, tomosynthesis was approved by the U.S. Food and Drug Administration (FDA) to be used in combination with standard digital mammography for breast cancer screening. The total radiation dose when tomosynthesis is added is about twice the current dose of digital mammography alone but remains below the limits set by the FDA.149

The STORM-2 trial reported that synthetic 2D-3D mammography yields similar breast cancer detection as dual-acquisition 2D-3D mammography with the advantage of reducing radiation exposure.150

Contrast-enhanced digital mammography (CEDM) was also approved by the FDA in 2001, which utilizes an iodinated contrast material and modified digital mammography units for imaging.148 CEDM has been shown to be feasible and detects breast cancers at a rate similar to MRI, which has potential to offer an alternative to MRI.151 The advantages of CEDM over MRI are that the use of compression limits motion, there is decrease in cost, decrease in exam time, and there is an option for patients who are unable to tolerate MRI or who due to various reasons cannot have MRI due to incompatibility, such as the presence of a pacemaker or tissue expanders.148,152

Ductography. The primary indication for ductography is nipple discharge, particularly when the fluid contains blood. Radiopaque contrast media is injected into one or more of the major ducts, and mammography is performed. A duct is gently enlarged with a dilator, and then a small, blunt cannula is inserted under sterile conditions into the nipple ampulla. With the patient in a supine position, 0.1 to 0.2 mL of dilute contrast media is injected, and CC and MLO mammographic views are obtained without compression. Intraductal papillomas are seen as small filling defects surrounded by contrast media (Fig. 17-22). Cancers may appear as irregular masses or as multiple intraluminal filling defects.

Ultrasonography. Second only to mammography in frequency of use for breast imaging, ultrasonography is an important method of resolving equivocal mammographic findings, defining cystic masses, and demonstrating the echogenic qualities of specific solid abnormalities. On ultrasound examination, breast cysts are well circumscribed, with smooth margins and an echo-free center (Fig. 17-23). Benign breast masses usually show smooth contours, round or oval shapes, weak internal echoes, and well-defined anterior and posterior margins (Fig. 17-24). Breast cancer characteristically has irregular walls (Fig. 17-25) but may have smooth margins with acoustic enhancement. Ultrasonography is used to guide fine-needle aspiration biopsy, core-needle biopsy, and needle localization of breast lesions. Its findings are highly reproducible, and it has a high patient acceptance rate, but it does not reliably detect lesions that are ≤1 cm in diameter. Ultrasonography can also be utilized to image the regional lymph nodes in patients with breast cancer (Fig. 17-26). The sensitivity of examination for the status of axillary nodes ranges from 35% to 82% and specificity ranges from 73% to 97%. The features of a lymph node involved with cancer include cortical thickening, change in shape of the node to more circular appearance, size larger
Figure 17-24. Ultrasonography images of benign breast tumors. A. Fibroadenoma. B. Intraductal papilloma (see arrow). (Used with permission from Dr. Anne Turnbull, Consultant Radiologist/Director of Breast Screening, Royal Derby Hospital, Derby, UK.)

than 10 mm, absence of a fatty hilum and hypoechoic internal echoes.153

**Magnetic Resonance Imaging.** In the process of evaluating magnetic resonance imaging (MRI) as a means of characterizing mammographic abnormalities, additional breast lesions have been detected. However, in the circumstance of negative findings on both mammography and physical examination, the probability of a breast cancer being diagnosed by MRI is extremely low. There is current interest in the use of MRI to screen the breasts of high-risk women and of women with a newly diagnosed breast cancer. In the first case, women who have a strong family history of breast cancer or who carry known genetic mutations require screening at an early age because mammographic evaluation is limited due to the increased breast density in younger women. In the second case, an MRI study of the contralateral breast in women with a known breast cancer has shown a contralateral breast cancer in 5.7% of these women (Fig. 17-27). MRI can also detect additional tumors in the index breast (multifocal or multicentric disease) that may be missed on routine breast imaging and this may alter surgical decision making (Fig. 17-28). In fact, MRI has been advocated by some for routine use in surgical treatment planning based on the fact that additional disease can be identified with this advanced imaging modality and the

Figure 17-23. Breast cyst. A. Simple cyst. Ultrasound image of the mass shows it to be anechoic with a well-defined back wall, characteristic of a cyst. B. Complex solid and cystic mass. C. Complex solid and cystic mass characteristic of intracystic papillary tumor. (Used with permission from Dr. Anne Turnbull, Consultant Radiologist/Director of Breast Screening, Royal Derby Hospital, Derby, UK.)
extent of disease may be more accurately assessed. A randomized trial performed in the United Kingdom (COMICE trial) that enrolled 1623 women did not show a decrease in rates of reoperation in those women randomized to undergo MRI in addition to mammography and ultrasonography (19%) compared to those undergoing standard breast imaging without MRI (19%). Houssami and colleagues performed a meta-analysis including two randomized trials and seven comparative cohort studies to examine the effect of preoperative MRI compared to standard preoperative evaluation on surgical outcomes. They reported that the use of MRI was associated with increased mastectomy rates. This is problematic because there is no evidence that the additional disease detected by MRI is of clinical or biologic significance, particularly in light of the low local-regional failure rates currently reported in patients undergoing breast conserving surgery who receive whole breast irradiation and systemic therapies. There is an ongoing trial in the Alliance for Clinical Trials in Oncology that is randomizing patients to preoperative MRI vs. standard imaging to assess the impact of MRI on local regional recurrence rates in patients with triple receptor negative and HER2 positive breast cancers.

The use of dedicated breast coils is mandatory in the MRI imaging of the breast. A BIRADS lexicon is assigned to each examination and an abnormality noted on MRI that is not seen on mammography requires a focused ultrasound examination for further assessment. If the abnormality is not seen on corresponding mammogram or ultrasound, then MRI-guided biopsy is necessary. Some clinical scenarios where MRI may be useful include the evaluation of a patient who presents with nodal metastasis from breast cancer without an identifiable primary tumor; to assess response to therapy in the setting of neoadjuvant

Figure 17-25. Ultrasonography images of malignant breast lesions. A. 25 mm irregular mass. B. Ultrasound 30 mm mass anterior to an implant. C. Ultrasound breast cancer with calcification. D. Ultrasound shows a 9 mm spiculated mass (see arrow) with attenuation. (Used with permission from Dr. Anne Turnbull, Consultant Radiologist/Director of Breast Screening, Royal Derby Hospital, Derby, UK.)
systemic treatment; to select patients for partial breast irradiation techniques; and evaluation of the treated breast for tumor recurrence.

**Breast Biopsy**

**Nonpalpable Lesions.** Image-guided breast biopsy specimens are frequently required to diagnose nonpalpable lesions. Ultrasound localization techniques are used when a mass is present, whereas stereotactic techniques are used when no mass is present (microcalcifications or architectural distortion only). The combination of diagnostic mammography, ultrasound or stereotactic localization, and fine-needle aspiration (FNA) biopsy achieves almost 100% accuracy in the preoperative diagnosis of breast cancer. However, although FNA biopsy permits cytologic evaluation, core-needle permits the analysis of breast tissue architecture and allows the pathologist to determine whether invasive cancer is present. This permits the surgeon and patient to discuss the specific management of a breast cancer before therapy begins. Core-needle biopsy is preferred over open biopsy for nonpalpable breast lesions because a single surgical procedure can be planned based on the results of the core biopsy. The advantages of core-needle biopsy include a low complication rate, minimal scarring, and a lower cost compared with excisional breast biopsy.

**Palpable Lesions.** FNA or core biopsy of a palpable breast mass can usually be performed in an outpatient setting. A 1.5-in, 22-gauge needle attached to a 10-mL syringe or a 14-gauge core biopsy needle is used. For FNA, use of a syringe holder...

---

**Figure 17-26.** Ultrasonography images of lymph nodes. A. Normal axillary lymph node (see arrows). B. Indeterminate axillary lymph node. C. Malignant appearing axillary lymph node. (Used with permission from Dr. Anne Turnbull, Consultant Radiologist/Director of Breast Screening, Royal Derby Hospital, Derby, UK.)

**Figure 17-27.** MRI examination revealing contralateral breast cancer (see arrows) in a patient diagnosed with unilateral breast cancer on mammography (two arrows). (Used with permission from Dr. Anne Turnbull, Consultant Radiologist/Director of Breast Screening, Royal Derby Hospital, Derby, UK.)
enables the surgeon performing the FNA biopsy to control the syringe and needle with one hand while positioning the breast mass with the opposite hand. After the needle is placed in the mass, suction is applied while the needle is moved back and forth within the mass. Once cellular material is seen at the hub of the needle, the suction is released and the needle is withdrawn. The cellular material is then expressed onto microscope slides. Both air-dried and 95% ethanol–fixed microscopic sections are prepared for analysis. When a breast mass is clinically and mammographically suspicious, the sensitivity and specificity of FNA biopsy approaches 100%. Core-needle biopsy of palpable breast masses is performed using a 14-gauge needle, such as the Tru-Cut needle. Automated devices also are available. Vacuum-assisted core biopsy devices (with 8–10 gauge needles) are commonly utilized with image guidance where between 4 and 12 samples can be acquired at different positions within a mass, area of architectural distortion or microcalcifications. If the target lesion was microcalcifications, the specimen should be radiographed to confirm appropriate sampling. A radiopaque marker should be placed at the site of the biopsy to mark the area for future intervention. In some cases the entire lesion is removed with the biopsy technique and clip placement allows for accurate targeting of the site for surgical resection. Tissue specimens are placed in formalin and then processed to paraffin blocks. Although the false-negative rate for core-needle biopsy specimens is very low, a tissue specimen that does not show breast cancer cannot conclusively rule out that diagnosis because a sampling error may have occurred. The clinical, radiographic, and pathologic findings should be in concordance. If the biopsy findings do not concur with the clinical and radiographic findings, the multidisciplinary team (including clinician, radiologist, and pathologist) should review the findings and decide whether or not to recommend an image-guided or open biopsy to be certain that the target lesion has been adequately sampled for diagnosis.

BREAST CANCER STAGING AND BIOMARKERS

Breast Cancer Staging
The clinical stage of breast cancer is determined primarily through physical examination of the skin, breast tissue, and regional lymph nodes (axillary, supraclavicular, and internal mammary). However, clinical determination of axillary lymph node metastases has an accuracy of only 33%. Ultrasound (US) is more sensitive than physical examination alone in determining axillary lymph node involvement during preliminary staging of breast carcinoma. FNA or core biopsy of sonographically indeterminate or suspicious lymph nodes can provide a more definitive diagnosis than US alone. Pathologic stage combines the findings from pathologic examination of the resected primary breast cancer and axillary or other regional lymph nodes. Fisher and colleagues found that accurate predictions regarding the occurrence of distant metastases were possible after resection and pathologic analysis of 10 or more levels I and II axillary lymph nodes. A frequently used staging system is the TNM (tumor, nodes, and metastasis) system. The American Joint Committee on Cancer (AJCC) has recently modified the TNM system for breast cancer to include both anatomic and biologic factors (Tables 17-10 and 17-11). Koscielny and colleagues demonstrated that tumor size correlates with the presence of axillary lymph node metastases (see Fig. 17-14B). Others have shown an association between tumor size, axillary lymph node metastases, and disease-free survival. One of the most important predictors of 10- and 20-year survival rates in breast cancer is the number of axillary lymph nodes involved with metastatic disease. Routine biopsy of internal mammary lymph nodes is not generally performed; however, it has been reported that in the context of a “triple node” biopsy approach either the internal mammary node or a low axillary node when positive alone carried the same prognostic weight. When both nodes were positive, the prognosis declined to the level associated with apical node positivity. A double node biopsy of the low axillary node and either the apical or the internal mammary node gave the same maximum prognostic information as a triple node biopsy. With the advent of sentinel lymph node dissection and the use of preoperative lymphoscintigraphy for localization of the sentinel nodes, surgeons have again begun to biopsy the internal mammary nodes but in a more targeted manner. The 8th edition of the AJCC staging system does allow for staging based on findings from the internal mammary sentinel nodes. Drainage to the internal mammary nodes is more frequent with central and medial quadrant cancers. Clinical or pathologic evidence of metastatic spread to supraclavicular lymph nodes is no longer considered stage IV disease, but routine scalene or supraclavicular lymph node biopsy is not indicated.

Biomarkers
Breast cancer biomarkers are of several types. Risk factor biomarkers are those associated with increased cancer risk. These include familial clustering and inherited germline abnormalities, proliferative breast disease with atypia, and mammographic density. Exposure biomarkers are a subset of risk factors that include measures of carcinogen exposure such as DNA adducts. Surrogate endpoint biomarkers are biologic alterations in tissue that occur between cancer initiation and development. These biomarkers are used as endpoints in short-term chemoprevention trials and include histologic changes, indices of proliferation, and genetic alterations leading to cancer. Prognostic biomarkers provide information regarding
### Primary tumor (T)

The T classification of the primary tumor is the same regardless of whether it is based on clinical or pathologic criteria, or both. Size should be measured to the nearest millimeter. If the tumor size is slightly less than or greater than a cutoff for a given T classification, it is recommended that the size be rounded to the millimeter reading that is closest to the cutoff. For example, a reported size of 1.1 mm is reported as 1 mm, or a size of 2.01 cm is reported as 2.0 cm. Designation should be made with the subscript “c” or “p” modifier to indicate whether the T classification was determined by clinical (physical examination or radiologic) or pathologic measurements, respectively. In general, pathologic determination should take precedence over clinical determination of T size.

<table>
<thead>
<tr>
<th>TX</th>
<th>Primary tumor cannot be assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>T0</td>
<td>No evidence of primary tumor</td>
</tr>
<tr>
<td>Tis (DCIS)*</td>
<td>Ductal carcinoma in situ</td>
</tr>
<tr>
<td>Tis (Paget)</td>
<td>Paget disease of the nipple NOT associated with invasive carcinoma and/or carcinoma in situ (DCIS) in the underlying breast parenchyma. Carcinomas in the breast parenchyma associated with Paget disease are categorized based on the size and characteristics of the parenchymal disease, although the presence of Paget disease should still be noted.</td>
</tr>
<tr>
<td>T1</td>
<td>Tumor ≤20 mm in greatest dimension</td>
</tr>
<tr>
<td>T1mi</td>
<td>Tumor ≤1 mm in greatest dimension</td>
</tr>
<tr>
<td>T1a</td>
<td>Tumor &gt;1 mm but ≤5 mm in greatest dimension (round any measurement &gt;1.0–1.9 mm to 2 mm).</td>
</tr>
<tr>
<td>T1b</td>
<td>Tumor &gt;5 mm but ≤10 mm in greatest dimension</td>
</tr>
<tr>
<td>T1c</td>
<td>Tumor &gt;10 mm but ≤20 mm in greatest dimension</td>
</tr>
<tr>
<td>T2</td>
<td>Tumor &gt;20 mm but ≤50 mm in greatest dimension</td>
</tr>
<tr>
<td>T3</td>
<td>Tumor &gt;50 mm in greatest dimension</td>
</tr>
<tr>
<td>T4</td>
<td>Tumor of any size with direct extension to the chest wall and/or to the skin (ulceration or macroscopic nodules); invasion of the dermis alone does not qualify as T4</td>
</tr>
<tr>
<td>T4a</td>
<td>Extension to the chest wall; invasion or adherence to pectoralis muscle in the absence of invasion of chest wall structures does not qualify as T4</td>
</tr>
<tr>
<td>T4b</td>
<td>Ulceration and/or ipsilateral macroscopic satellite nodules and/or edema (including peau d’orange) of the skin that does not meet the criteria for inflammatory carcinoma</td>
</tr>
<tr>
<td>T4c</td>
<td>Both T4a and T4b are present</td>
</tr>
<tr>
<td>T4d</td>
<td>Inflammatory carcinoma (see section “Rules for Classification”)</td>
</tr>
</tbody>
</table>

*Note: Lobular carcinoma in situ (LCIS) is a benign entity and is removed from TNM staging in the AJCC Cancer Staging Manual, 8th edition.

### Regional lymph nodes—Clinical (N)

| cNX* | Regional lymph nodes cannot be assessed (e.g., previously removed) |
| cN0 | No regional lymph node metastases (by imaging or clinical examination) |
| cN1 | Metastases to movable ipsilateral Level I, II axillary lymph node(s) |
| cN1mi** | Micrometastases (approximately 200 cells, larger than 0.2 mm, but none larger than 2.0 mm) |
| cN2 | Metastases in ipsilateral Level I, II axillary lymph nodes that are clinically fixed or matted; or in ipsilateral internal mammary nodes in the absence of axillary lymph node metastases |
| cN2a | Metastases in ipsilateral Level I, II axillary lymph nodes fixed to one another (matted) or to other structures |
| cN2b | Metastases only in ipsilateral internal mammary nodes in the absence of axillary lymph node metastases |
| cN3 | Metastases in ipsilateral infraclavicular (Level III axillary) lymph node(s) with or without Level I, II axillary lymph node involvement; or in ipsilateral internal mammary lymph node(s) with Level I, II axillary lymph node metastases; or metastases in ipsilateral supraclavicular lymph node(s) with or without axillary or internal mammary lymph node involvement |
| cN3a | Metastases in ipsilateral infraclavicular lymph node(s) |
| cN3b | Metastases in ipsilateral infraclavicular lymph node(s) |
| cN3c | Metastases in ipsilateral internal mammary lymph node(s) and axillary lymph node(s) |
| cN3d | Metastases in ipsilateral supraclavicular lymph node(s) |

*Note: (sn) and (f) suffixes should be added to the N category to denote confirmation of metastasis by sentinel node biopsy or fine needle aspiration/core needle biopsy respectively.

*the cNX category is used sparingly in cases where regional lymph nodes have previously been surgically removed or where there is no documentation of physical examination of the axilla.

**cN1mi is rarely used but may be appropriate in cases where sentinel node biopsy is performed before tumor resection, most likely to occur in cases treated with neoadjuvant therapy.

(Continued)
### Table 17-10

**TNM staging system for breast cancer (Continued)**

<table>
<thead>
<tr>
<th>Regional lymph nodes—Pathologic (pN)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pNX</td>
<td>Regional lymph nodes cannot be assessed (e.g., not removed for pathological study or previously removed)</td>
</tr>
<tr>
<td>pN0</td>
<td>No regional lymph node metastasis identified or ITCs only</td>
</tr>
<tr>
<td>pN0(i+)</td>
<td>ITCs only (malignant cell clusters no larger than 0.2 mm) in regional lymph node(s)</td>
</tr>
<tr>
<td>pN0(mol+)</td>
<td>Positive molecular findings by reverse transcriptase polymerase chain reaction (RT-PCR); no ITCs detected</td>
</tr>
<tr>
<td>pN1</td>
<td>Micrometastases; or metastases in 1–3 axillary lymph nodes; and/or clinically negative internal mammary nodes with micrometastases or macrometastases by sentinel lymph node biopsy</td>
</tr>
<tr>
<td>pN1mi</td>
<td>Micrometastases (approximately 200 cells, larger than 0.2 mm, but none larger than 2.0 mm)</td>
</tr>
<tr>
<td>pN1a</td>
<td>Metastases in 1–3 axillary lymph nodes, at least one metastasis larger than 2.0 mm</td>
</tr>
<tr>
<td>pN1b</td>
<td>Metastases in ipsilateral internal mammary sentinel nodes, excluding ITCs</td>
</tr>
<tr>
<td>pN1c</td>
<td>pN1a and pN1b combined</td>
</tr>
<tr>
<td>pN2</td>
<td>Metastases in 4–9 axillary lymph nodes; or positive ipsilateral internal mammary lymph nodes by imaging in the absence of axillary lymph node metastases</td>
</tr>
<tr>
<td>pN2a</td>
<td>Metastases in 4–9 axillary lymph nodes (at least one tumor deposit larger than 2.0 mm)</td>
</tr>
<tr>
<td>pN2b</td>
<td>Metastases in clinically detected internal mammary lymph nodes with or without microscopic confirmation; with pathologically negative axillary nodes</td>
</tr>
<tr>
<td>pN3</td>
<td>Metastases in 10 or more axillary lymph nodes; or in infraclavicular (Level III axillary) lymph nodes; or positive ipsilateral internal mammary lymph nodes by imaging in the presence of one or more positive Level I, II axillary lymph nodes; or in more than three axillary lymph nodes and micrometastases or macrometastases by sentinel lymph node biopsy in clinically negative ipsilateral internal mammary lymph nodes;</td>
</tr>
<tr>
<td>pN3a</td>
<td>or in ipsilateral supraclavicular lymph nodes</td>
</tr>
<tr>
<td>pN3b</td>
<td>Metastases in 10 or more axillary lymph nodes (at least one tumor deposit larger than 2.0 mm); or metastases to the infraclavicular (Level III axillary) lymph nodes</td>
</tr>
<tr>
<td>pN3c</td>
<td>or pN2a in the presence of cN2b (positive internal mammary lymph nodes by imaging); or pN2a in the presence of pNlb</td>
</tr>
<tr>
<td>pN3d</td>
<td>Metastases in ipsilateral supraclavicular lymph nodes</td>
</tr>
</tbody>
</table>

**Distant metastasis (M)**

<table>
<thead>
<tr>
<th>M0</th>
<th>No clinical or radiographic evidence of distant metastases*</th>
</tr>
</thead>
<tbody>
<tr>
<td>cM0(i+)</td>
<td>No clinical or radiographic evidence of distant metastases in the presence of tumor cells or deposits no larger than 0.2 mm detected microscopically or by molecular techniques in circulating blood, bone marrow, or other nonregional nodal tissue in a patient without symptoms or signs of metastases</td>
</tr>
<tr>
<td>cM1</td>
<td>Distant metastases detected by clinical and radiographic means</td>
</tr>
<tr>
<td>pM1</td>
<td>Any histologically proven metastases in distant organs; or if in non-regional nodes, metastases greater than 0.2 mm</td>
</tr>
</tbody>
</table>

*Note: (sn) and (f) suffixes should be added to the N category to denote confirmation of metastasis by sentinel node biopsy or FNA/core needle biopsy respectively, with NO further resection of nodes.

---

Cancer outcome irrespective of therapy, whereas predictive biomarkers provide information regarding response to therapy. Candidate prognostic and predictive biomarkers and biologic targets for breast cancer include (a) the steroid hormone receptor pathway; (b) growth factors and growth factor receptors such as human epidermal growth factor receptor 2 (HER2)/neu, epidermal growth factor receptor (EGFR), transforming growth factor, platelet-derived growth factor, and the insulin-like growth factor family; (c) indices of proliferation such as proliferating cell nuclear antigen (PCNA) and Ki-67; (d) indices of angiogenesis such as vascular endothelial growth factor (VEGF) and the angiogenesis index; (e) the mammalian target of rapamycin (mTOR) signaling pathway; (f) tumor-suppressor genes such as p53; (g) the cell cycle, cyclins, and cyclin-dependent kinases; (h) the proteasome; (i) the COX-2 enzyme; (j) the peroxisome proliferator-activated receptors (PPARs); and (k) indices of apoptosis and apoptosis modulators such as bcl-2 and the bax:bcl-2 ratio.

**Steroid Hormone Receptor Pathway.** Hormones play an important role in the development and progression of breast cancer. Estrogens, estrogen metabolites, and other steroid hormones such as progesterone all have been shown to have an effect. Breast cancer risk is related to estrogen exposure over time. In postmenopausal women, hormone replacement therapy consisting of estrogen plus progesterone increases the risk of breast cancer by 26% compared to placebo. Patients with hormone receptor-positive tumors survive two to three times longer after a diagnosis of metastatic disease than do patients with hormone receptor-negative tumors. Patients with tumors negative for both estrogen receptors and progesterone receptors are not considered candidates for hormonal therapy. Tumors positive
**Table 17-11**

<table>
<thead>
<tr>
<th>TNM stage groupings</th>
<th>When T is...</th>
<th>And N is...</th>
<th>And M is...</th>
<th>Then the stage group is...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tis</td>
<td>N0</td>
<td>M0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>N0</td>
<td>M0</td>
<td>IA</td>
<td></td>
</tr>
<tr>
<td>T0</td>
<td>N1mi</td>
<td>M0</td>
<td>IB</td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>N1mi</td>
<td>M0</td>
<td>IB</td>
<td></td>
</tr>
<tr>
<td>T0</td>
<td>N1</td>
<td>M0</td>
<td>II A</td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>N1</td>
<td>M0</td>
<td>II A</td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>N0</td>
<td>M0</td>
<td>II A</td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>N1</td>
<td>M0</td>
<td>II B</td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>N0</td>
<td>M0</td>
<td>II B</td>
<td></td>
</tr>
<tr>
<td>T0</td>
<td>N2</td>
<td>M0</td>
<td>III A</td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>N2</td>
<td>M0</td>
<td>III A</td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>N2</td>
<td>M0</td>
<td>III A</td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>N1</td>
<td>M0</td>
<td>III A</td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>N2</td>
<td>M0</td>
<td>III A</td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>N0</td>
<td>M0</td>
<td>III B</td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>N1</td>
<td>M0</td>
<td>III B</td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>N2</td>
<td>M0</td>
<td>III B</td>
<td></td>
</tr>
<tr>
<td>Any T</td>
<td>N3</td>
<td>M0</td>
<td>IIIC</td>
<td></td>
</tr>
<tr>
<td>Any T</td>
<td>Any N</td>
<td>M1</td>
<td>IV</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
1. T1 includes T1 mi.
2. T0 and T1 tumors with nodal micrometastases (N1mi) are staged as Stage IB.
3. T2, T3, and T4 tumors with nodal micrometastases (N1mi) are staged using the N1 category.
4. M0 includes M0(i+).
5. The designation pM0 is not valid; any M0 is clinical.
6. If a patient presents with M1 disease prior to neoadjuvant systemic therapy, the stage is Stage IV and remains Stage IV regardless of response to neoadjuvant therapy.
7. Stage designation may be changed if postsurgical imaging studies reveal the presence of distant metastases, provided the studies are performed within 4 months of diagnosis in the absence of disease progression, and provided the patient has not received neoadjuvant therapy.
8. Staging following neoadjuvant therapy is denoted with a “yc” or “yp” prefix to the T and N classification. There is no anatomic stage group assigned if there is a complete pathological response (pCR) to neoadjuvant therapy, for example, ypT0ypN0cM0.


For estrogen or progesterone receptors, a higher response rate to endocrine therapy than tumors that do not express estrogen or progesterone receptors. The determination of estrogen and progesterone receptor status used to require biochemical evaluation of fresh tumor tissue. Today, however, estrogen and progesterone receptor status can be measured in archived tissue using immunohistochemical techniques. Hormone receptor status also can be measured in specimens obtained with fine-needle aspiration biopsy or core-needle biopsy, and this can help guide treatment planning. Testing for estrogen and progesterone receptors should be performed on all primary invasive breast cancer specimens. The tumor hormone receptor status should be ascertained for both premenopausal and postmenopausal patients to identify patients who are most likely to benefit from endocrine therapy.

**Growth Factor Receptors and Growth Factors.** Overexpression of EGFR in breast cancer correlates with estrogen receptor–negative status and with p53 overexpression. Similarly, increased immunohistochemical membrane staining for the HER2 growth factor receptor in breast cancer is associated with mutated TP53, Ki67 overexpression, and estrogen receptor–negative status. HER2 is a member of the ErbB family of growth factor receptors in which ligand binding results in receptor homodimerization and tyrosine phosphorylation by tyrosine kinase domains within the receptor. Tyrosine phosphorylation is followed by signal transduction, which results in changes in cell behavior. An important property of this family of receptors is that ligand binding to one receptor type also may result in heterodimerization between two different receptor types that are coexpressed; this leads to transphosphorylation and transactivation of both receptors in the complex (transmodulation). In this context, the lack of a specific ligand for the HER2/neu receptor suggests that HER2/neu may function solely as a co-receptor, modulating signaling by other EGFR family members. HER2/neu is both an important prognostic factor and a predictive factor in breast cancer. When overexpressed in breast cancer, HER2/neu promotes enhanced growth and proliferation, and increases invasive and metastatic capabilities. Clinical studies have shown that patients with HER2/neu–overexpressing breast cancer have poorly differentiated tumors with high proliferation rates, positive lymph nodes, decreased hormone receptor expression, and an increased risk of recurrence and death due to breast cancer. Routine testing of the primary tumor specimen for HER2/neu expression should be performed on all invasive breast cancers. This can be done with immunohistochemical analysis to evaluate for overexpression of the cell-surface receptor at the protein level or by using fluorescence in situ hybridization to evaluate for gene amplification. While HER2/ERBB2 activation can also be assessed based on mRNA expression and reverse transcription polymerase chain reaction (RT-PCR) (OncoType Dx, Genomic Health), this approach is not recommended for clinical decision-making because of the high false negative rate. Patients whose tumors show HER2 amplification or HER2/neu protein overexpression are candidates for anti-HER2/neu therapy. Trastuzumab (Herceptin) is a recombinant humanized monoclonal antibody directed against HER2. Randomized clinical trials have demonstrated that single-agent trastuzumab therapy is well tolerated and active in the treatment of women with HER2/neu–overexpressing metastatic breast cancer. Subsequent adjuvant trials demonstrated that trastuzumab also was highly effective in the treatment of women with early-stage breast cancer when used in combination with chemotherapy. Patients who received trastuzumab in combination with chemotherapy had between a 40% and 50% reduction in the risk of breast cancer recurrence and approximately a one-third reduction in breast cancer mortality compared with those who received chemotherapy alone.

**Indices of Proliferation.** PCNA is a nuclear protein associated with a DNA polymerase whose expression increases in phase G1 of the cell cycle, reaches its maximum at the G1/S interface, and then decreases through G2. Immunohistochemical staining for PCNA outlines the proliferating compartments in...
breast tissue. Good correlation is noted between PCNA expression and (a) cell-cycle distributions seen on flow cytometry based on DNA content, and (b) uptake of bromodeoxyuridine and the proliferation-associated Ki67 antigen. Individual proliferation markers are associated with slightly different phases of the cell cycle and are not equivalent. PCNA and Ki67 expression are positively correlated with p53 overexpression, high S-phase fraction, aneuploidy, high mitotic index, and high histologic grade in human breast cancer specimens, and are negatively correlated with estrogen receptor content. Ki67 was included with three other widely measured breast cancer markers (ER, PR, and HER2) into a panel of four immunohistochemical makers (IHC4), which together provided similar prognostic information to that in the 21 Gene Recurrence Score (Oncotype DX, Genomic Health). While there has been significant interest in using Ki67 as a biomarker, and while the IHC4 panel would be much less expensive than the 21 Gene Recurrence Score, there remain issues regarding reproducibility across laboratories.

Indices of Angiogenesis. Angiogenesis is necessary for the growth and invasiveness of breast cancer and promotes cancer progression through several different mechanisms, including delivery of oxygen and nutrients and the secretion of growth-promoting cytokines by endothelial cells. VEGF induces its effect by binding to transmembrane tyrosine kinase receptors. Overexpression of VEGF in invasive breast cancer is correlated with increased microvessel density and recurrence in node-negative breast cancer. An angiogenesis index has been developed in which microvessel density (CD31 expression) is combined with expression of thrombospondin (a negative modulator of angiogenesis) and p53 expression. Both VEGF expression and the angiogenesis index may have prognostic and predictive significance in breast cancer. Bevacizumab (a monoclonal antibody to VEGF) was approved by the FDA for use in metastatic breast cancer in combination with paclitaxel chemotherapy. This approval was based on results from a phase 3 trial by the Eastern Cooperative Oncology Group. The group’s E2100 trial showed that when bevacizumab was added to paclitaxel chemotherapy, median progression-free survival increased to 11.3 months from the 5.8 months seen in patients who received paclitaxel alone. The results were not reproduced in other trials, and the indication for the drug was revoked by the FDA in 2011.

Indices of Apoptosis. Alterations in programmed cell death (apoptosis), which may be triggered by p53-dependent or p53-independent factors, may be important prognostic and predictive biomarkers in breast cancer. Bcl-2 family proteins appear to regulate a step in the evolutionarily conserved pathway for apoptosis, with some members functioning as inhibitors of apoptosis and others as promoters of apoptosis. Bcl-2 is the only oncogene that acts by inhibiting apoptosis rather than by directly increasing cellular proliferation. The death-signal protein bax is induced by genotoxic stress and growth factor deprivation in the presence of wild-type (normal) p53 and/or AP-1/fo. The bax to bcl-2 ratio and the resulting formation of either bax-bax homodimers, which stimulate apoptosis, or bax–bcl-2 heterodimers, which inhibit apoptosis, represent an intracellular regulatory mechanism with prognostic and predictive implications. In breast cancer, overexpression of bcl-2 and a decrease in the bax to bcl-2 ratio correlate with high histologic grade, the presence of axillary lymph node metastases, and reduced disease-free and overall survival rates. Similarly, decreased bax expression correlates with axillary lymph node metastases, a poor response to chemotherapy, and decreased overall survival.

The remaining biomarkers and biologic targets listed earlier are still in preclinical testing, and clinical trials are evaluating their importance in breast cancer for both prognostic and predictive purposes.

Coexpression of Biomarkers. Selection of optimal therapy for breast cancer requires both an accurate assessment of prognosis and an accurate prediction of response to therapy. The breast cancer markers that are most important in determining therapy are estrogen receptor, progesterone receptor, and HER2/neu. Clinicians evaluate clinical and pathologic staging and the expression of estrogen receptor, progesterone receptor, and HER2/neu in the primary tumor to assess prognosis and assign therapy. Adjuvant! Online (http://www.adjuvantonline.com) is one of a number of programs available to clinicians that incorporates clinical and pathologic factors for an individual patient and calculates risk of recurrence and death due to breast cancer and then provides an assessment of the reduction in risk of recurrence that would be expected with the use of combination chemotherapy, endocrine therapy, or both of these. Adjuvant! Online was developed using information from the SEER database, the EBCTCG overview analyses, and results from other individual published trials. The website is updated and modified as new information becomes available. Clinicopathologic factors are used to separate breast cancer patients into broad prognostic groups, and treatment decisions are made on this basis (Table 17-12). Other indices and programs that are validated and used include the Nottingham Prognostic Index, and PREDICT. When an approach, which combines prognostic factors is used, up to 70% of early breast cancer patients receive adjuvant chemotherapy that is either unnecessary or ineffective. As described earlier, a wide variety of biomarkers have been shown to individually predict prognosis and response to therapy, but they do not improve the accuracy of either the assessment of prognosis or the prediction of response to therapy.

As knowledge regarding cellular, biochemical, and molecular biomarkers for breast cancer have improved, prognostic indices have developed that combine the predictive power

### Table 17-12

<table>
<thead>
<tr>
<th>Traditional prognostic and predictive factors for invasive breast cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TUMOR FACTORS</strong></td>
</tr>
<tr>
<td>Nodal status</td>
</tr>
<tr>
<td>Tumor size</td>
</tr>
<tr>
<td>Histologic/nuclear grade</td>
</tr>
<tr>
<td>Lymphatic/vascular invasion</td>
</tr>
<tr>
<td>Pathologic stage</td>
</tr>
<tr>
<td>Hormone receptor status</td>
</tr>
<tr>
<td>DNA content (ploidy, S-phase fraction)</td>
</tr>
<tr>
<td>Extent of intraductal component HER2/neu expression</td>
</tr>
</tbody>
</table>

of several individual biomarkers with the relevant clinicopathologic factors.

Recent technological advances have enabled implementation of high-throughput gene expression assays in clinical practice. These assays enable detailed stratification of breast cancer patients for assessment of prognosis and for prediction of response to therapy. The Oncotype DX is a 21-gene RT-PCR–based assay that has been approved for use in newly diagnosed patients with node-negative, ER-positive breast cancer. A recurrence score is generated, and those patients with high recurrence scores are likely to benefit from chemotherapy, whereas those with low recurrence scores benefit most from endocrine therapy and may not require chemotherapy. Results from the Trial Assessing Individualized Options for Treatment for breast cancer (TAILORx), designed to prospectively validate the use of 21-gene expression assay, have shown that patients with low recurrence score (0 to 10) have a low rate of local-regional and distant recurrence (98.7%) and very good overall survival at 5 years (98%) with endocrine therapy alone without chemotherapy. This study has randomly assigned patients with an intermediate recurrence score (11 to 25) to endocrine therapy alone or to chemotherapy followed by endocrine therapy.

Additionally, retrospective analysis has shown that the 21-gene recurrence score can be used in postmenopausal patients with ER-positive tumors and 1 to 3 involved axillary lymph nodes to predict the benefit of chemotherapy in addition to endocrine therapy. Knowledge of the recurrence score has been shown to alter treatment recommendations by oncologists, and patients likewise change their decision to undergo treatment based on the risk of recurrence. The MammaPrint assay uses a 70-gene expression profile to assess the risk of distant metastasis. MammaPrint is FDA approved for use in stage-1 or stage-2, node negative, ER-positive or ER-negative breast cancers to identify patients with high or low risk of recurrence. Although fresh tissue was initially required to perform the assay, it has since been adapted for use in paraffin-embedded tissue samples. The prospective RASTER study reported that patients classified as low risk based on MammaPrint had a 97% distant recurrence-free interval at five years. Results of the prospective MINDACT (MicroarrayInNode negative and 1–3 positive lymph node Disease may Avoid ChemoTherapy) trial were recently reported. The study was designed to assess whether the 70-gene expression assay would help avoid chemotherapy in patients who are considered clinically high risk but categorized as low genomic risk based on the assay. A 5-year rate of distant metastasis-free survival of more than 92% was identified as the cutoff for the benefit of chemotherapy. At 5 years, the rate of survival without distant metastasis in patients with high clinical risk and low genomic risk was 94.7%, meeting the criteria for noninferiority. However, the rate of disease-free survival and overall survival was higher with chemotherapy in the intention to treat population.

**OVERVIEW OF BREAST CANCER THERAPY**

Before diagnostic biopsy, the surgeon must consider the possibility that a suspicious mass or mammographic finding may be a breast cancer. Once a diagnosis of breast cancer is made, the type of therapy offered to a breast cancer patient is determined by the stage of the disease, the biologic subtype, and the general health status of the individual. Laboratory tests and imaging studies are performed based on the initial stage as presented in Table 17-13. Before therapy is initiated, the patient and the surgeon must share a clear perspective on the planned course of treatment. Before initiating local therapy, the surgeon should determine the clinical stage, histologic characteristics, and appropriate biomarker levels.

### In Situ Breast Cancer (Stage 0)

Both LCIS and DCIS may be difficult to distinguish from atypical hyperplasia or from cancers with early invasion. Expert pathologic review is required in all cases. Bilateral mammography is performed to determine the extent of the in situ cancer and to exclude a second cancer. Because LCIS is considered a marker for increased risk rather than an inevitable precursor of invasive disease, the current treatment options for LCIS include observation, chemoprevention, and bilateral total mastectomy. The goal of treatment is to prevent or detect at an early stage the invasive cancer that subsequently develops in 25% to 35% of these women. There is no benefit to excising LCIS because the disease diffusely involves both breasts in many cases and the risk of developing invasive cancer is equal for both breasts. The use of tamoxifen as a risk-reduction strategy should be considered in women with a diagnosis of LCIS.

Women with DCIS and evidence of extensive disease (>4 cm of disease or disease in more than one quadrant) usually require mastectomy (Fig. 17-29). For women with limited disease, lumpectomy and radiation therapy are generally recommended. For nonpalpable DCIS, needle localization or other image-guided techniques are used to guide the surgical resection. Specimen mammography is performed to ensure that all visible evidence of cancer is excised. Adjuvant tamoxifen therapy is considered for DCIS patients with ER-positive
disease. The gold standard against which breast conservation therapy for DCIS is evaluated is mastectomy. Women treated with mastectomy have local recurrence and mortality rates of <2%. There is no randomized trial comparing mastectomy vs. breast conserving surgery, and none of the randomized trials of breast-conserving surgery with or without radiotherapy for DCIS were powered to show a difference in mortality. Women treated with lumpectomy and adjuvant radiation therapy in the initial clinical trials were noted to have a local recurrence rate that is increased compared to mastectomy. About 45% of these recurrences will be invasive cancer when radiation therapy is not used. The B-17 trial was conducted by the NSABP to assess the need for radiation in patients treated with breast conserving surgery for DCIS. Patients were randomly assigned to lumpectomy with radiation or lumpectomy alone, and after a mean follow-up time of 90 months rates of both ipsilateral noninvasive and invasive recurrences were significantly lower in patients who received radiation. However, in the B-17 trial the margins were not prospectively assessed, and it is estimated that up to half of the patients may have had tumor at the margin of resection. The benefit of the addition of radiation over breast-conserving surgery alone for DCIS has also been demonstrated in several other randomized trials where margins were prospectively assessed including the European Organization for Research and Treatment of Cancer (EORTC) protocol 10853; the United Kingdom, Australia, New Zealand DCIS Trial; and the Swedish Trial.

In 2016, the Society of Surgical Oncology (SSO), American Society for Radiation Oncology (ASTRO), and the American Society of Clinical Oncology (ASCO) established consensus guidelines on margins for patients with DCIS undergoing breast-conserving surgery. Based on a multidisciplinary consensus panel using a meta-analysis of margin width and ipsilateral breast tumor recurrence, a 2-mm margin was determined as adequate width for DCIS for patients undergoing breast-conserving surgery with whole-breast radiation therapy.

Despite the data from randomized trials showing a benefit in all patient subgroups with the addition of radiation in DCIS, there has been an interest in trying to define a subset where radiation could be avoided in order to minimize the cost and inconvenience associated with radiation. In addition, there have been several studies published where patients were treated with excision alone and never developed invasive breast cancer even at 25 years of follow-up. Silverstein and colleagues were proponents of avoiding radiation therapy in selected patients with DCIS who have widely negative margins after surgery. They reported that when greater than 10-mm margins were achieved, there was no additional benefit from radiation therapy. When margins were between 1 and 10 mm, there was a relative risk of local recurrence of 1.49, compared to 2.54 for those with margins less than 1 mm. These data suggested that appropriately selected patients with DCIS might not require postoperative radiation therapy.

The Eastern Cooperative Oncology Group (ECOG) initiated a prospective registry trial (ECOG 5194) to identify those patients who could safely undergo breast-conserving surgery without radiation. Eligible patients were those with low or intermediate grade DCIS measuring 2.5 cm or less who had negative margins of at least 3 mm and those with high-grade DCIS who had tumors measuring 1 cm or less with a negative margin of at least 3 mm. At a median follow-up of 6.2 years, patients with low or intermediate grade DCIS had an in-breast
recurrence rate of 6.1% while those with high-grade DCIS had a recurrence rate of 15.3%. Approximately 4% of patients developed a contralateral breast cancer during follow-up in both the low/intermediate and high-grade groups. This study identified an acceptable recurrence rate for those patients with low or intermediate grade DCIS treated with excision alone with a margin of at least 3 mm. In contrast, patients with high-grade DCIS had an unacceptably high local recurrence rate.

The Radiation Therapy Oncology Group (RTOG) initiated the 9804 trial for patients with “good risk” DCIS and randomized them to lumpectomy vs. lumpectomy with whole breast irradiation. Eligible patients were those with unicentric, low or intermediate grade DCIS measuring 2.5 cm or less with a margin of 3 mm or greater. The trial was closed early due to slow accrual; however, the results for 585 patients were recently reported with a median follow-up of 6.46 years. The local recurrence rate at 5 years was 0.4% for patients randomized to receive radiation and 3.2% for those who did not receive radiation.

Solin et al utilized samples from the ECOG 5194 trial to develop a quantitative multigene RT-PCR assay for predicting recurrence risk in patients with DCIS treated with surgery alone. They were able to define low, intermediate, and high risk groups using a DCIS Score. The DCIS Score was able to quantify the risk of recurrence in the breast for both DCIS and invasive events. This tool has recently been evaluated in another dataset and appears to be a promising tool for clinical use.

When selecting therapy for patients with DCIS, one must consider clinical and pathologic factors, including tumor size, grade, mammographic appearance, and patient preference. There is no single correct surgical treatment, and many patients will require extensive counseling in order to make a decision regarding surgical therapy. The role of axillary staging in patients with DCIS is limited. One consideration is for patients undergoing mastectomy. Since most lesions are currently diagnosed with needle core biopsy, there is about a 20% incidence of invasive breast cancer on final pathologic assessment of the primary tumor. Since it is not feasible to perform sentinel node dissection after mastectomy, most surgeons will recommend the use of sentinel node dissection at the time of mastectomy for DCIS.

Results from the NSABP B-24 trial reported a significant reduction in local recurrence after 5 years of tamoxifen in women with ER-positive DCIS. Based on this finding, some guidelines have advocated that all patients (women with ER-positive DCIS without contraindications to tamoxifen therapy) should be offered tamoxifen following surgery and radiation therapy for a duration of 5 years. The B-24 trial revealed a significant reduction in recurrence with adjuvant tamoxifen therapy for patients with DCIS; however, the results were not initially assessed based on ER status. There were 1804 women with DCIS randomized to lumpectomy and radiation with or without tamoxifen. The rate of breast cancer events was significantly lower in those who received tamoxifen at a median follow-up of 74 months (8.2% vs. 13.4%, P = 0.0009). Subsequently, Alred and colleagues evaluated 41% of patients with DCIS in the NSABP B-24 trial to determine the effect of tamoxifen based on ER status measured in the primary tumor. They found that 76% of women had DCIS that was ER-positive and these women had a greater reduction in ipsilateral breast tumor recurrence with tamoxifen than did patients with ER-negative DCIS (11% vs. 5.2%, P < 0.001). However, it should be noted that 15% of patients in B-24 had tumor at the resection margins. For these patients, tamoxifen could be viewed as treating what, by the current standard, would be viewed as inadequate local excision of the primary tumor.

Early Invasive Breast Cancer (Stage I, IIIA, or IIB)

There have been six prospective randomized trials comparing breast-conserving surgery to mastectomy in early stage breast cancer, and all have shown equivalent survival rates regardless of the surgical treatment type. One caveat, however, is that the majority of studies had a restriction of tumor size; most were either 2 cm or 2.5 cm, while the NSABP B-06 trial was 4 cm, and the NCI trial was up to 5 cm. NSABP B-06, which is the largest of all the breast conservation trials, compared total mastectomy to lumpectomy with or without radiation therapy in the treatment of women with stages I and II breast cancer. After 5- and 8-year follow-up periods, the disease-free (DFS), distant disease-free, and overall survival (OS) rates for lumpectomy with or without radiation therapy were similar to those observed after total mastectomy. However, the incidence of ipsilateral breast cancer recurrence was higher in the group not receiving radiation therapy. These findings supported the use of lumpectomy and radiation therapy in the treatment of stages I and II breast cancer and has since become the preferred method of treatment for women with early stage breast cancer who have unicentric disease and who are not known BRCA mutation carriers. Reanalysis of the B-06 study results was undertaken after 20 years of follow-up and confirmed that there was no difference in disease-free survival rates after total mastectomy or after lumpectomy with or without adjuvant radiation therapy. The in-breast recurrence rate was substantially higher in the lumpectomy alone group (39.2%) compared with the lumpectomy plus adjuvant radiation therapy group (14.3%), confirming the importance of radiation therapy in the management of patients with invasive disease. However, it should be noted that there were several criteria in the B-06 study. There was a specific lymphadenopathy exclusion criteria. Secondly, all patients randomized to breast-conserving surgery had a frozen section, and if the margins were involved, they were converted to mastectomy but were included in the analysis as having had a breast-conserving operation (on the basis of intention to treat). Finally, in the breast-conserving group recurrences in the treated breast were considered as a “nonevent.”

Data from all of the randomized trials where breast conservation was performed with or without radiation therapy have been examined by the EBCTCG. At 15 years of follow-up, the absolute reduction in mortality with the use of radiation therapy after lumpectomy was 5.1% in node-negative patients and 7.1% in node-positive patients. These data support the concept that the addition of radiation not only improves local control but also has an impact on survival. Similar to DCIS, clinicians have sought to identify subgroups of patients who may not benefit from the addition of radiation therapy, particularly older patients who may have a shorter life expectancy due to medical comorbidities. Randomized trials have shown that in selected patients with small, ER-positive, low-grade tumors, lumpectomy alone without radiation therapy may be appropriate. The Cancer and Leukemia Group B (CALGB) C9343 trial enrolled women over the age of 70 with T1N0 breast cancer and randomized them to lumpectomy with or without radiation therapy. All patients received adjuvant tamoxifen.
there were fewer local recurrences with radiation (1% vs. 4%, \(P < 0.001\)), there were no differences in DFS and OS. While long-term follow-up at 10 years showed fewer local recurrences with radiation (2% vs. 10%), there were no significant differences in time to distant metastasis, breast cancer–specific survival, or OS between the two groups. A trial similar to CALGB C9343 was conducted in Canada where they enrolled women age 50 years and older and randomized them to lumpectomy with or without radiation. Mean age was 68 years, and 80% of women had ER-positive tumors. Again, local recurrence rates were lower in women who received radiation (0.6% vs. 7.7%, \(P < 0.001\)); however, at a median follow-up of 5.6 years, there were no differences in DFS or OS. The PRIME-2 study enrolled women age 65 years or older with ER-positive, node-negative, up to 3 cm breast cancers, who had undergone breast-conserving surgery and were candidates for adjuvant endocrine treatment. They were assigned to receive whole-breast irradiation or no treatment. After a median follow-up of 5 years, ipsilateral breast tumor recurrence was 1.3% with radiation vs. 41% in those assigned to no radiotherapy. However, no differences in distant metastases, contralateral breast cancers, or overall survival were noted between the groups.\(^{234}\) These studies suggest that radiation can be avoided in select older patients with ER-positive, early-stage breast cancer.

Accelerated partial breast irradiation (APBI) is also an option for carefully selected patients with DCIS and early-stage breast cancer. Since the majority of recurrences after breast conservation occur in or adjacent to the tumor bed, there has been interest in limiting the radiation to the area of the primary tumor bed with a margin of normal tissue. APBI is delivered in an abbreviated fashion (twice daily for 5 days) and at a lower total dose compared with the standard course of 5 to 6 weeks of radiation (50 Gy with or without a boost) in the case of whole breast irradiation. Proponents have suggested that this shortened course of treatment may increase the feasibility of breast conservation for some women and may improve radiation therapy compliance. The RTOG 04-13/NSABP B-39 trial is a randomized comparison of whole breast irradiation to APBI in women with early stage breast cancer. The trial has completed accrual, and it will likely be several years before data are mature to report outcomes between the two radiation treatment strategies. TARGIT is another study that randomized 3451 patients in 33 centers in over 10 countries to intraoperative breast irradiation (IORT) or external beam radiotherapy (EBRT). The preliminary results were reported in 2012: with a median follow-up of 2.4 years, use of IORT had a recurrence rate of 3.3% vs. 1.3% with EBRT, a 2% increased recurrence risk.\(^{235,236}\) ASTRO developed guidelines for the use of APBI outside of clinical trials based on data reported from published studies.\(^{237,238}\) The ASTRO guidelines describe patients “suitable” for APBI to include women age 60 years or older with a unifocal, T1, ER-positive tumor with no lymphovascular invasion and margins of at least 2 mm. They describe a group where there is uncertainty about the appropriateness of APBI (“cautionary” group) to include patients with invasive lobular histology, a tumor size of 2.1 cm to 3 cm, ER-negative disease, focal lymphovascular invasion, or margins less than 2 mm. Finally, a group felt to be “unsuitable” for APBI includes those with T3 or T4 disease, ER-negative disease, multifocality, multicentricity, extensive LVI, or positive margins.

Currently, mastectomy with axillary staging and breast conserving surgery with axillary staging and radiation therapy are considered equivalent treatments for patients with stages I and II breast cancer. Breast conservation is considered for all patients because of the important cosmetic advantages and equivalent survival outcomes; however, this approach is not advised in women who are known BRCA mutation carriers due to the high lifetime risk for development of additional breast cancers. Relative contraindications to breast conservation therapy include (a) prior radiation therapy to the breast or chest wall, (b) persistently positive surgical margins after reexcision, (c) multicentric disease, and (d) scleroderma or lupus erythematosus.

For most patients with early-stage disease, reconstruction can be performed immediately at the time of mastectomy. Immediate reconstruction allows for skin-sparing, thus optimizing cosmetic outcomes. Skin-sparing mastectomy with immediate reconstruction has been popularized over the past decade as reports of low local-regional failure rates have been reported and reconstructive techniques have advanced. There is a growing interest in the use of nipple-areolar sparing mastectomy with reports suggesting the oncologic safety of this approach in early stage breast cancer. Patients who are planned for postmastectomy radiation therapy may not be ideal candidates for nipple-sparing mastectomy because of the effects of radiation on the preserved nipple. In addition to providing optimal cosmesis from preservation of the skin and/or the nipple-areolar complex, immediate reconstruction allows patients to wake up with a breast mound, which provides some psychological benefit for the patient. Immediate reconstruction is also more economical as both the extirpative and reconstructive surgery are combined in one operation.

Immediate reconstruction can be performed using implants or autologous tissue; tissue flaps commonly used include the transverse rectus abdominis myocutaneous flap, deep inferior epigastric perforator flap, and latissimus dorsi flap (with or without an implant). If postmastectomy radiation therapy is needed, a tissue expander can be placed at the time of mastectomy to save the shape of the breast and reduce the amount of skin replacement needed at the time of definitive reconstruction. The expander can be deflated at the initiation of radiation therapy to allow for irradiation of the chest wall and regional nodal basins. Removal of the tissue expander and definitive reconstruction, usually with autologous tissue, can proceed 6 months to 1 year after completion of radiation therapy.

Axillary lymph node status has traditionally been an important determinant in staging and prognosis for women with early stage breast cancer. Historically, axillary lymph node dissection (ALND) was utilized for axillary staging and regional control by removing involved lymph nodes. Randomized trials evaluating immediate ALND over ALND performed in a delayed fashion once clinically palpable axillary disease became evident have not shown any detriment in survival.\(^{239}\) With increased mammographic screening and detection of smaller, node-negative breast cancers, it became clear that routine use of ALND for axillary staging was not necessary in up to 75% percent of women with operable breast cancer presenting with a negative axilla at the time of screening. Lymphatic mapping and sentinel lymph node (SLN) dissection were initially developed for assessment of patients with clinically node-negative melanoma. Given the changing landscape of newly diagnosed breast cancer patients with a clinically node-negative axilla, surgeons quickly began to explore the utility of SLN dissection as a replacement for ALND in axillary staging.
In the early 1990s, David Krag at the University of Vermont began performing SLN dissection with injection of a radioisotope in the primary tumor site and localizing the SLN node with a handheld gamma probe. He was able to identify a SLN in 18 of 22 patients examined, and the SLN was positive in all 7 patients with positive lymph nodes. Giuliano and colleagues initiated a pilot study in 1991 to examine the use of SLN dissection using blue dye in patients with clinically negative nodes. They reported successful identification of a SLN in 114 (65.5%) of 174 patients, and in 109 (95.6%), the SLN accurately predicted the status of the axillary nodes. These studies along with initial work by Doug Reintgen and Charles Cox at the Moffitt Cancer Center and Umberto Veronesi and his colleagues at the European Institute of Oncology in Milan led the way toward validation of the technique in large single institution and multicenter studies.

Following validation of the technique of SLN dissection for staging of the axilla by multiple centers, randomized trials were initiated in order to determine if SLN dissection could replace ALND in the contemporary management of breast cancer patients. The ALMANAC trial randomized 1031 patients with primary operable breast cancer to SLN dissection vs. standard axillary surgery. The incidence of lymphedema and sensory loss for the SLN group was significantly lower than with the standard axillary treatment. At 12 months, drain usage, length of hospital stay, and time to resumption of normal day-to-day activities after surgery were also statistically significantly lower in the SLN group. The NSABP B-32 trial compared clinically node-negative patients undergoing SLN dissection followed by ALND with patients undergoing SLN dissection with ALND only if a SLN was positive for metastatic disease. A total of 5611 patients were randomized with a SLN identification rate of 97% and a false-negative rate of 9.7%. A total of 26% of these clinically node-negative patients had a positive SLN. Over 60% of patients with positive SLNs had no additional positive lymph nodes within the ALND specimen. The B-32 trial and other randomized trials demonstrated no difference in DFS, OS, and local-regional recurrence rates between patients with negative SLNs who had SLN dissection alone compared with those who underwent ALND. Most important, patients who had SLN dissection alone were found to have decreased morbidity (arm swelling and range of motion) and improved quality of life vs. patients who underwent ALND.

The American College of Surgeons Oncology Group (ACOSOG) initiated the Z0010 and Z0011 trials in order to evaluate the incidence and prognostic significance of occult metastases identified in the bone marrow and SLNs (Z0010) of early-stage clinically node-negative patients and to evaluate the utility of ALND in patients with clinical T1-2, N0 breast cancer with 1 or 2 positive SLNs for patients treated with breast-conserving surgery and whole breast irradiation (WBI) (Z0011). The Z0010 study enrolled 5539 patients with clinical T1-2 breast cancer planned for breast conserving surgery and WBI. Of these patients, 24% proved to have positive SLNs based on standard pathologic assessment, and of the negative SLNs subjected to immunohistochemical staining for cytokeratin, 10.5% proved to have occult metastasis. Of the patients who had bone marrow aspiration, 3.0% had immunohistochemically detected tumor cells in the bone marrow. Although the presence of disease in the bone marrow identified a population at high risk for recurrence, neither immunohistochemical detection of disease in the SLNs or the bone marrow was statistically significant on multivariable analysis with clinicopathologic and treatment factors included. The investigators concluded that routine use of immunohistochemistry to detect occult disease in SLNs is not warranted.

The Z0011 trial was a companion study to Z0010 and was designed to study the role of completion ALND on survival in women with positive SLNs. Patients were not eligible if they received neoadjuvant chemotherapy or neoadjuvant hormonal therapy or if their treatment plan included mastectomy, lumpectomy without radiation, or lumpectomy with APBI. WBI was to be administered using standard tangential fields without specific treatment of the axilla or additional fields targeting other nodal basins. Patients with 1 or 2 positive SLNs were randomized to completion ALND or no further surgery. Adjuvant systemic therapy recommendations were left to the treating clinicians. After median follow-up of 6.3 years, there was no difference between patients randomized to ALND and those randomized to no further surgery (SLN only) in terms of OS (91.9% and 92.5%, respectively; \( P = 0.25 \)) or DFS (82.2% and 83.8%, respectively; \( P = 0.14 \)). The low local regional failure rates and similar survival outcomes were recently reported with 10-year follow-up.

The morbidity of SLN dissection alone vs. SLN dissection with completion ALND has been reported by the ACOSOG investigators. Immediate effects of SLN dissection in the Z0010 trial included wound infection in 1%, axillary seroma in 7.1%, and axillary hematoma in 1.4%. At 6 months following surgery, axillary paresthesias were noted in 8.6% of patients, decreased range of motion in the upper extremity was reported in 3.8%, and 6.9% of patients had a change in the arm circumference of \( >2 \) cm on the ipsilateral side, which was reported as lymphedema. Younger patients were more likely to report paresthesias, whereas increasing age and body mass index were more predictive of lymphedema. When adverse surgical effects were examined in the Z0011 trial, patients undergoing SLN dissection with ALND had more wound infections, seromas, and paresthesias than those women undergoing SLN dissection alone. Lymphedema at 1 year after surgery was reported by 13% in the SLN plus ALND group but only 2% in the SLN dissection alone group. Arm circumference measurements were greater at 1 year in patients undergoing SLN dissection plus ALND, but the difference between study groups was not statistically significant. This supports the results published from the ALMANAC trial.

Prior to the publication of ACOSOG Z0011, completion ALND was standard of care for patients with positive SLNs. Since the reporting of ACOSOG Z0011, the National Comprehensive Cancer Network (NCCN) guidelines now state that there was no OS difference for patients with 1 or 2 positive SLNs treated with breast-conserving surgery who underwent completion ALND vs. those who had no further axillary surgery. In addition, the American Society of Breast Surgeons issued a consensus statement supporting omission of ALND for patients who meet Z0011 criteria. The results of ACOSOG Z0011 have revolutionized management of the axilla and changed practice such that selected patients with axillary metastasis can now avoid ALND if they have clinical and pathologic features similar to those patients enrolled on Z0011. However, there have been some concerns raised about the Z0011 study that include the fact that the study only recruited about half of
the intended patients and that there was no standardization of whether or not patients received irradiation to the low axilla when the radiation oncologist irradiated the breast. These issues have thus far limited the uptake of the results of Z0011 by some centers.

The International Breast Cancer Study Group (IBCSG) 23-01 trial was similar in design to Z0011 but enrolled only patients with micrometastases in the SLNs. Patients with SLN micrometastases were randomized to ALND vs. no further surgery. Unlike Z0011, the 23-01 trial did not exclude patients treated with mastectomy. Approximately 9% of patients randomized to each study arm underwent mastectomy. The investigators published the primary and secondary endpoints of the trial showing no differences in OS or local-regional recurrence between the study arms. However, as with the Z0011 trial, some concerns have been raised regarding the 23-01 study. For example, in the statistics on the primary endpoint, local recurrence included contralateral breast cancer and other tumor types as events. No hypothesis was presented as to why the difference in axillary surgery should impact on either of these events. Including these events therefore reduced the power of the study to show a statistical difference between treatment arms. There is also concern that the study appears underpowered to show a meaningful difference in overall survival.

Most pathology laboratories perform a more detailed analysis of the SLN than is routinely done for axillary nodes recovered from a levels I and II dissection. This can include examining thin sections of the node with step sectioning at multiple levels through the paraffin blocks or performing immunohistochemical staining of the SLN for cytokeratin or a combination of these techniques. The results of ACOSOG Z0010 and NSABP B-32 showed no clinically meaningful difference in survival based on detection of occult metastases in the SLNs using immunohistochemical staining and do not support the routine use in SLN processing. The type of intraoperative assessment of SLNs also varies for different clinicians and pathology laboratories. Some centers prefer to use touch preparation cytologic analysis of the SLNs, whereas others use frozen-section analysis, and the sensitivity and specificity of these assays vary considerably. The GeneSearch Breast Lymph Node Assay is a real-time reverse-transcriptase polymerase chain reaction assay that detects breast tumor cell metastasis in lymph nodes through the identification of the gene expression markers mammaglobin and cytokeratin 19. These markers are present in higher levels in breast tissue and not in nodal tissue (cell type-specific messenger RNA). The GeneSearch breast lymph node assay generates expression data for genes of interest, which are then evaluated against predetermined criteria to provide a qualitative (positive/negative) result. The assay is designed to detect foci that correspond to metastases that are seen with examination by standard hematoxylin and eosin staining and measure >0.2 mm. The GeneSearch assay results have been compared with permanent-section histologic analysis and frozen-section analysis of sentinel nodes in a prospective trial, and the assay was approved by the FDA for the intraoperative assessment of sentinel nodes. When a positive node is identified intraoperatively by touch preparation, frozen-section analysis, or GeneSearch assay, the surgeon can proceed with immediate ALND. With the findings of ACOSOG Z0011 that there is not a survival benefit to the use of ALND in selected patients, many surgeons have abandoned the intraoperative evaluation of SLNs. There are a number of nomograms and predictive models designed to determine which patients with a positive SLN are at risk for harboring additional positive non-SLNs in the axilla. These tools can be helpful in determining the likelihood of additional disease in the axilla and may be used clinically to counsel patients.

In patients who present with axillary lymphadenopathy that is confirmed to be metastatic disease on FNA or core biopsy, SLN dissection is not necessary, and patients can proceed directly to ALND or be considered for preoperative systemic therapy (see “Neoadjuvant [Preoperative] Chemotherapy” under “Nonsurgical Breast Cancer Therapies”). Initially there was controversy about the suitability of SLN dissection in women with larger primary tumors (T3) and those treated with neoadjuvant chemotherapy. The American Society of Clinical Oncology has included SLN dissection is its guidelines as appropriate for axillary staging in these patients. If an SLN cannot be identified, then ALND is generally performed for appropriate staging. However, this is not universally accepted, and there are as yet no randomized studies that have assessed how a patient with a locally advanced cancer at presentation should be treated if SLN dissection reveals no metastases or micrometastases after neoadjuvant therapy.

The ASCO guidelines suggest that adjuvant chemotherapy should be considered for patients with positive lymph nodes, ER-negative disease, HER2-positive disease, Adjuvant! Online mortality greater than 10%, grade 3 node-negative tumors >5 mm, triple-negative tumors, lymphovascular invasion, or estimated distant relapse risk of greater than 15% at 10 years based on the 21 gene recurrence score assay. Adjuvant endocrine therapy is considered for women with hormone receptor-positive cancers, and an aromatase inhibitor is recommended if the patient is postmenopausal. HER2/neu status is determined for all patients with newly diagnosed invasive breast cancer and when positive, should be used to guide systemic therapy recommendations. The FDA approved trastuzumab in November 2006 for use as part of a treatment regimen containing doxorubicin, cyclophosphamide, and paclitaxel for treatment of HER2/neu-positive, node-positive breast cancer. Subsequently, the BCIRG 006 study reported that giving trastuzumab concurrently with docetaxel and carboplatin appeared as effective as giving trastuzumab following an anthracycline containing regimen. In addition to trastuzumab, pertuzumab has also recently been FDA approved for adjuvant use in patients with HER2 amplified breast cancers with high risk of recurrence.

Advanced Local-Regional Breast Cancer (Stage IIIA or IIIB)

Women with stage IIIA and IIIB breast cancer have advanced local-regional breast cancer but have no clinically detected distant metastases (Fig. 17-30). In an effort to provide optimal local-regional disease-free survival as well as distant disease-free survival for these women, surgery is integrated with radiation therapy and chemotherapy (Fig. 17-31). However, it should be noted that these patients have an increased risk of distant metastasis that is often highlighted by radiological evidence when staging PET or CT and bone scans are performed. Thus, the paradigm for small screen detected cancers where cure can be expected in >90% of patients, often by local treatment alone, is not appropriate for patients with locally advanced disease.

Preoperative (also known as neoadjuvant) chemotherapy should be considered in the initial management of patients with
locally advanced stage III breast cancer, especially those with estrogen receptor negative tumors. Chemotherapy is used to maximize distant disease-free survival, whereas radiation therapy is used to maximize local-regional control and disease-free survival.

In selected patients with stage IIIA cancer, preoperative chemotherapy can reduce the size of the primary cancer and permit breast-conserving surgery. Investigators from the MD Anderson Cancer Center reported that low local-regional failure rates could be achieved in selected patients with stage III disease treated with preoperative chemotherapy followed by breast-conserving surgery and radiation. The 5-year actuarial ipsilateral breast tumor recurrence-free survival rates in this study were 95%. They noted that the ipsilateral breast tumor recurrence rates increased when patients had clinical N2 or N3 disease, >2 cm of residual disease in the breast at surgery, a pattern of multifocal residual disease in the breast at surgery, and lymphovascular space invasion in the primary tumor. This study demonstrated that breast-conserving surgery can be used for appropriately selected patients with locally advanced breast cancer who achieve a good response with preoperative chemotherapy. However, the Oxford overview of all randomized studies of neoadjuvant therapy (vs. adjuvant therapy) reported a hazard ratio of 1.5 (i.e., 50% increase) in local recurrence rates.

Figure 17-30. Locally advanced breast cancer. A. Mammography of the right breast reveals a large tumor with enlarged axillary lymph nodes. B. Imaging of the left breast is normal. (Used with permission from Dr. Anne Turnbull, Consultant Radiologist/Director of Breast Screening, Royal Derby Hospital, Derby, UK.)

Figure 17-31. Treatment pathways for stage IIIA and stage IIIB breast cancer.
A meta-analysis reported a hazard ratio of 1.3. These studies included some patients treated with radiation therapy alone without resection of the primary tumor bed, which results in higher local failure rates. These findings are important in view of the previous findings that the avoidance of recurrence in a conserved breast avoids about one breast cancer death over the next 15 years for every four such recurrences avoided. The German Breast Cancer Group recently reported their local recurrence rate in 5535 patients in seven studies. With a median of 46 months (range 1–127) follow-up the local recurrence rates ranged from 7.6% to 19.5% for T1-T4 tumors and from 6.4% to 17.9% for N0-N3 tumors treated with neoadjuvant therapy. For patients with stage IIIA disease who experience minimal response to chemotherapy and for patients with stage IIIB breast cancer, preoperative chemoradiation can decrease the local-regional cancer burden enough to permit subsequent modified radical mastectomy to establish local-regional control. In both stages II A and IIIB disease, surgery is followed by adjuvant radiation therapy. However there is a small percentage of age of patients who experience progression of disease during neoadjuvant therapy, and therefore the surgeon should review patients with the oncologist at regular points during the neoadjuvant regimen.

For selected clinically indolent, ER-positive, locally advanced tumors, primary endocrine therapy may be considered, especially if the patient has other comorbid conditions. A series of 195 patients with ER-positive, locally advanced breast cancer treated by endocrine therapy—median age 69 years, median tumor size 6 cm, median follow-up 61 months—reported a 5-year overall survival of 76%, a breast cancer–specific survival of 86%, and a metastasis-free survival of 77%. The median time to an alternative treatment was 48 months. Given that this was a 20-year series, the number of such patients is small but should be considered when the clinician is discussing treatment options. Results from the ACOSOG Z1031 trial suggest that neoadjuvant endocrine therapy is a good option for tumor downstaging in patients with strongly ER-positive tumors. The preoperative endocrine prognostic index (PEPI score) can be calculated based on pathologic findings from surgery following neoadjuvant endocrine therapy. This can help guide decision-making regarding the need for systemic chemotherapy in this patient population.

### Internal Mammary Lymph Nodes

Metastatic disease to internal mammary lymph nodes may be occult, may be evident on chest radiograph or CT scan, or may present as a painless parasternal mass with or without skin involvement. There is no consensus regarding the need for internal mammary lymph node radiation therapy in women who are at increased risk for occult involvement (cancers involving the medial aspect of the breast, axillary lymph node involvement) but who show no signs of internal mammary lymph node involvement. Systemic chemotherapy and radiation therapy are indicated in the treatment of grossly involved internal mammary lymph nodes.

### Distant Metastases (Stage IV)

Treatment for stage IV breast cancer is not curative but may prolong survival and enhance a woman’s quality of life. Endocrine therapies that are associated with minimal toxicity are preferred to cytotoxic chemotherapy in ER-positive disease. Appropriate candidates for initial endocrine therapy include women with hormone receptor-positive cancers who do not have immediately life threatening disease (or “visceral crisis”). This includes not only women with bone or soft tissue metastases but also women with limited visceral metastases. Symptoms per se (e.g., breathlessness) are not in themselves an indication for chemotherapy. For example, breathlessness due to a pleural effusion can be treated with percutaneous drainage, and if the breathlessness is relieved, the patient should be commenced on endocrine therapy; if the breathlessness is due to lymphangitic spread, then chemotherapy would be the treatment of choice. The same approach should be taken to other symptoms such as pain. Systemic chemotherapy is indicated for women with hormone receptor-negative cancers, “visceral crisis,” and hormone-refractory metastases. Women with stage IV breast cancer may develop anatomically localized problems that will benefit from individualized surgical or radiation treatment, such as brain metastases, pleural effusion, pericardial effusion, biliary obstruction, ureteral obstruction, impending or existing pathologic fracture of a long bone, spinal cord compression, and painful bone or soft tissue metastases. Bisphosphonates or anti-RANKL (receptor activator of nuclear factor kappa-B ligand) agent, denosumab, which may be given in addition to chemotherapy or endocrine therapy, should be considered in women with bone metastases. Whether to perform surgical resection of the local-regional disease in women with stage IV breast cancer has been debated after several reports have suggested that women who undergo resection of the primary tumor have improved survival over those who do not. Khan and associates used the National Cancer Data Base to identify patterns of treatment in women with metastatic breast cancer and found that those who had surgical resection with negative margins had a better prognosis than those women who did not have surgical therapy. Gnerlich et al reported similar findings using the SEER database, and there have been several reports subsequent to this study from single institutions that have confirmed these findings. Some have suggested that the finding of improved survival is due to selection bias and that local therapy should be reserved for palliation of symptoms. A randomized trial through ECOG (E2108) was designed to address this question. The surgical management of patients with stage IV disease should be addressed by obtaining multidisciplinary input and by considering the treatment goals of each individual patient and the patient’s treating physicians.

### Local-Regional Recurrence

Women with local-regional recurrence of breast cancer may be separated into two groups: those who have had mastectomy and those who have had lumpectomy. Women treated previously with mastectomy undergo surgical resection of the local-regional recurrence and appropriate reconstruction. Chemotherapy and antiestrogen therapy are considered, and adjuvant radiation therapy is given if the chest wall has not previously received radiation therapy or if the radiation oncologist feels that given the time from previous treatment there is scope for further radiation therapy, particularly if this is palliative. Women treated previously with a breast-conservation procedure undergo a mastectomy and appropriate reconstruction. Chemotherapy and antiestrogen therapy are considered depending of the hormone receptor status and HER2 status of the tumor.

### Breast Cancer Prognosis

Survival rates for women diagnosed with breast cancer in the United States can be obtained from the SEER Program of the
Surgical Techniques in Breast Cancer Therapy

Excisional Biopsy With Needle Localization

Excisional biopsy implies complete removal of a breast lesion with a margin of normal-appearing breast tissue. In the past, surgeons would obtain prior consent from the patient, allowing mastectomy if the initial biopsy results confirmed cancer. Today it is important to consider the options for local therapy (lumpectomy vs. mastectomy with or without reconstruction) and the need for nodal assessment with SLN dissection. Needle-core biopsy is the preferred diagnostic method, and excisional biopsy should be reserved for those cases in which the needle biopsy results are discordant with the imaging findings or clinical examination (Fig. 17-32). In general, circumareolar incisions can be used to access lesions that are subareolar or within a short distance of the nipple-areolar complex. Elsewhere in the breast, incisions can be placed along the lines of tension in the skin that are generally concentric with the nipple-areola complex. In the lower half of the breast, the use of radial incisions typically provides the best outcome. When the tumor is quite distant from the central breast, the biopsy incision can be excised separately from the primary mastectomy incision, should a mastectomy be required. Radial incisions in the upper half of the breast are not recommended because of possible scar contracture resulting in displacement of the ipsilateral nipple-areola complex. Similarly, curvilinear incisions in the lower half of the breast may displace the nipple-areolar complex downward.

After excision of a suspicious breast lesion, the specimen should be x-rayed to confirm that the lesion has been excised with appropriate margins. The biopsy tissue specimen is oriented for the pathologist using sutures, clips, or dyes. Additional margins (superior, inferior, medial, lateral, superficial, and deep) may be taken from the surgical bed if the specimen x-ray shows the lesion is close to one or more margins. Some surgeons also take additional shavings from the margins as one approach to confirm complete excision of the suspicious lesion. Electrocautery or absorbable ligatures are used to achieve wound hemostasis. Cosmesis may be facilitated by approximation of the surgical defect using 3-0 absorbable sutures. A running subcuticular closure of the skin using 4-0 or 5-0 absorbable monofilament sutures is performed. Wound drainage is usually not required.

Excisional biopsy with needle or seed localization requires a preoperative visit to the mammography suite for placement of a localization wire or a radioactive or magnetic seed that can be detected intraoperatively with a handheld probe. The lesion can also be targeted by sonography in the imaging suite or in the operating room. The lesion to be excised is accurately localized by mammography, and the tip of a thin wire hook or a seed is positioned close to the lesion (Fig. 17-33). Using the wire hook as a guide, or detection of the seed with a handheld probe, the surgeon subsequently excises the suspicious breast lesion while removing a margin of normal-appearing breast tissue. Before the patient leaves the operating room, specimen radiography is performed to confirm complete excision of the suspicious lesion (Fig. 17-34).

Figure 17-32. Lesion to be targeted for excisional biopsy. A. Craniocaudal view of the left breast demonstrating 2 lesions (arrows) to be targeted for needle localization and excision. B. Oblique view demonstrating target lesions. (Used with permission from Dr. Anne Turnbull, Consultant Radiologist/Director of Breast Screening, Royal Derby Hospital, Derby, UK.)
Figure 17-33. Wire localization procedure. Mammographic images of hookwire in place targeting lesions for excision in the left breast (A) and the right breast (B). (Used with permission from Dr. Anne Turnbull, Consultant Radiologist/Director of Breast Screening, Royal Derby Hospital, Derby, UK.)

Figure 17-34. Specimen mammography. Specimen mammograms demonstrating excision of targeted (A) density, (B) calcifications, and (C) spiculated mass seen on preoperative imaging. (Used with permission from Dr. Anne Turnbull, Consultant Radiologist/Director of Breast Screening, Royal Derby Hospital, Derby, UK.)
Sentinel Lymph Node Dissection

Sentinel lymph node (SLN) dissection is primarily used to assess the regional lymph nodes in women with early breast cancers who are clinically node-negative by physical examination and imaging studies.²⁷¹-²⁷⁹ This method also is accurate in women with larger tumors (T3 N0), but nearly 75% of these women will prove to have axillary lymph node metastases on histologic examination, and wherever possible it is better to identify them preoperatively as this will allow a definitive procedure for known axillary disease. SLN dissection has also been reported to be accurate for staging of the axilla after chemotherapy in women with clinically node-negative disease at initial presentation.²⁸⁰,²⁸¹ Tan et al in a review and meta-analysis of 449 cases of SLN biopsy in clinically lymph node-negative disease reported a sensitivity of 93%, giving a false negative rate of 7% with a negative predictive value of 94% and an overall accuracy of 95%.²⁸² Clinical situations where SLN dissection is not recommended include patients with inflammatory breast cancers, those with biopsy proven metastasis, DCIS without mastectomy, or prior axillary surgery. Although limited data are available, SLN dissection appears to be safe in pregnancy when performed with radioisotope alone.

Evidence from large prospective studies suggests that the combination of intraoperative gamma probe detection of radioactive colloid and intraoperative visualization of blue dye (isosulfan blue dye or methylene blue) is more accurate for identification of SLNs than the use of either agent alone. Some surgeons use preoperative lymphoscintigraphy, although it is not required for identification of the SLNs. On the day before surgery, or the day of surgery, the radioactive colloid is injected either in the breast parenchyma around the primary tumor or prior biopsy site, into the subareolar region, or subdermally in proximity to the primary tumor site. With a 25-gauge needle, 0.5 mCi of 0.2-μm technetium 99m-labeled sulfur colloid is injected for same-day surgery, or a higher dose of 2.5 mCi of technetium-labeled sulfur colloid is administered when the isotope is to be injected on the day before surgery. Subdermal injections are given in proximity to the cancer site or in the subareolar location. Later, in the operating room, 3 to 5 mL of blue dye is injected either in the breast parenchyma or in the subareolar location. It is not recommended that the blue dye be used in a subdermal injection because this can result in tattooing of the skin (isosulfan blue dye) or skin necrosis (methylene blue). For nonpalpable cancers, the injection of the technetium-labeled sulfur colloid solution can be guided by ultrasound or by mammographic guidance. In women who have undergone previous excisional biopsy, the injections are made in the breast parenchyma around the biopsy cavity but not into the cavity itself. Women are told preoperatively that the isosulfan blue dye injection will cause a change in the color of their urine and that there is a very small risk of allergic reaction to the dye (1 in 10,000). Anaphylactic reactions have been documented, and some groups administer a regimen of antihistamine, steroids, and a histamine H-2 receptor antagonist preoperatively as a prophylactic regimen to prevent allergic reactions. The use of radioactive colloid is safe, and radiation exposure is very low. Sentinel node dissection can be performed in pregnancy with the radioactive colloid without the use of blue dye.

A hand-held gamma counter is used to transcutaneously identify the location of the SLN. This can help to guide placement of the incision. A 3- to 4-cm incision is made in line with that used for an axillary dissection, which is a curved transverse incision in the lower axilla just below the hairline. After dissecting through the subcutaneous tissue, the surgeon dissects through the axillary fascia, being mindful to identify blue lymphatic channels. Following these channels can lead directly to the SLN and limit the amount of dissection through the axillary tissues. The gamma probe is used to facilitate the dissection and to pinpoint the location of the SLN. As the dissection continues, the signal from the probe increases in intensity as the SLN is approached. The SLN also is identified by visualization of blue dye in the afferent lymph vessel and in the lymph node itself.

Before the SLN is removed, a 10-second in vivo radioactivity count is obtained. After removal of the SLN, a 10-second ex vivo radioactive count is obtained, and the node is then sent to the pathology laboratory for either permanent- or frozen-section analysis. The lowest false-negative rates for SLN dissection have been obtained when all blue lymph nodes and all lymph nodes with counts >10% of the 10-second ex vivo count of the SLN are harvested (“10% rule”). Based on this, the gamma counter is used before closing the axillary wound to measure residual radioactivity in the surgical bed. A search is made for additional SLNs if the counts remain high. This procedure is repeated until residual radioactivity in the surgical bed is less than 10% of the 10-second ex vivo count of the most radioactive SLN and all blue nodes have been removed. Studies have demonstrated that 98% of all positive SLNs will be recovered with the removal of four SLNs; therefore, it is not necessary to remove greater than four SLNs for accurate staging of the axilla.

Results from the NSABP B-32 trial showed that the false-negative rate for SLN dissection is influenced by tumor location, type of diagnostic biopsy, and number of SLNs removed at surgery.²⁴³ The authors reported that tumors located in the lateral breast were more likely to have a false-negative SLN. This may be explained by difficulty in discriminating the hot spot in the axilla when the radioisotope has been injected at the primary tumor site in the lateral breast. Those patients who had undergone an excisional biopsy before the SLN procedure were significantly more likely to have a false-negative SLN. This report further confirms that surgeons should use needle biopsy for diagnosis whenever possible and reserve excisional biopsy for the rare situations in which needle biopsy findings are non-diagnostic or discordant. Finally, removal of a larger number of SLNs at surgery appears to reduce the false-negative rate. In B-32, the false-negative rate was reduced from 17.7% to 10% when two SLNs were recovered and to 6.9% when three SLNs were removed. Yi and associates reported that the number of SLNs that need to be removed for accurate staging is influenced by individual patient and primary tumor factors.²⁸³

In the B-32 trial, SLNs were identified outside the levels I and II axillary nodes in 1.4% of cases. This was significantly influenced by the site of radioisotope injection. When a subareolar or periareolar injection site was used, there were no instances of SLNs identified outside the level I or II axilla, compared with a rate of 20% when a peritumoral injection was used. This supports the overall concept that the SLN is the first site of drainage from the lymphatic vessels of the primary tumor. Although many patients will have similar drainage patterns from injections given at the primary tumor site and at the subareolar plexus, some patients will have extra-axillary drainage, either alone or in combination with axillary node drainage, and this is best assessed with a peritumoral injection of the radioisotope. Kong et al reported that internal mammary node drainage on preoperative lymphoscintigraphy was associated with
worse distant disease-free survival in early-stage breast cancer patients.284

Breast Conservation

Breast conservation involves resection of the primary breast cancer with a margin of normal-appearing breast tissue, adjuvant radiation therapy, and assessment of regional lymph node status.285,286 Resection of the primary breast cancer is alternatively called segmental mastectomy, lumpectomy, partial mastectomy, wide local excision, and takedown. For many women with stage I or II breast cancer, breast-conserving therapy (BCT) is preferable to total mastectomy because BCT produces survival rates equivalent to those after total mastectomy while preserving the breast.287 Six prospective randomized trials have shown that overall and disease-free survival rates are similar with BCT and mastectomy; however, three of the studies showed higher local-regional failure rates in patients undergoing BCT. In two of these studies, there were no clear criteria for histologically negative margins.285,287 Data from the EBCTCG meta-analysis revealed that the addition of radiation reduces recurrence by half and improves survival at year 15 by about a sixth.288 When all of this information is taken together, BCT is considered to be oncologically equivalent to mastectomy.

In addition to being equivalent to mastectomy in terms of oncologic safety, BCT appears to offer advantages over mastectomy with regard to quality of life and aesthetic outcomes. BCT allows for preservation of breast shape and skin as well as preservation of sensation, and it provides an overall psychologic advantage associated with breast preservation.

Breast conservation surgery is currently the standard treatment for women with stage 0, I, or II invasive breast cancer. Women with DCIS require only resection of the primary cancer and adjuvant radiation therapy without assessment of regional lymph nodes. When a lumpectomy is performed, a curvilinear incision lying concentric to the nipple-areola complex is made in the skin overlying the breast cancer when the tumor is in the upper aspect of the breast. Radial incisions are preferred when the tumor is in the lower aspect of the breast. Skin excision is not necessary unless there is direct involvement of the overlying skin by the primary tumor. The breast cancer is removed with an envelope of normal-appearing breast tissue that is adequate to achieve a cancer-free margin. Significant controversy has existed on the appropriate margin width for BCT.267 However, recently the SSO and ASTRO developed a consensus statement, supported by data from a systematic review data, encouraging “no tumor on ink” to be the standard definition of a negative margin for invasive stages I and II breast cancer in patients who undergo breast conserving surgery with whole-breast irradiation. The meta-analysis found that increasing the margin width does not affect local recurrence rates as long as the inked or transected margin is microscopically negative.267,268 Specimen X-ray should routinely be performed to confirm the lesion has been excised. Specimen orientation is performed by the surgeon. Additional margins from the surgical bed are taken as needed to provide a histologically negative margin. Requests for determination of ER, PR, and HER2 status are conveyed to the pathologist.

It is the surgeon’s responsibility to ensure complete removal of cancer in the breast. Ensuring surgical margins that are free of breast cancer will minimize the chances of local recurrence and will enhance cure rates. If negative margins are not obtainable with reexcision, mastectomy is required. SLN is performed before removal of the primary breast tumor. When indicated, intraoperative assessment of the sentinel node can proceed while the segmental mastectomy is being performed.

The use of oncoplastic surgery can be entertained at the time of segmental mastectomy or at a later time to improve the overall aesthetic outcome. The use of oncoplastic techniques range from a simple reshaping of breast tissue to local tissue rearrangement to the use of pedicled flaps or breast reduction techniques. The overall goal is to achieve the best possible aesthetic result. In determining which patients are candidates for oncoplastic breast surgery, several factors should be considered, including the extent of the resection of breast tissue necessary to achieve negative margins, the location of the primary tumor within the breast, and the size of the patient’s breast and body habitus. Oncoplastic techniques are of prime consideration when (a) a significant area of breast skin will need to be resected with the specimen to achieve negative margins; (b) a large volume of breast parenchyma will be resected resulting in a significant defect; (c) the tumor is located between the nipple and the inframammary fold, an area often associated with unfavorable cosmetic outcomes; or (d) excision of the tumor and closure of the breast may result in malpositioning of the nipple.

Mastectomy and Axillary Dissection

A skin-sparing mastectomy removes all breast tissue, the nipple-areola complex, and scars from any prior biopsy procedures.293,294 There is a recurrence rate of less than 6% to 8%, comparable to the long-term recurrence rates reported with standard mastectomy, when skin-sparing mastectomy is used for patients with Tis to T3 cancers. A total (simple) mastectomy without skin sparing removes all breast tissue, the nipple-areola complex, and skin. An extended simple mastectomy removes all breast tissue, the nipple-areola complex, skin, and the level I axillary lymph nodes. A modified radical (“Patey”) mastectomy removes all breast tissue, the nipple-areola complex, skin, and the levels I, II, and III axillary lymph nodes; the pectoralis major that was divided and removed by Patey may be simply divided, giving improved access to level III nodes, and then left in situ, or occasionally the axillary clearance can be performed without dividing pectoralis minor. The Halsted radical mastectomy removes all breast tissue and skin, the nipple-areola complex, the pectoralis major and pectoralis minor muscles, and the levels I, II, and III axillary lymph nodes. The use of systemic chemotherapy and hormonal therapy as well as adjuvant radiation therapy for breast cancer have nearly eliminated the need for the radical mastectomy.

Nipple-areolar sparing mastectomy has been popularized over the last decade especially for risk-reducing mastectomy in high risk women. For those patients with a cancer diagnosis, many consider the following factors for eligibility: tumor located more than 2 to 3 cm from the border of the areola, smaller breast size, minimal ptosis, no prior breast surgeries with periareolar incisions, body mass index less than 40 kg/m², no active tobacco use, no prior breast irradiation, and no evidence of collagen vascular disease.

For a variety of biologic, economic, and psychosocial reasons, some women desire mastectomy rather than breast conservation. Women who are less concerned about cosmesis may view mastectomy as the most expeditious and desirable therapeutic option because it avoids the cost and inconvenience of radiation therapy. Some women whose primary breast cancers cannot be excised with a reasonable cosmetic result or those who have extensive microcalcifications are best treated with
mastectomy. Similarly, women with large cancers that occupy the subareolar and central portions of the breast and women with multicentric primary cancers also undergo mastectomy.

**Modified Radical Mastectomy**

A modified radical mastectomy preserves the pectoralis major muscle with removal of levels I, II, and III (apical) axillary lymph nodes. The operation was first described by David Patey, a surgeon at St Bartholomew’s Hospital London, who reported a series of cases where he had removed the pectoralis minor muscle allowing complete dissection of the level III axillary lymph nodes while preserving the pectoralis major and the lateral pectoral nerve. A modified radical mastectomy permits preservation of the medial (anterior thoracic) pectoral nerve, which courses in the lateral neurovascular bundle of the axilla and usually penetrates the pectoralis minor to supply the lateral border of the pectoralis major. Anatomic boundaries of the modified radical mastectomy are the anterior margin of the latissimus dorsi muscle laterally, the midline of the sternum medially, the subclavius muscle superiorly, and the caudal extension of the breast 2 to 3 cm inferior to the inframammary fold inferiorly. Skin-flap thickness varies with body habitus but ideally is 7 to 8 mm inclusive of skin and telasubcutanea (Fig. 17-35). Once the skin flaps are fully developed, the fascia of the pectoralis major muscle and the overlying breast tissue are elevated off the underlying musculature, which allows for the complete removal of the breast (Fig. 17-36).

Subsequently, an axillary lymph node dissection is performed. The most lateral extent of the axillary vein is identified, and the areolar tissue of the lateral axillary space is elevated as the vein is cleared on its anterior and inferior surfaces. The areolar tissues at the junction of the axillary vein and the anterior edge of the latissimus dorsi muscle, which include the lateral and subscapular lymph node groups (level I), are cleared. Care is taken to preserve the thoracodorsal neurovascular bundle. The dissection then continues medially with clearance of the central axillary lymph node group (level II). The long thoracic nerve of Bell is identified and preserved as it travels in the investing fascia of the serratus anterior muscle. Every effort is made to preserve this nerve because permanent disability with a winged scapula and shoulder weakness will follow denervation of the serratus anterior muscle. Patey divided the pectoralis minor and removed it to allow access right up to the apex of the axilla. The pectoralis minor muscle is usually divided at the tendinous portion near its insertion onto the coracoid process (Fig. 17-37 inset), which allows dissection of the axillary vein medially to the costoclavicular (Halsted’s) ligament. Finally, the breast and axillary contents are removed from the surgical bed and are sent for pathologic assessment. In his modified radical mastectomy, Patey removed the pectoralis minor muscle. Many surgeons now divide only the tendon of the pectoralis minor muscle at its insertion onto the coracoid process while leaving the rest of the muscle intact, which still provides good access to the apex of the axilla.

![Figure 17-35. Modified radical mastectomy: elevation of skin flaps. Skin flaps are 7 to 8 mm in thickness, inclusive of the skin and telasubcutanea. (Visual Art: © 2013. The University of Texas MD Anderson Cancer Center.)(fig17-35)](image)

![Figure 17-36. Modified radical mastectomy after resection of breast tissue. The pectoralis major muscle is cleared of its fascia as the overlying breast is elevated. The latissimus dorsi muscle is the lateral boundary of the dissection. (Visual Art: © 2013. The University of Texas MD Anderson Cancer Center.)](image)
Seromas beneath the skin flaps or in the axilla represent the most frequent complication of mastectomy and axillary lymph node dissection, reportedly occurring in as many as 30% of cases. The use of closed-system suction drainage reduces the incidence of this complication. Catheters are retained in the wound until drainage diminishes to <30 mL per day. Wound infections occur infrequently after a mastectomy, and the majority are a result of skin-flap necrosis. Cultures of specimens taken from the infected wound for aerobic and anaerobic organisms, debridement, and antibiotic therapy are effective management. Moderate or severe hemorrhage in the postoperative period is rare and is best managed with early wound exploration for control of hemorrhage and reestablishment of closed-system suction drainage. The incidence of functionally significant lymphedema after a modified radical mastectomy is approximately 20% but can be as high as 50% to 60% when postoperative radiation is employed. Extensive axillary lymph node dissection, the delivery of radiation therapy, the presence of pathologic lymph nodes, and obesity are predisposing factors. Patients should be referred to physical therapy at the earliest signs of lymphedema to prevent progression to the later stages. The use of individually fitted compressive sleeves and complex decongestive therapy may be necessary.

Reconstruction of the Breast and Chest Wall

The goals of reconstructive surgery after a mastectomy for breast cancer are wound closure and breast reconstruction, which is either immediate or delayed. In most cases, wound closure after mastectomy is accomplished with simple approximation of the wound edges. However, if a more radical removal of skin and subcutaneous tissue is necessary, a pedicled myocutaneous flap from the latissimus dorsi muscle is generally the best approach for wound coverage. A skin graft provides functional coverage that will tolerate adjuvant radiation therapy; however, this is not preferred because poor graft adherence may delay delivery of radiation therapy. Breast reconstruction after risk-reducing mastectomy or after mastectomy for early-stage breast cancer may be performed at the same time as the mastectomy. This allows for a skin-sparing mastectomy to be performed, which offers the best overall cosmetic outcomes. Reconstruction can proceed with an expander/implant reconstruction or with autologous tissue such as a pedicled myocutaneous flap or a free flap using microvascular techniques. In patients with locally advanced breast cancer, reconstruction is often delayed until after completion of adjuvant radiation therapy to ensure that local-regional control of disease is obtained. The expected use of postmastectomy radiotherapy should also be considered as a reason for delayed reconstruction as radiotherapy to a reconstructed breast has been reported to result in inferior cosmetic outcomes. Consideration can be made for placement of a tissue expander to allow for skin-sparing, but this should be discussed with the radiation oncologist and other members of the treatment team. If chest wall coverage is needed to replace a large skin or soft tissue defect, many different types of myocutaneous flaps are employed, but the latissimus dorsi and the rectus abdominis myocutaneous flaps are most frequently used. The latissimus dorsi myocutaneous flap consists of a skin paddle based on the underlying latissimus dorsi muscle, which
is supplied by the thoracodorsal artery with contributions from the posterior intercostal arteries. A transverse rectus abdominis myocutaneous (TRAM) flap consists of a skin paddle based on the underlying rectus abdominis muscle, which is supplied by vessels from the deep inferior epigastric artery. The free TRAM flap uses microvascular anastomoses to establish blood supply to the flap. When the bony chest wall is involved with cancer, resection of a portion of the bony chest wall is indicated. If only one or two ribs are resected and soft tissue coverage is provided, reconstruction of the bony defect is usually not necessary because scar tissue will stabilize the chest wall. If more than two ribs are sacrificed, it is advisable to stabilize the chest wall with prosthetic material, which is then covered with soft tissue by using a latissimus dorsi or TRAM flap.

NONSURGICAL BREAST CANCER THERAPIES

Radiation Therapy

Radiation therapy is used for all stages of breast cancer depending on whether the patient is undergoing BCT or mastectomy. Adjunct radiation for patients with DCIS and early-stage breast cancer have been described previously in this chapter. Those women treated with mastectomy who have cancer at the surgical margins are at sufficiently high risk for local recurrence to warrant the use of adjuvant radiation therapy to the chest wall postoperatively. Women with metastatic disease involving four or more axillary lymph nodes and premenopausal women with metastatic disease involving one to three lymph nodes also are at increased risk for recurrence and are candidates for the use of chest wall and supraclavicular lymph node radiation therapy. In advanced local-regional breast cancer (stage IIIA or IIIB), women are at high risk for recurrent disease after surgical therapy, and adjuvant radiation therapy is used to reduce the risk of recurrence. Current recommendations for stages IIIA and IIIB breast cancer are (a) adjuvant radiation therapy to the breast and supraclavicular lymph nodes after neoadjuvant chemotherapy and segmental mastectomy with or without axillary lymph node dissection, (b) adjuvant radiation therapy to the chest wall and supraclavicular lymph nodes after neoadjuvant chemotherapy and mastectomy with or without axillary lymph node dissection, and (c) adjuvant radiation therapy to the chest wall and supraclavicular lymph nodes after segmental mastectomy or mastectomy with axillary lymph node dissection and adjuvant chemotherapy. Data from the EBCCTCG has shown improvements in local-regional control and survival in patients treated with mastectomy and postmastectomy radiation therapy for one to three positive axillary lymph nodes. This data is based on clinical trials from the era of axillary lymph node dissection for staging prior to the routine use of sentinel lymph node dissection. It is likely that the volume of disease in the earlier trials was greater overall than what is currently seen in patients who have small volume metastases detected at sentinel node dissection. It is important to include all multidisciplinary team members (medical oncology, plastic surgery, radiation oncology, and surgical oncology) regarding the risks and benefits of postmastectomy radiation therapy in patients with one to three positive nodes.

The use of partial breast irradiation (APBI) for patients treated with breast-conserving surgery has also been previously described. APBI can be delivered via brachytherapy, external beam radiation therapy using 3D conformal radiation, or intensity-modulated radiation therapy. Although initial results are promising in highly selected low-risk populations, use of APBI should be based on current guidelines or offered in the setting of a prospective trial.

Chemotherapy Adjuvant

Chemotherapy. The Early Breast Cancer Trialists’ Collaborative Group overview analysis of adjuvant chemotherapy demonstrated reductions in the odds of recurrence and death in women ≤70 years of age with stage I, IIA, or IIB breast cancer. For those ≥70 years of age, the lack of definitive clinical trial data regarding adjuvant chemotherapy prevented definitive recommendations. Adjuvant chemotherapy is of minimal benefit to women with negative nodes and cancers ≥0.5 cm in size and is not recommended. Women with negative nodes and cancers 0.6 to 1.0 cm are divided into those with a low risk of recurrence and those with unfavorable prognostic features that portend a higher risk of recurrence and a need for adjuvant chemotherapy. Adverse prognostic factors include blood vessel or lymph vessel invasion, high nuclear grade, high histologic grade, HER2/neu overexpression, and negative hormone receptor status. American Society of Clinical Oncology guidelines suggest that adjuvant chemotherapy should be considered for patients with positive lymph nodes, HER2-positive disease, Adjuvant! Online mortality greater than 10%, grade 3 lymph node negative tumors >5 mm, triple-negative tumors, lympho-vascular invasion, or estimated distant relapse risk of greater than 15% at 10 years based on 21 gene recurrence score. Adjuvant chemotherapy is recommended by the NCCN for women with unfavorable prognostic features. Table 17-14 lists the frequently used chemotherapy regimens for breast cancer. For women with hormone receptor-negative cancers that are >1 cm in size, adjuvant chemotherapy is appropriate.

### Table 17-14

Adjuvant chemotherapy regimens for breast cancer

<table>
<thead>
<tr>
<th>HER-2 NEGATIVE</th>
<th>HER-2 POSITIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Preferred</strong></td>
<td><strong>HER-2-positive</strong></td>
</tr>
<tr>
<td>Dose dense AC → Paclitaxel every 2 weeks</td>
<td>AC → T + trastuzumab +/− pertuzumab (T = paclitaxel) TCH (docetaxel, carboplatin, trastuzumab +/− pertuzumab)</td>
</tr>
<tr>
<td>Dose dense AC → Paclitaxel weekly TC (T = docetaxel)</td>
<td>Other Regimens</td>
</tr>
<tr>
<td>Other Regimens</td>
<td>Docetaxel + cyclophosphamide + trastuzumab</td>
</tr>
<tr>
<td>CMF</td>
<td>FEC → Docetaxel + cyclophosphamide + trastuzumab</td>
</tr>
<tr>
<td>AC → Docetaxel every 3 weeks</td>
<td>FEC → Docetaxel + trastuzumab + pertuzumab</td>
</tr>
<tr>
<td>AC → Paclitaxel weekly TAC (T = docetaxel)</td>
<td>FEC → Paclitaxel + trastuzumab + pertuzumab</td>
</tr>
</tbody>
</table>

A = Adriamycin (doxorubicin); C = cyclophosphamide; E = epirubicin; F = 5-fluorouracil; M = methotrexate; T = Taxane (docetaxel or paclitaxel); → = followed by.

However, women with node-negative hormone receptor–positive cancers and T1 tumors are candidates for antiestrogen therapy with or without chemotherapy. Assessment of overall risk using known prognostic factors or additional testing such as the 21-gene recurrence score assay can help to guide decision making regarding chemotherapy in patients with node-negative, ER-positive breast cancer. For special-type cancers (tubular, mucinous, medullary, etc), which are usually strongly estrogen receptor positive, adjuvant antiestrogen therapy should be advised for cancers >1 cm. For women with node-positive tumors or with a special-type cancer that is >3 cm, the use of chemotherapy is appropriate; those with hormone receptor–positive tumors should receive antiestrogen therapy.

For stage IIIA breast cancer, preoperative chemotherapy with an anthracycline and taxane-containing regimen followed by either a modified radical mastectomy or segmental mastectomy with axillary dissection followed by adjuvant radiation therapy should be considered, especially for estrogen receptor negative disease. While the same regimen may be considered for estrogen receptor positive disease, it is known that these tumors respond less well to chemotherapy with <10% pCR rate overall and <3% pCR rate for lobular cancers. Other options such as neoadjuvant endocrine therapy followed by local-regional treatment or in some cases primary endocrine therapy may be considered depending on other tumor characteristics and the patient’s comorbid conditions and preference.

Neoadjuvant (Preoperative) Chemotherapy. In the early 1970s, the National Cancer Institute in Milan, Italy, initiated two prospective randomized multimodality clinical trials for women with T3 or T4 breast cancer.310 The best results were achieved when surgery was interposed between chemotherapy courses, with 82% local-regional control and 25% having a 5-year disease-free survival. The NSABP B-18 trial evaluated the role of neoadjuvant chemotherapy in women with operable stages II and III breast cancer.296 Women entered into this study were randomly assigned to receive either surgery followed by chemotherapy or neoadjuvant chemotherapy followed by surgery. There was no difference in the 5-year disease-free survival rates for the two groups, but after neoadjuvant chemotherapy there was an increase in the number of lumpectomies performed and a decreased incidence of node positivity. It was suggested that neoadjuvant chemotherapy be considered for the initial management of breast cancers judged too large for initial lumpectomy.

Several prospective clinical trials have evaluated the neoadjuvant approach, and two meta-analyses have been performed, each showing that neoadjuvant vs. adjuvant chemotherapy are equivalent in terms of OS.262,311 These analyses also evaluated local-regional recurrence (LRR) and found that there was an increase in LRR rates for patients receiving neoadjuvant chemotherapy when radiation therapy was used alone without surgery after completion of chemotherapy. Mittendorf and colleagues evaluated a contemporary series of almost 3000 patients treated with breast conserving surgery and radiation therapy who received either neoadjuvant or adjuvant chemotherapy for breast cancer.312 They found that the risk of LRR was driven by biologic factors and disease stage and was not impacted by the timing of chemotherapy delivery. These data highlight the importance of the multidisciplinary management of patients with breast cancer in achieving the best outcomes.

The use of neoadjuvant chemotherapy offers the opportunity to observe the response of the intact primary tumor and any regional nodal metastases to a specific chemotherapy regimen.279 For patients whose tumors remain stable in size or even progress with the initial neoadjuvant chemotherapy regimen, a new regimen may be considered that uses another class of agents, although there is no randomized data confirming this will improve outcome.

After treatment with neoadjuvant chemotherapy, patients are assessed for clinical and pathologic response to the regimen. Patients whose tumors achieve a pathologic complete response to neoadjuvant chemotherapy have been shown to have statistically improved survival outcomes to those of patients whose tumors demonstrate only a partial response, remain stable, or progress on treatment. Researchers at MD Anderson Cancer Center have shown that residual cancer burden (RCB)—categorized into four classes, RCB-0 or pathologic complete response, RCB-1, RCB-2, and RCB-3—is predictive of 10-year relapse-free survival with neoadjuvant chemotherapy in triple negative, ER-positive, and HER2-positive tumors.313 Patients who experience progression of disease during neoadjuvant chemotherapy have the poorest survival.314,315 This means that while patients who achieve a pCR will have a better prognosis based on their response to neoadjuvant chemotherapy. Equally other patients will have a poorer prognosis compared to when they started neoadjuvant therapy based on the nonresponse to treatment. Consequently, the FDA has supported the use of the neoadjuvant platform and pathologic response rates as an endpoint for mechanism of accelerated approval for new agents in high risk early stage breast cancer, though the short-term endpoints (i.e., pCR) have not been shown to correlate with long-term outcomes (i.e., disease free survival and overall survival).

Current NCCN recommendations for treatment of operable advanced local-regional breast cancer are neoadjuvant chemotherapy with an anthracycline-containing or taxane-containing regimen or both, followed by mastectomy or lumpectomy with axillary lymph node dissection if necessary, followed by adjuvant radiation therapy. For patients with HER2-positive breast cancer, trastuzumab and pertuzumab can be combined with chemotherapy in the preoperative setting to increase pathologic complete response rates. For inoperable stage IIIA and for stage IIIB breast cancer, neoadjuvant chemotherapy is used to decrease the local-regional cancer burden. This may then permit subsequent modified radical or radical mastectomy, which is followed by adjuvant radiation therapy.

Nodal Evaluation in Patients Receiving Neoadjuvant Chemotherapy. The management of the axilla after neoadjuvant chemotherapy has not been specifically addressed in randomized trials. Standard practice has been to perform an axillary lymph node dissection after chemotherapy or to perform a sentinel lymph node dissection before chemotherapy for nodal staging before chemotherapy is initiated. A number of small single-institution studies, one multicenter study, and a recent meta-analysis have explored the use of SLN dissection at the completion of chemotherapy. The published results from these studies have demonstrated the feasibility of SLN dissection in breast cancer patients after neoadjuvant chemotherapy. A review of 14 studies with 818 patients showed a false negative rate of 11% with an overall accuracy of 94%.280,281,316 While SLN dissection has been accepted for assessment of the axilla in the clinically node-negative axilla after neoadjuvant chemotherapy, clinicians have been slower to adopt this approach for axillary staging after chemotherapy in patients who started with initial node-positive disease. Several clinical
trials have been performed to evaluate the accuracy of SLN dissection in patients with documented axillary metastases at initial presentation, including ACOSOG Z1071, SENTINA, and SN FNAC. ACOSOG Z1071 (Alliance) analyzed women with clinical T0–T4, N1–N2, M0 breast cancer who underwent both SLN surgery and axillary lymph node dissection (ALND).317 The primary endpoint was the false-negative rate (FNR) of SLN surgery after chemotherapy with clinically node-positive disease with a prespecified endpoint of 10% considered to be an acceptable rate. However, the FNR was found to be 12.6%, though it was lower when dual-agent mapping technique was used and at least three or more SLNs removed.317 The SENTINA and SN FNAC trials had findings similar to Z1071. The results from Z1071 were further analyzed to determine if a clip was placed in the positive node at initial diagnosis and if the clipped node location at surgery (SLN or ALND) was evaluated. Indeed, this showed that identification of the clipped node during the surgical procedure further decreased the FNR.318 The results from the ACOSOG Z1071 (Alliance) trial, in cases presenting with cN1 disease and at least two SLN resections and clipped node was within the SLN specimen, showed that the FNR was 6.8%.318 Caudle et al at MD Anderson Cancer Center performed a prospective study of patients with biopsy-confirmed nodal metastases with a clip placed in the biopsy-proven lymph node, who were treated with neoadjuvant chemotherapy; at the time of surgery these patients underwent SLN dissection with targeting and removal of the clipped node (targeted axillary dissection [TAD]).319 TAD includes SLN surgery and selective localization and removal of the clipped node, with the goal to determine if pathologic changes in the clipped node accurately reflect the status of the nodal basin, and proposing that TAD improves the FNR compared to SLN surgery alone.319 In patients undergoing SLN surgery and ALND (n = 118), the FNR was 10.1% (95% CI, 4.2–19.8), and adding evaluation of the clipped node reduced the FNR to 1.4% (95% CI, 0.03–7.3; P = 0.3). TAD followed by ALND was performed in 85 patients, with an FNR of 2.0% (1 of 50; 95% CI, 0.05–10.7).319 Although the use of dual tracer technique, retrieval of three or more SLNs, and TAD improve axillary staging after neoadjuvant chemotherapy, there is no long-term data about the oncologic safety of omitting ALND in patients who convert from cN1 to cN0 disease at this time.

Neoadjuvant Endocrine Therapy. While initially used in elderly women who were deemed poor candidates for surgery or cytotoxic chemotherapy, neoadjuvant endocrine therapy is being increasingly evaluated in clinical trials. As clinicians have gained experience with neoadjuvant treatment strategies, it is now clear from examination of predictors of complete pathologic response that ER-positive tumors do not shrink in response to chemotherapy as readily as ER-negative tumors.320 Indeed, the pCR rate in ER-negative tumors is approximately three times that of ER-positive tumors. Fisher et al examined the results of the NSABP B-14 and B-20 trials and found that, as age increased, women obtained less benefit from chemotherapy. They recommended that factors including tumor estrogen receptor concentration, nuclear grade, histologic grade, tumor type, and markers of proliferation should be considered in these patients before choosing between the use of chemotherapy and hormonal therapy. If in fact the tumor is estrogen-receptor rich, these patients may benefit more from endocrine therapy in the neoadjuvant setting than they might if they received standard chemotherapy. Neoadjuvant endocrine therapy has been shown to shrink tumors, enabling breast-conserving surgery in women with hormone receptor-positive disease who otherwise would have to be treated with mastectomy, although long-term recurrence rates have not been reported.265 The IMPACT trial evaluated neoadjuvant use of tamoxifen or anastrozole or both in combination in postmenopausal women with ER-positive operable or locally advanced breast cancer.313 While there were no significant differences in objective tumor response among tamoxifen, anastrozole, or a combination of the two, in patients who were initially deemed as mastectomy candidates, only 31% had breast-conserving surgery with tamoxifen, whereas 44% underwent breast-conserving surgery with anastrozole. invasive lobular cancers in particular have been shown to respond poorly to neoadjuvant chemotherapy and may have better response to neoadjuvant endocrine therapy.322-324 A meta-analysis evaluating the response rate and rate of breast conservation surgery with the use of neoadjuvant endocrine therapy compared to combination chemotherapy was recently reported. This meta-analysis included nearly 3500 patients across 20 studies.325 Interestingly, aromatase inhibitors had a similar response, and breast conservation rates in comparison with combination chemotherapy albeit with lower toxicity suggest that neoadjuvant endocrine therapy is an appropriate alternative in ER-positive breast cancers. However, the incidence of complete pathological response was low (<10%) with both approaches. Also, aromatase inhibitors were associated with significantly higher response and breast conservation rates compared with tamoxifen. The ALTER-NATE (Alternate Approaches for Clinical Stage II or III Estrogen Receptor Positive Breast Cancer Neoadjuvant Treatment in Postmenopausal Women) trial is currently evaluating neoadjuvant endocrine therapy with fulvestrant or anastrozole or in combination.

Increasing knowledge of secondary resistance mechanisms to endocrine therapy and cross talk between ER and the PI3K/Akt/mTOR pathway have led to the evaluation of PI3K pathway inhibitors in combination with endocrine therapy. Postmenopausal women with ER-positive early breast cancers were treated with letrozole or letrozole in combination with everolimus, a mTOR inhibitor, in a randomized, phase 2 clinical trial. Clinical response and antiproliferative response, characterized by reduction in Ki67, was superior in the combination arm, suggesting that everolimus can increase efficacy of neoadjuvant letrozole.326 The LORLEI study is evaluating the use of taselisib, a PI3K inhibitor in combination with letrozole compared with letrozole alone. With the approval of CDK 4/6 inhibitors in the metastatic setting, clinical trials are evaluating the use of CDK inhibitors in combination with neoadjuvant endocrine therapy. Neoadjuvant anastrozole in combination with palbociclib, a CDK4/6 inhibitor, has been shown to significantly reduce Ki67, suggesting that CDK4/6 inhibition can increase the efficacy of neoadjuvant endocrine therapy.

With the use of neoadjuvant chemotherapy or endocrine therapy, observation of the response of the intact tumor and/or nodal metastases to a specific regimen could ultimately help to define which patients will benefit from specific therapies in the adjuvant setting. In adjuvant trials the primary endpoint is typically survival, whereas in neoadjuvant trials the endpoints have more often been clinical or pathologic response rates. There are a number of clinical trials underway comparing neoadjuvant chemotherapy and endocrine therapy regimens with pretreatment and posttreatment biopsy samples obtained from the primary tumors in all of the participants. These samples are being subjected to intensive genomic and proteomic analyses that may
help to define a more personalized or individualized approach to breast cancer treatment in the future.

**Antiestrogen Therapy**

**Tamoxifen.** Within the cytosol of breast cancer cells are specific proteins (receptors) that bind and transfer steroid moieties into the cell nucleus to exert specific hormonal effects. The most widely studied hormone receptors are the estrogen receptor and progesterone receptor. Hormone receptors are detectable in >90% of well-differentiated ductal and lobular invasive cancers. Although the receptor status may remain the same between the primary cancer and metastatic disease in the same patient in the majority of cases, there are instances where the status is changed in the metastatic focus; therefore, biopsy of newly diagnosed metastatic disease should be considered for assessment of hormone receptor and HER2 status.

After binding to estrogen receptors in the cytosol, tamoxifen blocks the uptake of estrogen by breast tissue. Clinical responses to antiestrogen and >60% of women with hormone receptor-positive breast cancers but in <10% of women with hormone receptor-negative breast cancers. A meta-analysis by the Early Breast Cancer Trialists’ Collaborative Group showed that adjuvant therapy with tamoxifen for 5 years reduced breast cancer mortality by about a third through the first 15 years of follow-up. This mortality benefit continues to be statistically significant in the second and third 5-year periods (i.e., years 5–9 and 10–15) when the patients are no longer receiving endocrine treatment—the so-called carry-over effect. The analysis also showed a 39% reduction in the risk of cancer in the contralateral breast. The antiestrogens do have defined toxicity, including bone pain, hot flashes, nausea, vomiting, and fluid retention. Thrombotic events occur in <3% of treated women. Cataract surgery is more frequently performed in patients receiving tamoxifen. The Stockholm trial showed that 5 years of tamoxifen was associated with a significant reduction in locoregional recurrences and distant metastasis in postmenopausal women with ER-positive breast cancer. However, an increase in endometrial cancers was observed with long-term tamoxifen use. The NSABP B14 trial evaluated 10 years of tamoxifen compared to 5 years. However, the study was terminated based on interim analyses indicating no additional benefit from tamoxifen beyond 5 years. The ATLAS trial also evaluated the use of tamoxifen for 5 years vs. 10 years in nearly 13,900 women across the world. This study showed that continuing tamoxifen for 10 years vs. 5 years produced a significant reduction in recurrence and mortality. Interestingly, the benefit was not seen in the second 5 years (i.e., years 5–9) while the patients were on treatment, but it was seen from years 10 to 15. One reason the NSABP B14 study was led to conclude that 10 years of tamoxifen was not beneficial was that the follow-up time was shorter. Results of the ATLAS study were also corroborated by the ATtom study. Similarly, extended adjuvant therapy with letrozole after 5 years of tamoxifen was shown to improve disease-free survival without improvement in overall survival except in node-positive patients.

Tamoxifen therapy is also considered for women with DCIS that is found to be ER-positive. The goals of such therapy are to decrease the risk of an ipsilateral recurrence after breast conservation therapy for DCIS and to decrease the risk of a primary invasive breast cancer or a contralateral breast cancer event. Consequently, tamoxifen is not recommended for patients who have had bilateral mastectomies with ER-positive DCIS. With the use of aromatase inhibitors in postmenopausal women, use of adjuvant tamoxifen has increasingly been limited to premenopausal women.

**Aromatase Inhibitors.** In postmenopausal women, aromatase inhibitors are now considered first-line therapy in the adjuvant setting. Currently, three third-generation aromatase inhibitors are approved for clinical use: the reversible nonsteroidal inhibitors anastrozole and letrozole and the irreversible steroidal inhibitor exemestane. While all the aromatase inhibitors have been shown to have similar efficacy with a similar spectrum of adverse effects, the Early Breast Cancer Trialists’ Collaborative Group meta-analyses of 31,920 postmenopausal women with ER-positive early breast cancers treated with tamoxifen or aromatase inhibitors demonstrated that 5 years of aromatase inhibitors reduced the rate of recurrence by 30% and 10-year breast cancer mortality by about 15% compared to 5 years of tamoxifen. The NSABP B42 study evaluated whether an additional 5 years of letrozole improved disease-free survival in postmenopausal women who have completed 5 years of tamoxifen or an aromatase inhibitor. After a median follow-up of 6.9 years, while extended letrozole significantly improved breast cancer-free interval, no improvement in disease-free survival, the primary endpoint, was observed. Recently, the results of the MA-17R study, designed to assess the efficacy of adjuvant letrozole for 10 years, were reported. Similar to NSABP B42, extended letrozole improved disease-free survival without significant improvement in overall survival. Patients who are node-positive, have received adjuvant chemotherapy, with prior receipt of tamoxifen are likely to benefit from long-term use of an aromatase inhibitor.

The aromatase inhibitors are less likely than tamoxifen to cause endometrial cancer but do lead to changes in bone mineral density that may result in osteoporosis and an increased rate of fractures in postmenopausal women. The risk of osteoporosis can be averted by treatment with bisphosphonates. Joint pains are a side effect that affects a significant number of patients. Node-negative and node-positive breast cancer patients whose tumors express hormone receptors should be considered for endocrine therapy in the adjuvant setting. Women with hormone receptor–positive cancers achieve significant reduction in risk of recurrence of breast cancer and mortality from breast cancer through the use of endocrine therapies.

For postmenopausal women with ER-positive, HER2-negative, metastatic breast cancer, available endocrine therapies include nonsteroidal aromatase inhibitors (anastrozole and letrozole); steroidal aromatase inhibitors (exemestane); serum ER modulators (tamoxifen or toremifene); ER down-regulators (fulvestrant); progestin (megestrol acetate); androgens (fluoxymesterone); and high-dose estrogen (ethinyl estradiol). A third generation nonsteroidal aromatase inhibitor or palbociclib, the CDK 4/6 inhibitor, in combination with letrozole may be considered as a treatment option for first-line therapy. Activation of CDK4/CDK6 cell cycle signaling axis has been implicated in mediating endocrine resistance. Consequently, PALOMA-1 evaluated the safety and efficacy of palbociclib in combination with letrozole vs. letrozole alone as first-line treatment for patients with ER-positive, HER2-negative advanced breast cancer. Median progression-free survival (PFS) was doubled with the combination compared to letrozole alone (20.2 months vs. 10.2 months for the letrozole). Based on this, the FDA approved palbociclib in combination with letrozole for the treatment of postmenopausal women with ER-positive, HER2-negative advanced breast cancer as initial treatment. The
benefit of palbociclib in combination with letrozole was subsequently confirmed in a phase 3 trial (PFS 24.8 months vs. 14.5 months for letrozole). Two additional CDK4/6 inhibitors, ribociclib and abemaciclib, have been approved for use in combination with endocrine therapy for patients with hormone receptor–positive advanced breast cancer.

On the other hand, PALOMA-3 compared the combination of palbociclib and fulvestrant to fulvestrant alone in pre- or postmenopausal ER-positive, HER2-negative metastatic breast cancer patients, whose disease progressed on prior endocrine therapy. Premenopausal women also received the GnRH agonist, goserelin. The median PFS was 9.2 months for the combination compared to 3.8 months with fulvestrant alone. Thus, fulvestrant with palbociclib is a potential option for women with metastatic breast cancer who have progressed on prior endocrine therapy. Additionally, abemaciclib in combination with fulvestrant or as single agent is approved for use in ER-positive advanced breast cancers previously treated with endocrine therapy.

In premenopausal women with stage IV ER-positive breast cancer without previous exposure to endocrine therapy, initial treatment with tamoxifen or ovarian suppression/ablation plus aromatase inhibitor with or without CDK4/6 inhibitors are reasonable options.

Activation of the PI3K/mammalian target of rapamycin (mTOR) signal transduction pathway has also been implicated in secondary resistance to estrogen targeting. BOLEbLERO-2 evaluated the use of exemestane in combination with everolimus in postmenopausal women with ER-positive tumors who had progressed or recurred on a nonsteroidal aromatase inhibitor. An improvement in PFS was observed with combination compared to exemestane alone (11 vs. 4.1 months) leading to FDA approval. Similar improvement in PFS was observed with a combination of tamoxifen and everolimus. However, a phase 3 trial of letrozole in combination with temsirolimus, an mTOR inhibitor, did not show any improvement in PFS in aromatase inhibitor–naive metastatic postmenopausal women. Trials evaluating the adjuvant use of mTOR inhibitors and CDK 4/6 inhibitors are currently in progress.

Women whose tumors respond to an endocrine therapy with either shrinkage of their breast cancer (objective response) or long-term stabilization of disease (stable disease) are considered to represent “clinical benefit” and should receive additional endocrine therapy at the time of progression because their chances of a further response remain high. Patients whose tumors progress de novo on an endocrine agent have a low rate of clinical benefit (<20%) to subsequent endocrine therapy; the choice of endocrine or chemotherapy should be considered based on the disease site and extent as well as the patient’s general condition and treatment preference.

The adjuvant use of aromatase inhibitors and recent advances in tumor genome sequencing technologies have enabled the identification of secondary ESR1 mutations. These mutations, typically present in the ligand binding domains, lead to ligand-independent activation of the receptor, mediate resistance to aromatase inhibitors, and are associated with shorter survival. Reported incidence of these mutations is variable (20%–30%) based on prior exposure to aromatase inhibitors and are uncommon in primary breast cancers. Clinical trials evaluating novel selective estrogen receptor degraders with potential activity against these mutations are in progress.

Ablative Endocrine Therapy

In the past, adrenalectomy and/or hypophysectomy were the primary endocrine modalities used to treat metastatic breast cancer, but today these approaches are seldom used. In women who are premenopausal at diagnosis, ovarian ablation can be accomplished by oophorectomy or ovarian radiation. Ovarian suppression can be accomplished by the use of gonadotrophin-hormone releasing hormone agonists, such as goserelin or leuprolide. Evaluation of the combination of goserelin with tamoxifen vs. cyclophosphamide/methotrexate/fluorouracil chemotherapy in premenopausal ER-positive early-stage breast cancers showed that relapse-free survival was superior with endocrine therapy combination, with a similar trend in overall survival. Data from the SOFT and TEXT trials on adjuvant endocrine therapy show that exemestane plus ovarian suppression significantly reduces recurrences as compared with tamoxifen plus ovarian suppression. In these trials, ovarian suppression was achieved with the use of the gonadotropin-releasing hormone agonist triptorelin, oophorectomy, or ovarian irradiation. The disease-free survival was 89% in the tamoxifen plus ovarian suppression group, while it was 93% in exemestane plus ovarian suppression group; however, there was no significant differences in overall survival. In the SOFT trial, while tamoxifen plus ovarian suppression was not superior to tamoxifen alone in terms of disease-free survival, improved outcomes were observed in ovarian suppression in women with a high risk of recurrence. In women who received no adjuvant chemotherapy, no meaningful benefit was obtained with ovarian suppression. Thus, ovarian suppression in combination with an aromatase inhibitor can be considered in select premenopausal women with high-risk features (age <40 years, positive lymph nodes) who warranted adjuvant chemotherapy.

Anti-HER2 Therapy

The determination of tumor HER-2 expression or gene amplification for all newly diagnosed patients with breast cancer is now recommended. It is used to assist in the selection of adjuvant chemotherapy in both node-negative and node-positive patients. Trastuzumab was initially approved for the treatment of HER2/neu-positive breast cancer in patients with metastatic disease. Once efficacy was demonstrated for patients with metastatic disease, the NSABP and the North Central Cancer Treatment Group conducted phase 3 trials that evaluated the impact of adjuvant trastuzumab therapy in patients with early-stage breast cancer. After approval from the FDA, these groups amended their adjuvant trastuzumab trials (B-31 and N9831, respectively), to provide for a joint efficacy analysis. The first joint interim efficacy analysis demonstrated an improvement in 3-year disease-free survival from 75% in the control arm to 87% in the trastuzumab arm (hazard ratio = 0.48, P < .0001). There was an accompanying 33% reduction in mortality in the patients who received trastuzumab (hazard ratio = 0.67, P = 0.015). The magnitude of reduction in hazard for disease-free survival events crossed prespecified early reporting boundaries, so the data-monitoring committees for both groups recommended that randomized accrual to the trials be ended, and the results were subsequently published.

While anthracycline-based adjuvant chemotherapy was considered preferable in HER2-positive breast cancer, the BCIRG 006 compared the use of anthracycline with taxane and trastuzumab (AC-TH) versus taxane, carboplatin chemotherapy with trastuzumab (TCH). With 10 years of follow-up, no statistical significance with regard to disease-free and overall survival was observed. However, there was a similar trend in overall survival.
survival was observed for anthracycline-based chemotherapy. While anthracycline chemotherapy was numerically superior, this was accompanied by an increase in the incidence of leukemia and congestive heart failure. A year of adjuvant trastuzumab is considered standard of care. Two years of adjuvant trastuzumab has been shown to be more effective, although it is associated with more toxicity than 1 year of trastuzumab.\textsuperscript{357} On the other hand, the PHARE trial examined 6 months vs. standard 12 months of trastuzumab. After 3.5 years of follow-up, the study failed to demonstrate that 6 months was noninferior compared to the standard therapy.\textsuperscript{358} Patients with HER2-positive tumors benefit if trastuzumab is added to taxane chemotherapy. Because of overlapping cardiotoxicities, trastuzumab is not usually given concurrently with anthracyclines.

Buzdar and colleagues reported the results of a randomized neoadjuvant trial of trastuzumab in combination with sequential paclitaxel followed by FEC-75 (5-fluorouracil, epirubicin, cyclophosphamide) vs. the same chemotherapy regimen without trastuzumab in 42 women with early-stage operable breast cancer. The pathologic complete response rates in this trial increased from 25% to 66.7% when chemotherapy was given concurrently with trastuzumab.\textsuperscript{301} A subsequent report that included additional patients treated with concurrent chemotherapy and trastuzumab further confirmed the high pathologic complete response rates and continued to show that cardiac function was preserved.\textsuperscript{302}

While novel agents have been approved for the treatment of women with metastatic HER2-positive breast cancers, currently trastuzumab is the only HER2-targeted agent approved for use in the adjuvant setting. Lapatinib is a dual tyrosine kinase inhibitor that targets both HER2 and EGFR. It was approved for use with capecitabine in patients with HER2-positive metastatic disease. Adjuvant lapatinib was shown to be inferior to trastuzumab, and the combination of lapatinib with trastuzumab did yield a significant improvement in disease-free survival compared to trastuzumab alone. Ado-trastuzumab emtansine (T-DM1) is approved for HER2-positive metastatic breast cancer patients who have previously received trastuzumab and a taxane either separately or in combination. T-DM1 is an antibody drug conjugate that incorporates the HER2 targeted activity of trastuzumab with the cytotoxic activity of DM1, a microtubule inhibitory agent leading to apoptosis.\textsuperscript{359}

Pertuzumab is a humanized monoclonal antibody that binds at a different epitope of the HER2 extracellular domain (subdomain II) and prevents dimerization of HER2 with other members of the family, primarily HER3. In the metastatic setting, it is approved in combination with trastuzumab and docetaxel for patients with metastatic HER2-positive breast cancer who have not received prior HER2-targeted therapy or chemotherapy for metastatic disease.\textsuperscript{360} In the neoadjuvant setting, pertuzumab is approved in combination with trastuzumab and docetaxel in HER2-positive, early stage breast cancers that are greater than 2 cm or node-positive. However, this approval is based on improvement in pathologic complete response rate, and not data based on improvement in event free or overall survival.\textsuperscript{361,362} In the NeoSphere trial, neoadjuvant use of pertuzumab with trastuzumab and docetaxel led to nearly a 17% increase in pathologic complete response in the breast ($P = .0141$).\textsuperscript{361} While in the TRYPHAENA study, pathologic complete responses ranging from 57% to 66% were observed with neoadjuvant pertuzumab and trastuzumab combination given with anthracycline-containing or nonanthracycline-containing chemotherapy.\textsuperscript{362} With the use of dual antibody therapy, currently there is significant interest in identifying patients who can avoid chemotherapy and potentially be treated with HER2-targeted agents alone. The NeoSphere study showed 27% pathologic complete response in HER2-positive, ER-negative, breast cancer patients treated with pertuzumab and trastuzumab alone. Pertuzumab was recently FDA approved in combination with trastuzumab and chemotherapy in the adjuvant setting in HER2 amplified breast cancers with high risk of recurrence. Approval is based on APHINITY trial showing that the addition of pertuzumab improved invasive disease free survival (7.1%) compared to placebo (8.7%) (HR 0.82, 95% CI: 0.67, 1.00; $p = 0.047$). Overall survival data is not mature.

The ExteNET study evaluated the use of neratinib, an irreversible inhibitor of EGFR, HER2, and HER4, in HER2-positive early stage patients who have completed adjuvant trastuzumab. A year of neratinib after completion of chemotherapy and trastuzumab-based adjuvant therapy significantly improved 2-year disease-free survival, the primary endpoint.\textsuperscript{363} After two years, invasive disease free survival was 94.2% in patients treated with neratinib compared with 91.9% in those receiving placebo (HR 0.66; 95% CI: 0.49, 0.90, $p = 0.008$) leading to FDA approval for HER2 amplified breast cancers following a year of adjuvant trastuzumab.

In addition to amplifications or copy number alterations, activating mutations or single nucleotide variants in HER2 have been described (2%).\textsuperscript{364} Typically observed in ER-positive breast cancers, a higher prevalence of HER2 mutations have been reported in invasive lobular carcinomas, particularly in the pleomorphic subtype.\textsuperscript{365} These mutations, usually exclusive with HER2 amplification, are observed in kinase or extracellular domains and predict for responses or resistance to HER2-targeting agents.\textsuperscript{366,367} A phase 2 trial of neratinib in HER2-mutated metastatic breast cancers showed a clinical benefit rate of 36% with one complete response and one partial response in a heavily pretreated population. A clinical trial evaluating the combination of neratinib with fulvestrant, in HER2-mutated, ER-positive breast cancers, is in progress.

### SPECIAL CLINICAL SITUATIONS

#### Nipple Discharge

**Unilateral Nipple Discharge.** Nipple discharge is a finding that can be seen in a number of clinical situations. It may be suggestive of cancer if it is spontaneous, unilateral, localized to a single duct, present in women $\geq 40$ years of age, bloody, or associated with a mass. A trigger point on the breast may be present so that pressure around the nipple-areolar complex induces discharge from a single duct. In this circumstance, mammography and ultrasound are indicated for further evaluation. A ductogram also can be useful and is performed by cannulating a single discharging duct with a small nylon catheter or needle and injecting 1.0 mL of water-soluble contrast solution. Nipple discharge associated with a cancer may be clear, bloody, or serous. Testing for the presence of hemoglobin is helpful, but hemoglobin may also be detected when nipple discharge is secondary to an intraductal papilloma or duct ectasia. Definitive diagnosis depends on excisional biopsy of the offending duct and any associated mass lesion. A 3.0 lacrimal duct probe can be used to identify the duct that requires excision. Another approach is to inject methylene blue dye within
the duct after ductography. The nipple must be sealed with collodion or a similar material so that the blue dye does not discharge through the nipple but remains within the distended duct facilitating its localization. Localization with a wire or seed is performed when there is an associated mass that lies >2.0 to 3.0 cm from the nipple.

**Bilateral Nipple Discharge.** Nipple discharge is suggestive of a benign condition if it is bilateral and multiduct in origin, occurs in women ≤39 years of age, or is milky or blue-green. Prolactin-secreting pituitary adenomas are responsible for bilateral nipple discharge in <2% of cases. If serum prolactin levels are repeatedly elevated, plain radiographs of the sellar turcica are indicated, and thin section CT scan is required. Optical nerve compression, visual field loss, and infertility are associated with large pituitary adenomas.

**Axillary Lymph Node Metastases in the Setting of an Unknown Primary Cancer**

A woman who presents with an axillary lymph node metastasis is consistent with a breast cancer metastasis has a 90% probability of harboring an occult breast cancer. However, axillary lymphadenopathy is the initial presenting sign in only 1% of breast cancer patients. Fine-needle aspiration biopsy or core-needle biopsy can be used to establish the diagnosis when an enlarged axillary lymph node is identified. When metastatic cancer is found, immunohistochemical analysis may classify the cancer as epithelial, melanocytic, or lymphoid in origin. The presence of hormone receptors (estrogen or progesterone receptors) suggests metastasis from a breast cancer but is not diagnostic. The search for a primary cancer includes careful examination of the thyroid, breast, and pelvis, including the rectum. The breast should be examined with diagnostic mammography, ultrasonography, and MRI to evaluate for an occult primary lesion. Further radiologic and laboratory studies should include chest radiography and liver function studies. Additional imaging of the chest, abdomen, and skeleton may be indicated if the extent of nodal involvement is consistent with stage III breast cancer. Suspicious findings on mammography, ultrasonography, or MRI necessitate breast biopsy. When a breast cancer is found, treatment consists of an axillary lymph node dissection with a mastectomy or preservation of the breast followed by whole-breast radiation therapy. Chemotherapy and endocrine therapy should be considered.

**Breast Cancer During Pregnancy**

Breast cancer occurs in 1 of every 3000 pregnant women, and axillary lymph node metastases are present in up to 75% of these women. The average age of the pregnant woman with breast cancer is 34 years. Fewer than 25% of the breast nodules developing during pregnancy and lactation will be cancerous. Ultrasonography and needle biopsy specimens are used in the diagnosis of these nodules. Mammography is rarely indicated because of its decreased sensitivity during pregnancy and lactation; however, the fetus can be shielded if mammography is needed. Approximately 30% of the benign conditions encountered will be unique to pregnancy and lactation (galactoceles, lobular hyperplasia, lactating adenoma, and mastitis or abscess). Once a breast cancer is diagnosed, complete blood count, chest radiography (with shielding of the abdomen), and liver function studies are performed.

Because of the potential deleterious effects of radiation therapy on the fetus, radiation cannot be considered until the fetus is delivered. A modified radical mastectomy can be performed during the first and second trimesters of pregnancy, even though there is an increased risk of spontaneous abortion after first-trimester anesthesia. During the third trimester, lumpectomy with axillary node dissection can be considered if adjuvant radiation therapy is deferred until after delivery. Lactation is suppressed. Chemotherapy administered during the first trimester carries a risk of spontaneous abortion and a 12% risk of birth defects. There is no evidence of teratogenicity resulting from administration of chemotherapeutic agents in the second and third trimesters. For this reason, many clinicians now consider the optimal strategy to be delivery of chemotherapy in the second and third trimesters as a neoadjuvant approach, which allows local therapy decisions to be made after the delivery of the baby. Pregnant women with breast cancer often present at a later stage of disease because breast tissue changes that occur in the hormone-rich environment of pregnancy obscure early cancers. However, pregnant women with breast cancer have a prognosis, stage by stage, that is similar to that of nonpregnant women with breast cancer.

**Male Breast Cancer**

Fewer than 1% of all breast cancers occur in men. The incidence appears to be highest among North Americans and the British, in whom breast cancer constitutes as much as 1.5% of all male cancers. Jewish and African-American men have the highest incidence. Male breast cancer is preceded by gynecomastia in 20% of men. It is associated with radiation exposure, estrogen therapy, testicular feminizing syndromes, and Klinefelter’s syndrome (XXY). Breast cancer is rarely seen in young males and has a peak incidence in the sixth decade of life. A firm, nontender mass in the male breast requires investigation. Skin or chest wall fixation is particularly worrisome.

DCIS makes up <15% of male breast cancer, whereas infiltrating duct carcinoma makes up >85%. Special-type cancers, including infiltrating lobular carcinoma, have occasionally been reported. Male breast cancer is staged in the same way as female breast cancer, and stage by stage, men with breast cancer have the same survival rate as women. Overall, men do worse because of the more advanced stage of their cancer (stage II, III or IV) at the time of diagnosis. The treatment of male breast cancer is surgical, with the most common procedure being a modified radical mastectomy. SLN dissection has been shown to be feasible and accurate for nodal assessment in men presenting with a clinically node-negative axilla. Adjuvant radiation therapy is appropriate in cases in which there is a high risk for local-regional recurrence. Approximately 80% of male breast cancers are hormone receptor–positive, and adjuvant tamoxifen is considered. Systemic chemotherapy is considered for men with hormone receptor-negative cancers and for men with large primary tumors, multiple positive nodes, and locally advanced disease.

**Phyllodes Tumors**

The nomenclature, presentation, and diagnosis of phyllodes tumors (including cystosarcoma phylloides) have posed many problems for surgeons. These tumors are classified as benign, borderline, or malignant. Borderline tumors have a greater potential for local recurrence.

Mammographic evidence of calcifications and morphologic evidence of necrosis do not distinguish between benign, borderline, and malignant phyllodes tumors. Consequently, it is difficult to differentiate benign phyllodes tumors from the
malignant variant and from fibroadenomas. Phyllodes tumors are usually sharply demarcated from the surrounding breast tissue, which is compressed and distorted. Connective tissue composes the bulk of these tumors, which have mixed gelatinous, solid, and cystic areas. Cystic areas represent sites of infarction and necrosis. These gross alterations give the gross cut tumor surface its classical leaf-like (phyllodes) appearance. The stroma of a phyllodes tumor generally has greater cellular activity than that of a fibroadenoma. After microdissection to harvest clusters of stromal cells from fibroadenomas and from phyllodes tumors, molecular biology techniques have shown the stromal cells of fibroadenomas to be either polyclonal or monoclonal (derived from a single progenitor cell), whereas those of phyllodes tumors are always monoclonal.

Most malignant phyllodes tumors (Fig. 17-38) contain liposarcomatous or rhabdomyosarcomatous elements rather than fibrosarcomatous elements. Evaluation of the number of mitoses and the presence or absence of invasive foci at the tumor margins may help to identify a malignant tumor. Small phyllodes tumors are excised with a margin of normal-appearing breast tissue. When the diagnosis of a phyllodes tumor with suspicious malignant elements is made, reexcision of the biopsy specimen site to ensure complete excision of the tumor with a 1-cm margin of normal-appearing breast tissue is indicated. Large phyllodes tumors may require mastectomy. Axillary dissection is not recommended because axillary lymph node metastases rarely occur.

**Inflammatory Breast Carcinoma**

Inflammatory breast carcinoma (stage IIIB) accounts for <3% of breast cancers. This cancer is characterized by the skin changes of brawny induration, erythema with a raised edge, and edema (peau d’orange). Permeation of the dermal lymph vessels by cancer cells is seen in skin biopsy specimens. There may be an associated breast mass (Fig. 17-39). The clinical differentiation of inflammatory breast cancer may be extremely difficult, especially when a locally advanced scirrhous carcinoma invades dermal lymph vessels in the skin to produce peau d’orange and lymphangitis (Table 17-15). Inflammatory breast cancer also may be mistaken for a bacterial infection of the breast. More than 75% of women who have inflammatory breast cancer present with palpable axillary lymphadenopathy, and distant metastases also are frequently present. A PET-CT scan should be considered at the time of diagnosis to rule out concurrent metastatic disease. A report of the SEER program described distant metastases at diagnosis in 25% of white women with inflammatory breast carcinoma.

Surgery alone and surgery with adjuvant radiation therapy have produced disappointing results in women with inflammatory breast cancer. However, neoadjuvant chemotherapy with an anthracycline-containing regimen may affect dramatic regressions in up to 75% of cases. Tumors should be assessed for HER2 and hormone receptors with treatment dictated based on receptor status. Modified radical mastectomy is performed after demonstrated response to systemic therapy to remove residual cancer from the chest wall and axilla. Adjuvant chemotherapy may be indicated depending on final pathologic assessment of the breast and regional nodes. Finally, the chest wall and the

---

**Figure 17-38.** A. Malignant phyllodes tumor (cystosarcoma-phyllodes). B. Histologic features of a malignant phyllodes tumor (hematoxylin and eosin stain, ×100).

**Figure 17-39.** Inflammatory breast carcinoma. Stage IIIB cancer of the breast with erythema, skin edema (peau d’orange), nipple retraction, and satellite skin nodules.
Table 17-15

<table>
<thead>
<tr>
<th>INFLAMMATORY</th>
<th>NONINFLAMMATORY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dermal lymph vessel invasion is present with or without inflammatory changes.</td>
<td>Inflammatory changes are present without dermal lymph vessel invasion.</td>
</tr>
<tr>
<td>Cancer is not sharply delineated.</td>
<td>Cancer is better delineated.</td>
</tr>
<tr>
<td>Erythema and edema frequently involve &gt;33% of the skin over the breast.</td>
<td>Erythema is usually confined to the lesion, and edema is less extensive.</td>
</tr>
<tr>
<td>Lymph node involvement is present in &gt;75% of cases.</td>
<td>Lymph nodes are involved in approximately 50% of the cases.</td>
</tr>
<tr>
<td>Distant metastases are more common at the initial presentation (25% of cases).</td>
<td>Distant metastases are less common at presentation.</td>
</tr>
</tbody>
</table>


supraclavicular, internal mammary, and axillary lymph node basins receive adjuvant radiation therapy. This multimodal approach results in 5-year survival rates that approach 30%. Patients with inflammatory breast cancer should be encouraged to participate in clinical trials.

Rare Breast Cancers

Squamous Cell (Epidermoid) Carcinoma. Squamous cell (epidermoid) carcinoma is a rare cancer that arises from metaplasia within the duct system and generally is devoid of distinctive clinical or radiographic characteristics. Regional metastases occur in 25% of patients, whereas distant metastases are rare.

Adenoid Cystic Carcinoma. Adenoid cystic carcinoma is very rare, accounting for <0.1% of all breast cancers. It is typically indistinguishable from adenoid cystic carcinoma arising in salivary tissues. These cancers are generally 1 to 3 cm in diameter at presentation and are well circumscribed. Axillary lymph node metastases are rare, but deaths from pulmonary metastases have been reported.

Apocrine Carcinomas. Apocrine carcinomas are well-differentiated cancers that have rounded vesicular nuclei and prominent nucleoli. There is a very low mitotic rate and little variation in cellular features. However, apocrine carcinomas may display an aggressive growth pattern.

Sarcomas. Sarcomas of the breast are histologically similar to soft tissue sarcomas at other anatomic sites. This diverse group includes fibrosarcoma, malignant fibrous histiocytoma, liposarcoma, leiomyosarcoma, malignant schwannoma, rhabdomyosarcoma, osteogenic sarcoma, and chondrosarcoma. The clinical presentation is typically that of a large, painless breast mass with rapid growth. Diagnosis is by core-needle biopsy or by open incisional biopsy. Sarcomas are graded based on cellularity, degree of differentiation, nuclear atypia, and mitotic activity. Primary treatment is wide local excision, which may necessitate mastectomy. Axillary dissection is not indicated unless there is biopsy proven lymph node involvement. Angiosarcomas are classified as de novo, as postradiation, or as arising in association with postmastectomy lymphedema. In 1948, Stewart and Treves described lymphangiosarcoma of the upper extremity in women with ipsilateral lymphedema after radical mastectomy. Angiosarcoma is now the preferred name. The average interval between modified radical or radical mastectomy and the development of an angiosarcoma is 7 to 10 years. Sixty percent of women developing this cancer have a history of adjuvant radiation therapy. Forequarter amputation may be necessary to palliate the ulcerative complications and advanced lymphedema.

Lymphomas. Primary lymphomas of the breast are rare, and there are two distinct clinicopathologic variants. One type occurs in women ≤39 years of age, is frequently bilateral, and has the histologic features of Burkitt’s lymphoma. The second type is seen in women ≥40 years of age and is usually of the B-cell type. Breast involvement by Hodgkin’s lymphoma has been reported. An occult breast lymphoma may be diagnosed after detection of palpable axillary lymphadenopathy. Treatment depends on the stage of disease. Lumpectomy or mastectomy may be required. Axillary dissection for clearance of disease may be necessary. Recurrent or progressive local-regional disease is best managed by chemotherapy and radiation therapy. The prognosis is favorable, with 5- and 10-year survival rates of 74% and 51%, respectively. More recently anaplastic large cell lymphoma has been described in association with breast implants for cosmetic or reconstructive purposes. This disease is treated with complete excision of the implant capsule with any associated soft tissue mass. More advanced cases may require systemic therapy and radiation treatment.

REFERENCES

Entries highlighted in bright blue are key references.


SPECIFIC CONSIDERATIONS


Disorders of the Head and Neck
Antoine Eskander, Stephen Y. Kang, Michael S. Harris, Bradley A. Otto, Oliver Adunka, Randal S. Weber, and Theodoros N. Teknos

Infectious processes of the ear may be considered by their location (external, middle, or inner ear), their time course (acute or chronic), and the presence of complications. The external ear or pinna consists of a cartilaginous framework, perichondrium, and a relatively thin layer of skin. Erysipelas (St Anthony’s Fire) or impetigo are causes of external ear infection affecting the dermis or hypodermis of the auricle, typically caused by Streptococcus pyogenes or Staphylococcus aureus, respectively, that may be encountered posttraumatically or related to ear piercing. Treatment is oral antibiotic therapy targeting these organisms. History and clinical features such as presence of bullae and golden crusting distinguish erysipelas and impetigo from other benign entities causing erythema and edema of the auricle, such as relapsing polychondritis, which is typically diffuse, lobule-sparing, and steroid-responsive.

**Acute otitis externa**, often referred to as “swimmer’s ear,” denotes infection of the skin of the external auditory canal. Typically, the pathology is incited by moisture within the canal leading to skin maceration and pruritus. Subsequent trauma to the canal skin by scratching (i.e., instrumentation with a cotton swab or fingernail), erodes the normally protective skin/cerumen barrier. Hearing aid use and comorbid dermatologic conditions such as eczema or other forms of dermatitis may similarly serve as predisposing factors. The milieu of the external ear canal—dark, warm, humid—is ideal for rapid microbial proliferation. The most common offending organism is Pseudomonas aeruginosa, although other bacteria and fungi may also be involved. Symptoms and signs of otitis externa include itching during the initial phases and pain with marked swelling of the canal soft tissues as the infection progresses. Treatment involves removal of debris under otomicroscopy and application of appropriate ototopical antimicrobials, such as neomycin/polymyxin or quinolone-containing eardrops. The topical steroid component of these drops (e.g., hydrocortisone or dexamethasone) addresses swelling and, as a result, decreases the intense pain associated with this infection. In cases of marked ear canal edema, the use of an otowick is required to facilitate delivery of ototopical medication medially into the ear canal. Fungal infections may call for the addition of 2% acetic acid to reestablish the premorbid pH balance. Patients with otitis externa should also be instructed to keep the ear dry. Systemic antibiotics are reserved for those with severe infections, diabetics, and immunosuppression.
Malignant otitis externa, a fulminant necrotizing infection of the soft tissues of the external ear canal combined with osteomyelitis of the temporal bone, is a potentially life-threatening form of otitis externa seen most commonly among elderly patients with insulin-dependent diabetes mellitus or immunodeficiency. The classic physical finding is granulation tissue along the floor of the external auditory canal near the bony cartilaginous junction. Symptoms include persistent otalgia for longer than one month and purulent otorrhea. Biopsy is called for in order to exclude malignancy. Computed tomography (CT) and magnetic resonance imaging (MRI) define the extension of disease. Technetium 99-m scans are useful in gauging extend of bony involvement in early disease. Gallium-67 scans are valuable for monitoring disease during the course of treatment and for determining duration of antibiotic therapy. These patients require aggressive medical therapy including ototopical and IV antibiotics targeting Pseudomonas. Other gram-negative bacteria and fungi are occasionally implicated, necessitating culture-directed therapy. Patients who do not respond to medical management require surgical debridement. This condition may progress to involvement of the adjacent skull base and soft tissues, meningitis, brain abscess, and death.

Acute otitis media (AOM) typically implies a bacterial infection of the middle ear. This diagnosis accounts for 25% of pediatric antibiotic prescriptions and is the most common bacterial infection of childhood. Most cases occur before 2 years of age and are secondary to immaturity of the Eustachian tube. Well-recognized contributing factors include upper respiratory viral infection and daycare attendance, as well as craniofacial conditions affecting Eustachian tube function, such as cleft palate.

It is important to distinguish between acute otitis media and otitis media with effusion (OME). The later denotes uninfected serous fluid accumulation within the middle ear space. In children not already considered “at risk” for developmental difficulties, OME is generally observed for resolution for a period of 3 months. Age-appropriate hearing testing should be performed when OME persists for ≥3 months or at any time when language delay, learning problems, or a significant hearing loss is suspected. In the absence of these factors, the child with OME should be reexamined at 3- to 6-month intervals until the effusion is no longer present or until significant hearing loss is identified or structural abnormalities of the eardrum or middle ear are suspected. When hearing, speech, or structural concerns exist, myringotomy with tympanostomy tube placement is indicated.

Signs and symptoms of infectious otitis media occurring for <3 weeks denote AOM. In this phase, otalgia and fever are the most common symptoms and physical exam reveals a bulging, opaque tympanic membrane (Fig. 18-1). If the process lasts 3 to 8 weeks, it is deemed subacute. Chronic otitis media, lasting more than 8 weeks, usually results from an unresolved acute otitis media. The most common organisms responsible are Streptococcus pneumoniae, Haemophilus influenzae, and Moraxella catarrhalis.

In order to minimize antibiotic resistance and obviate complications of antimicrobial therapy such as allergic reaction and diarrhea, guidelines have been established for the treatment of AOM. Pain associated with AOM should be recognized and treated with oral analgesics. In children older than 6 months who are not otherwise considered “high risk” for complications (e.g., immunocompromised, previous cochlear implantation, developmental anomalies of the inner ear) with symptoms consistent with unilateral AOM without otorrhea, an initial period of observation is offered. If initial observation is selected by the physician and family, a mechanism for reexamination in 48 to 72 hours to evaluate for clinical improvement must be in place. When these criteria are not met, or clinical improvement is not observed within 48 to 72 hours, oral antibiotics are begun. First-line therapy is high-dose amoxicillin or amoxicillin-clavulanate, for β-lactamase coverage. Chronic otitis media is frequently...
treated with myringotomy and tube placement (Fig. 18-2). This treatment is indicated for frequent acute episodes and in the setting of COME as discussed previously. The purpose of this procedure is to remove the effusion and provide a route for middle ear ventilation. Episodes of AOM following tube placement are still possible. Myringotomy tubes, however, allow for prevention of painful tympanic membrane distension, risk of perforation and other complications, and permit delivery of ototopicals into the middle ear space, in most cases obviating the need for systemic antibiotic therapy.

Spontaneous tympanic membrane perforation during acute otitis media provides for drainage of purulent fluid and middle ear ventilation and frequently results in immediate resolution of severe pain. In the majority of cases, these perforations will heal spontaneously after the infection has resolved. Chronic otitis media, however, may be associated with nonhealing tympanic membrane perforations. Patients may have persistent otorrhea, which is treated with topical drops. Preparations containing aminoglycoside are avoided because this class of drugs is toxic to the inner ear. Solutions containing alcohol or acetic acid may be irritating or caustic to the middle ear and are also avoided in the setting of a perforation. Nonhealing perforation requires surgical closure (tympanoplasty) after medical treatment of any residual acute infection.

Chronic inflammatory changes from otitis media intersect with and share common etiological factors with cholesteatoma. Cholesteatoma is an epidermoid cyst of the middle ear and/or mastoid cavity that develops as result of Eustachian tube dysfunction. While several theories exist regarding causes of cholesteatoma, most cholesteatoma arises from squamous epithelium drawn into the middle ear via retraction pockets, most commonly in the pars flaccida. Squamous epithelium may also migrate into the middle ear via a perforation. Chronic mastoiditis that fails medical management or is associated with cholesteatoma is treated by mastoidectomy. Chronic inflammation and destruction of middle ear structures by osteolytic enzymes of cholesteatoma matrix may also be associated with erosion of the ossicular chain, which can be reconstructed with various prostheses or autologous ossicular replacement techniques.

Complications of otitis media with or without cholesteatoma may be grouped into two categories: intratemporal (otologic) and intracranial. Fortunately, complications are rare in the antibiotic era, but mounting antibiotic resistance necessitates an increased awareness of these conditions. Intratemporal complications include acute coalescent mastoiditis, petrositis, facial nerve paralysis, and labyrinthitis. In acute coalescing mastoiditis, destruction of the bony lamellae by an acute purulent process results in severe pain, fever, and fluctuance behind the ear. The mastoid air cells coalesce into one common space filled with pus. Mastoid infection may also spread to the petrous apex, causing retro-orbital pain and sixth-nerve palsy. These diagnoses are confirmed by computed tomographic scan. Facial nerve paralysis may also occur secondary to an acute inflammatory process in the middle ear or mastoid.

Intratemporal complications of otitis media are managed by myringotomy tube placement in addition to appropriate IV antibiotics. In acute coalescent mastoiditis and petrositis, mastoidectomy is also performed as necessary to drain purulent foci. Labyrinthitis refers to inflammation of the inner ear. Most cases are idiopathic or are secondary to viral infections of the endolympathic space. The patient experiences vertigo together with sensorineural hearing loss, and symptoms may smolder over several weeks. Labyrinthitis associated with middle ear infection may be serous or suppurative. In the former case, bacterial products and/or inflammatory mediators transudate into the inner ear via the round window membrane, establishing an inflammatory process therein. Total recovery is eventually possible after the middle ear is adequately treated.

Suppurative labyrinthitis, however, is a much more toxic condition in which the acute purulent bacterial infection extends into the inner ear and causes marked destruction of the sensory hair cells and neurons of the eighth-nerve ganglion. This condition may be a harbinger for meningitis and must be treated rapidly. The goal of management of inner ear infection, which occurs secondary to middle ear infection, is to “sterilize” the middle ear space with antibiotics and the placement of a myringotomy tube.

The most common intracranial complication of otitis media is meningitis. Otologic meningitis in children is most commonly associated with an H. influenzae type B infection. Other intracranial complications include epidural abscess, subdural abscess, brain abscess, otitic hydrocephalus, and sigmoid sinus thrombophlebitis. In these cases, the otogenic source must be urgently treated with antibiotics and myringotomy tube placement. Mastoidectomy and neurosurgical consultation may be necessary.

**Facial Nerve Disorders.** Bell’s palsy is the most common etiology of facial nerve weakness/paralysis and is clinically distinct from that occurring as a complication of otitis media in that the otologic exam is normal. Bell’s palsy is rapid, unilateral and, historically, considered idiopathic. It is now accepted, however, that the majority of these cases represent a viral neuropathy caused by herpes simplex. It is critical that clinicians distinguish Bell’s palsy from other causes of facial weakness/palsy. Alternative diagnoses are suggested by weakness/paralysis that arise gradually (rather than <72 hours), is bilateral, is accompanied by other neurological deficits, or does not show some recovery within 2 to 3 weeks and complete recovery at 3 to 4 months. Treatment includes oral steroids plus antiviral therapy (i.e., valacyclovir). Complete recovery is the norm, but it does not occur universally, and selected cases may benefit from surgical decompression of the nerve via exposure in the mastoid and middle cranial fossa.
Varicella zoster virus may also cause facial nerve paralysis when the virus reactivates from dormancy in the nerve. This condition, known as Ramsay Hunt syndrome, is characterized by severe otalgia followed by the eruption of vesicles of the external ear and the soft palate. Treatment is similar to Bell’s palsy, but full recovery is only seen in approximately two-thirds of cases.

Traumatic facial nerve injuries may occur secondary to accidental trauma or surgical injury. Iatrogenic facial nerve trauma most often occurs during mastoidectomy, most commonly to the vertical segment of the nerve. Detailed knowledge of facial nerve anatomy and adjunctive use of nerve integrity monitoring systems are imperative in this context. When the facial nerve is injured during an operative procedure, it is explored. Injury to >50% of the neural diameter of the facial nerve is addressed either with primary reanastomosis or reconstructed with the use a nerve graft. Complete recovery of nerve function is uncommon in these cases.

Lesions of the Internal Auditory Canal and Cerebellopontine Angle. The most common lesion affecting the internal auditory canal (IAC) and the cerebellopontine angle (CPA) is vestibular schwannoma (formerly referred to as “acoustic neuroma”). Less commonly encountered lesions of the IAC and CPA include meningioma and epidermoid tumors. Vestibular schwannomas are benign tumors that comprise 60% to 92% of all CPA lesions and 6% to 10% of intracranial tumors. They demonstrate an average growth rate of 1 to 2 mm per year. Vestibular schwannomas are most commonly unilateral and sporadic; bilateral tumors are the hallmark of neurofibromatosis type 2 (NF2), an autosomal dominant condition linked to mutation of a tumor suppressor gene mapped to chromosome 22. The most common presenting symptoms of vestibular schwannoma are asymmetric sensorineural hearing loss and speech perception deficits often out of proportion to degree of hearing loss indicated by audiometry. Unilateral tinnitus is also frequently reported. Disequilibrium or, less commonly, episodic vertigo may be present. Facial nerve weakness or paralysis is rare. Larger tumors may feature facial numbness and loss of the cornea reflex from compression of the trigeminal nerve. Very large lesions can lead to brainstem compression, obstructive hydrocephalus, and death.

Gadolinium-enhancement on T1-weighted MRI is the gold standard for diagnosis and detects even very small tumors (Fig. 18-3). The conventional armamentarium for vestibular...

Figure 18-3. A. Axial T1 magnetic resonance imaging (MRI) post-contrast showing left cerebellopontine angle tumor with avid gadolinium enhancement. Minimal internal auditory canal involvement is noted. B. Axial T2 MRI showing left cerebellopontine angle tumor with thin cerebrospinal fluid cleft between tumor and brainstem/cerebellum. C. Axial T1 MRI post-contrast showing left cerebellopontine angle tumor with avid gadolinium enhancement. The lesion is confined to the internal auditory canal with minimal cerebellopontine angle involvement. D. Intraoperative phono during microsurgical resection via translabyrinthine approach. Black arrow indicates cochlear nerve.
schwannoma includes observation, microsurgical resection, and stereotactic radiation. Management of patients with vestibular schwannomas involves weighing a multitude of variables particular to the tumor (location, size, growth pattern), the patient (age, overall health, individual wishes), and the interaction between tumor and patient (symptoms currently experienced, symptoms likely to develop with lesion progression, degree of residual hearing). For patients who have hearing that may still benefit from acoustic amplification using a hearing aid, either a retrosigmoid or a middle fossa approach may be offered, depending on tumor location, size, patient preference, and provider experience. For patients without serviceable hearing preoperatively, a translabyrinthine approach is most commonly offered.

Sinonasal Inflammatory Disease

Rhinosisinusitis. Rhinosinusitis is defined as symptomatic inflammation of the nasal cavity and paranasal sinuses. Rhinosinusitis is preferred over sinusitis because sinusitis almost always is accompanied by inflammation of the contiguous nasal mucosa. Rhinosinusitis is a significant health burden, affecting nearly 12% of the population. Rhinosinusitis is the fifth most common diagnosis responsible for antibiotic prescription and accounts for more than 20% of all antibiotics prescribed to adults. Rhinosinusitis may be broadly classified based on duration of symptomatology. Symptoms lasting <4 weeks may be classified as acute rhinosinusitis (ARS), while symptoms lasting >12 weeks may be classified as chronic rhinosinusitis (CRS). Rhinosinusitis lasting between 4 and 12 weeks has historically been defined as “subacute,” although the current clinical practice guideline published by the American Academy of Otolaryngology—Head and Neck Surgery does not distinguish rhinosinusitis in this time frame, noting that this group likely represents crossover symptoms from one of the other two subclasses. Hence, the decision on how to manage this group of patients must be individualized.

Acute Rhinosinusitis. Acute rhinosinusitis most commonly occurs in the setting of a viral upper respiratory tract infection (URI). Although it is believed that acute bacterial rhinosinusitis (ABRS) typically follows a viral URI, it has been estimated that only up to 2% of viral URIs lead to ABRS. The most common viruses involved in ARS include rhinovirus, influenza virus, and parainfluenza virus. It is not known whether the viral URI preceedes or only occurs along with ABRS. Regardless, viral infection leads to mucosal edema with sinus ostium obstruction, mucus stasis, tissue hypoxia, ciliary dysfunction, and epithelial damage, which may enhance bacterial adherence. Other conditions that may contribute to ABRS should be investigated, especially in the setting of recurrent ABRS. Such conditions include foreign body, sinus fungal ball (with bacterial secondary infection), and periapical dental disease (Figs. 18-4 and 18-5).

The symptomatic criteria used to define ABRS include up to 4 weeks of purulent nasal drainage accompanied by nasal obstruction, facial pain with pressure and fullness, or both.
Other historical factors that may predict the development of ABRS include persistence of symptoms beyond 10 days, or worsening of symptoms, following initial improvement, within 10 days (“double worsening”). Although routine head and neck examination may identify anteriorly or posteriorly draining purulent secretions, the utilization of a rigid endoscope may improve diagnostic sensitivity and may also facilitate culture acquisition (Fig. 18-6).

The management of ABRS is heavily dependent on antibiotics, either culture-directed or empirically chosen to cover the most common isolates of ABRS, including S. pneumoniae, H. influenzae, and M. catarrhalis. Nosocomial ABRS more commonly involves P. aeruginosa or S. aureus. Methicillin-resistant S. aureus (MRSA) has been isolated with increasing frequency. Other treatments include topical and systemic decongestants, nasal saline spray, topical nasal steroids, and oral steroids in selected cases. In the acute setting, surgery is reserved for complications or pending complications, which may include extension to the eye (orbital cellulitis or abscess) or the intracranial space (meningitis or intracranial abscess).

Chronic Rhinosinusitis. Chronic rhinosinusitis (CRS) is characterized by symptomatic inflammation of the nose and paranasal sinuses lasting over 12 weeks. CRS has been clinically classified into two main groups: those with CRS with nasal polyps (CRSsNP) tend to exhibit a Th2-biased inflammatory profile, and those with CRS without nasal polyps (CRSsNP) tend to exhibit a Th1-biased profile. Although the etiology of CRS is unclear and the development of the clinical subtypes may be distinct, there exists significant overlap not only in physiologic manifestations but also in symptomatology. Hence, the sinonasal cavities of patients with both subtypes of CRS tend to exhibit mucosal edema, ostial obstruction, ciliary dysfunction, and an abhorrent inflammatory milieu.

Two of the following symptomatic criteria must be present to diagnose CRS: purulent nasal drainage, nasal obstruction, facial pain-pressure-fullness, and decreased sense of smell. These patients may also experience acute exacerbation, generally signified by an escalation of symptoms. Frequently, this is due to bacterial infection. However, patients with acute exacerbation of CRS may be distinguished from patients with recurrent acute bacterial rhinosinusitis (four or more episodes of ABRS per year) through baseline comparison: patients with CRS are symptomatic, even while at baseline, while patients with recurrent acute bacterial sinusitis are normal at baseline. As with ARS, the diagnosis of CRS requires objective confirmation utilizing either nasal endoscopy, CT scans, or, less commonly, MRI.

Nasal endoscopy is a critical element of the diagnosis of CRS. Abnormalities that may confirm the diagnosis of CRS include

- Purulent mucus in the middle meatus or anterior ethmoid region
- Edema in the middle meatus or ethmoid region
- Polyps in nasal cavity or the middle meatus

In addition to establishing the diagnosis, nasal endoscopy can be valuable in antibiotic selection by facilitating specific culture acquisition. Furthermore, simple polypectomy or steroid injection can be performed under topical anesthesia in the appropriate clinical setting.

Imaging is also an important clinical tool in the diagnosis of CRS. In general, CT is the modality of choice for diagnosis and management of CRS. Usual diagnostic criteria include mucosal thickening, sinus opacification, and bony remodeling (erosion or hyperostosis). It should be underscored, however, that CT scan is not the positive gold standard because many asymptomatic patients will demonstrate findings on a sinus CT scan, and many patients with presumed sinusitis will have negative findings. CT scan has excellent negative predictive value when performed in the setting of active symptoms. Thus, if a patient complains of rhinosinusitis-like symptoms but has no specific physical (endoscopic) findings, and the scan
is negative, other diagnoses (e.g., allergic rhinitis, migraine headache, tension headaches, and laryngopharyngeal reflux) should be sought. This has led to the utility of point-of-care CT (POC-CT) scan that can be performed in the physician’s office. POC-CT utilizes cone beam technology, which acquires the equivalent of >100 axial slices in approximately 1 minute at an effective resolution of 0.3 mm or less. The equipment occupies a room of 8’ × 10’ and can thus be accommodated in almost any office setting (Fig. 18-7). Perhaps most important, the radiation dosing for even the most sophisticated protocol is 0.17 mSv, which is <10% the dose of a conventional head CT and equivalent to approximately 20 days of background radiation. One theoretical shortcoming of this technology is that it does not permit soft tissue imaging. This is seldom a concern in sinonasal evaluation, as this is typically undertaken in bone windows. The acquired data are immediately formatted into triplanar (axial, sagittal, coronal) reconstructions and is also compatible with devices used for intraoperative stereotactic navigation, which can be used to confirm relationships between the disease process, medial orbital wall, and skull base during surgery (Figs. 18-8 and 18-9).

Medical management of CRS is heavily dependent on topical intranasal therapy. The reasons for this lie not only in established effectiveness but also in tolerability and safety—the chronic nature of CRS generally lends to requisite long-term medication administration despite other measures such as surgery. Nasal irrigation and topical nasal steroids are commonplace in the management of CRSwNP and CRSsNP. Oral steroids have demonstrated effectiveness in patients with CRSwNP, although the role in CRSsNP is less clear. Although otolaryngologists commonly utilize antibiotics in the management of CRS, indications and administration practices are not uniform. Oral antibiotic therapy given for short duration (<4 weeks) is generally useful in the management of acute exacerbation related to bacterial infection. Long-term utilization of antibiotics may be necessary in the setting of chronic infection or osteomyelitis. Additionally, long-term macrolide administration may be utilized for anti-inflammatory effects in the appropriate clinical setting.

In most cases, patients considering endoscopic sinus surgery (ESS) for CRS should have significant residual
symptomatology despite medical therapy. However, there currently exists no consensus regarding what constitutes a “maximum” course of medical therapy. It should be noted that unless there is suspicion of neoplasm or pending complication of rhinosinusitis, the decision to proceed with surgery is highly individualized. This is because surgery for uncomplicated CRS is elective, and patients who “fail” medical management will exhibit significant variability in symptoms, physical signs, and CT findings. Furthermore, ESS is not necessarily curative—the intent of ESS is to remove the symptoms related to CRS rather than cure the underlying condition itself.

Surgery is typically preformed endoscopically where the goals are to remove polyps, enlarge or remove obstructing tissue surrounding the natural sinus ostia (Fig. 18-10), and remove chronically infected bone and mucosa to promote both ventilation and drainage of the sinus cavities. Inspissated mucin or pus is drained and cultured. Eventual resolution of the chronic inflammatory process can be attained with a combination of meticulous surgery and directed medical therapy, although the patient must understand that surgery may not alter the underlying immunologic pathophysiology. In cases where resection of inflammatory tissue and polyps are not required, recent trends have also included use of angioplasty-type balloons to dilate sinus ostia. The exact role for this technology is unclear, but it appears to have promise in outpatient office management of patients with focal or limited obstructive pathology.

**Endoscopic Skull Base Surgery.** Over the past three decades, the development and expansion of multidisciplinary skull base teams has become somewhat commonplace at large academic institutions. Facilitated mainly by growing cooperation between otolaryngologists and neurosurgeons, a variety of approaches that utilize the sinonasal corridor to treat a plethora of pathologic processes of the anterior skull base have been developed.

Technological advances in endoscopy, instrumentation, and imaging have also facilitated the development of endoscopic endonasal approaches (EEAs), allowing team members to work simultaneously while maintaining optimal visualization of the relevant anatomy and freedom of movement within the corridor. Although historically the sphenoid sinus has been the common access route in the management of sellar pathology, a series of modular approaches of varied complexity have been developed that have broadened the reach of EEAs to address lesions at virtually all comportments of the ventral skull base, from the crista galli to the anterior arch of C2.22

One of the key tenets of the EEA is that the sinonasal corridor presents the most prudent and safest path to the lesion of interest. Accordingly, the EEA is generally chosen for lesions adjacent to the skull base, without intervening brain parenchyma, cranial nerves, major vessels, or other important anatomical structures. Currently, EEAs are utilized to treat a significant number of pathologic process involving the skull base, including: cerebrospinal fluid leaks, encephaloceles, meningoceles, pseudomeningoceles, benign intracranial tumors (Fig. 18-11), benign sinonasal tumors, malignant sinonasal tumors, and inflammatory or traumatic conditions leading to compression at the craniovertebral junction. Although EEAs tend to be considered “minimally invasive,” the corridor created in the sinonasal cavity is nonetheless comprehensive enough to
Figure 18-10. A. Endoscopic view of the right nasal cavity demonstrating the uncinate process (U), ethmoid bulla (EB), middle turbinate (MT), inferior turbinate (IT), and nasal septum (S). B. Endoscopic view of a microdebrider being used to widen the right maxillary sinus ostium.

Figure 18-11. Preoperative coronal (A) and sagittal (B) magnetic resonance images of a large olfactory groove meningioma removed using endoscopic endonasal approach. Postoperative coronal (C) and sagittal (D) images demonstrating removal of the tumor. The skull base can be reconstructed using local flaps (most commonly a nasoseptal flap pedicled on the posterior nasal artery).
provide maximal freedom of movement for the critical component of the case (i.e., tumor resection near vital structures). Once the corridor is created by the otolaryngologist, the neurosurgeon joins, and a two-person, three- to four-hand technique is utilized to address the lesion of interest and reconstruct the skull base (Fig. 18-12).

Despite the relatively confined aperture provided by the nostrils, even large tumors can be removed using EEAs, albeit via piecemeal removal. For malignant tumors, this has required a philosophical shift whereby en bloc resection of the entire tumor is replaced by piecemeal removal of the bulk of the tumor followed by complete resection of the pedicle with sufficient margins. Outcomes utilizing EEAs for resection of malignant tumors, when chosen appropriately, parallel those of traditional open approaches. However, EEAs are not favored over traditional approaches when oncological principles would otherwise need to be violated.

**Pharyngeal and Adenotonsillar Disease**

Waldeyer’s ring consists of the palatine tonsils between the anterior and posterior tonsillar pillars, the lingual tonsils (lymphoid tissue in the base of tongue), and the adenoid located in the nasopharynx. These four main sites of Waldeyer’s ring are connected by other minor lymphoid tissue along the posterior and lateral pharyngeal wall completing the ring. These are all considered mucosa-associated lymphoid tissue (MALT). These tissues react to inflammatory disease, infection, trauma, acid reflux, and radiotherapy. Even the vibratory effects of chronic snoring have been implicated in the development of adenotonsillar disease. Inflammation of these tissues can lead to referred pain through cranial nerves IX and X to the throat and ear. Adenotonsillar tissue does not have any afferent lymphatics and receives antigen presentation directly, with appropriate production of memory cells. However, there is no clear immune compromise after removal.

**Microbiology and Complications.** Adenotonsillar infections present with three temporal patterns: acute, recurrent acute, and chronic. Acute infection is typically viral in origin but secondary bacterial invasion may initiate chronic disease. Viruses do not cause chronic infections; however, Epstein-Barr Virus (EBV) can cause significant hypertrophy. Systemic EBV infection, also known as mononucleosis, can mimic bacterial pharyngitis, but the progression of signs and symptoms demonstrates lymphadenopathy, splenomegaly, and hepatitis. This can be diagnosed on bloodwork (heterophile antibody or atypical lymphocytes). The most common bacterial causes of acute tonsillitis are group A β-hemolytic streptococcus species (GABHS) and *S. pneumoniae.* If GABHS is confirmed, then antibiotic therapy is warranted in the pediatric population to decrease the risk (3%) of developing rheumatic fever. A positive test for GABHS historically meant a throat swab with culture and sensitivity; however, rapid antigen assays have been demonstrated to be reasonably sensitive and specific (85% and 95%, respectively), thus largely replacing cultures. If the rapid assay is negative, then a culture is warranted. The remainder of the bacteriology for adenotonsillar disease is similar to otitis media and sinusitis, which includes *H. influenzae* and *M. catarrhalis.* Atypical infections include *Corynebacterium diphtheria,* *Neisseria gonorrhoeae,* and *Chlamydia trachomatis.*

Complications of GABHS pharyngitis, typically from *S. pyogenes,* can be systematic and include poststreptococcal glomerulonephritis, scarlet fever, and rheumatic fever. Antibiotic therapy does not decrease the incidence of glomerulonephritis. Scarlet fever, caused by blood-borne streptococcal toxins, causes a strawberry tongue and a punctate rash on the trunk that spreads distally while sparing the palms and soles. Peritonsillar abscess is also a common complication that is treated in an ambulatory setting through a transoral approach after appropriate topicalization and local anesthetic. Deep neck space infections are rare from pharyngitis but can occur from odontogenic and salivary gland infections. These typically require a transcervical approach for incision and drainage.

**Adenoids and Adenoidectomy.** Acute adenoiditis typically presents with purulent rhinorrhea, nasal obstruction, and fever and can be associated with otitis media, particularly in the pediatric population. Recurrent acute adenoiditis is defined as four or more acute infections in a 6-month period, but in an adult, this may be difficult to distinguish from recurrent acute sinusitis, and endoscopy with or without imaging of the sinuses may be warranted to distinguish between the two diagnoses. Chronic adenoiditis presents with persistent nasal discharge, halitosis, chronic congestion, and postnasal drip. In children, obstructive adenoid hyperplasia often requires surgical intervention to help relieve obstructive symptoms such as snoring, obligate mouth breathing, and hyponasal voice.

The management of adenoid disease is slightly different than that for tonsillar disease. Chronic infection can be treated with antibiotics, although this often does not lead to a full resolution of symptoms. If the adenoid bed appears hyperplastic on lateral X-ray imaging or endoscopy, a 2-month trial of nasal steroids may be helpful. Adenoidectomy is indicated for recurrent and chronic infections that have failed conservative management. These infections are not limited to the adenoid bed but also involve the sinuses and the middle ear. Adenoidectomy with a myringotomy and ventilation tube placement is beneficial for recurrent or chronic otitis media in children because the

**Figure 18-12.** Two-surgeon, three- to four-hand technique utilized in endoscopic endonasal surgery.
Adenoid functions as a reservoir for bacteria that can enter the middle ear through the Eustachian tube.\textsuperscript{25}

Adenoidectomy is also the first line of surgical management for children with chronic sinusitis because the adenoid can obstruct mucociliary clearance from the sinonasal tract into the choana and ultimately into the pharynx. Patients with obstructive systems attributable to the adenoids and suspected benign or malignant neoplasms of the adenoid bed are also candidates. However, the procedure is contraindicated in patients with velopalatine insufficiency (VPI) and in patients with a cleft palate. Prior to adenoidectomy, patients should be examined for a submucous cleft, a lack of midline muscular tissue of the soft palate. Clinical signs of this include a bifid uvula, a translucent portion of the muscular diastasis of the soft palate (zona pellicuda), and a palpable notched hard palate.\textsuperscript{26} A number of different methods can be used to perform an adenoidectomy: cold steel, suction coagulator, microdebrider, and coblation. Adenoid regrowth and bleeding rates are both low, and no study has been able to demonstrate the superiority of one technique over the other for either outcome.\textsuperscript{27,28} Adenoidectomy is not without complications though, beyond VPI and bleeding, halitosis and adenoid bed regrowth (\textminus1\%) are common complications. Rare complications include torticollis secondary to inflammation of the prevertebral fascia, nasopharyngeal stenosis, and cervical spine subluxation, which is more common in patients with Down syndrome.

**Tonsils and Tonsillectomy** Patients with acute tonsillitis present with sore throat, fever, dysphagia, and tender cervical nodes with erythematous or exudative tonsils. The Centor Criteria is used to identify the likelihood of bacterial infection in adult patients complaining of sore throat in the emergency department or walk-in clinic, a point is given for each of the following: fever, tonsillar exudate, lymphadenopathy, and lack of cough.\textsuperscript{29-31} A score of 0 to 1 warrants no treatment, a score of 2 to 3 warrants GABHS testing, and a score of 4 warrants initiation of antibiotic therapy. First-line treatment is with penicillin or a cephalosporin; however, in those with an allergy, a macrolide can be considered. Documentation of recurrent acute infections should include a temperature ($>38.3^\circ$C), cervical adenopathy, tonsillar exudate, and a positive test for GABHS.

According to the American Academy of Otolaryngology—Head and Neck Surgery (AAO-HNS) clinical practice guideline on tonsillectomy in children, tonsillectomy is indicated when children have more than 7 documented episodes per year, 5 episodes per year in the past 2 years, or 3 episodes per year in the past 3 years.\textsuperscript{32} Tonsillectomy can still be considered in children who do not meet these criteria if they have multiple antibiotic allergies or intolerances, have a history of peritonsillar abscess after the acute inflammation has resolved, or have PFAPA (periodic fever, apthous stomatitis, pharyngitis, and adenitis). A peritonsillar abscess is an infection of the peritonsillar salivary gland (Weber’s gland), located between the tonsil capsule and the muscles of the tonsillar fossa. In selected cases of active peritonsillar abscess, tonsillectomy is required in the acute setting to treat systemic toxicity or impending airway compromise. Multiple techniques have been described, including electrocautery, sharp dissection, laser, and radiofrequency ablation. There is no consensus as to the best method.

**Sleep Disordered Breathing and Adenotonsillar Disease.** Patients with sleep-disordered breathing (SDB) and tonsillar hypertrophy may also benefit from tonsillectomy if they have growth retardation, poor school performance, enuresis, or behavioral problems. The benefits may be accentuated in children with abnormal polysomnography; however, DB may require further treatment after tonsillectomy when it is multifactorial. Clinical documentation of tonsillar grade/size is based on the percentage of the transverse oropharyngeal space measured between the anterior tonsillar pillars: grade 1 $<25\%$; grade 2+ 25\% to 49\%; grade 3+ 50\% to 74\%; grade 4+ $\geq75\%$ or more sometimes referred to as “kissing tonsils.”\textsuperscript{33} Tonsillectomy is effective for control of SDB in 60\% to 70\% of patients with tonsillar hypertrophy, although this much lower (10\%–25\%) in obese children, and it is therefore not curative in obese children but may improve some of their symptoms nonetheless. In patients with Down syndrome, obesity, craniofacial abnormalities, neuromuscular disorders, sickle cell disease, or mucopolysaccharidoses, polysomnography (PSG) should be performed prior to tonsillectomy.\textsuperscript{33} When the need for surgery is uncertain or when there is a discordance between tonsillar size on physical examination and the reported severity of SDB, physicians should advocate for PSG prior to tonsillectomy. Tonsillectomy, usually with adenoídectomy if the adenoids are enlarged, is often performed on an outpatient basis unless the patient has documented or strongly suspected obstructive sleep apnea (OSA), is $<3$ years of age, or has severe OSA (in children, an apnea-hypopnea index $\geq10$ or more, oxygen saturation $<80\%$, or both). Other reasons for admission include a home $>1$ hour from a hospital, patients with craniofacial abnormalities, or any other medical issue. There is strong evidence to suggest the routine administration of a single intraoperative dose of IV dexamethasone in children undergoing tonsillectomy, though antibiotics should not be administered or prescribed perioperatively in children. The complications from tonsillectomy include perioperative bleeding (3\%–5\%), airway obstruction, death, and readmission from postoperative dysphagia leading to dehydration.\textsuperscript{34} It is recommended that surgeons calculate and quote their own primary and secondary posttonsillectomy hemorrhage rates yearly.\textsuperscript{35} A rare but serious complication in patients with obstructive adenotonsillar disease post adenotonsillectomy is postobstructive pulmonary edema syndrome, which presents with decreased oxygen saturation and frothy, blood-tinged oral secretions. Patients usually recover with reintubation, positive pressure, diuresis, and supportive care.

**Multilevel Sleep Surgery.** SDB surgery is often multilevel and is not limited to adenotonsillar disease. Patients with nasal obstruction may benefit from septoplasty and tracheotomy reduction, although in the adult population this is most commonly used to allow patients to tolerate their OSA appliances. Similarly, patients with significant lingual tonsillar hypertrophy and a large base of tongue may benefit from a base of tongue reduction, tongue base advancement, or geniohyoidopexy. A base of tongue reduction alone does not often provide enough apnea-hypopnea index reduction (30\%–60\%) for resolution of symptoms and is fraught with a high morbidity rate.\textsuperscript{35} Rarely, maxillomandibular advance is required to open up the retrolingual space. In patients with life threatening symptoms (right heart failure/cor pulmonale, oxygen saturation $<70\%$, comorbid cardiopulmonary disease) who have failed other measures, the only “cure” for OSA is a tracheotomy.

**Other Tonsillar Pathology.** Unilateral tonsillar hypertrophy is mostly benign but can also be the result of *Mycobacterium tuberculosis*, atypical mycobacterium, fungi, or Actinomyces. With the epidemic rise in incidence of oropharyngeal
cancers, neoplasms (squamous cell carcinoma and lymphoma) have increasingly also presented as tonsillar asymmetry.\(^{36}\) Management of these lesions is dependent on the pretest probability of malignancy and the type of malignancy. If squamous cell carcinoma is suspected, then a biopsy alone is sufficient so as to not impact the possibility of other future surgical interventions such as transoral robotic surgery. If lymphoma or a nonmalignant pathology is suspected, tonsillectomy is often recommended for diagnostic and therapeutic reasons, and the specimen should be sent fresh to pathology for a lymphoma protocol workup, bacterial and fungal culture, and gram stain. Pharyngitis may also be seen in immune-mediated conditions such as erythema multiforme, bullous pemphigoid, and pemphigus vulgaris.

### Benign Conditions of the Larynx

Hoarseness is the most common presenting symptom for patients with a voice complaint. Other complaints include breathiness, weakness/hypophonia, aphonia, and pitch breaks. Voice disorders affect a large range of patient ages, occupations, and socioeconomic statuses and affect both genders equally. They can be associated with dysphagia, globus sensation, laryngopharyngeal reflux (LPR) disease and, rarely, airway obstruction.\(^{37}\) Smoking can both cause and aggravate preexisting benign laryngeal conditions and raises the suspicion of malignancy often requiring a biopsy to exclude this diagnosis.

Any discussion of laryngeal disorders should start with a review of the anatomy of the vocal cords (Fig. 18-13). The true vocal cords are formed from stratified squamous epithelium, beneath which is the superficial lamina propria (in Reinke’s space). Beneath this is the ligament that includes the middle and deep lamina propria. Beneath this ligament is the muscular layer that includes the thyroarytenoid muscle or vocalis. The cover-body theory describes the freely mobile cover (mucosa and Reinke’s space) over the more rigid body (vocalis and vocalis).\(^{38}\)

Membranous vocal cord lesions have been notoriously difficult to classify reliably; however, increased availability of videoendoscopic examination and standardized definitions have improved the classification of these lesions.\(^{39}\) These lesions are usually mid cord because that is the site of maximal lateral displacement and amplitude. Vocal fold nodules are typically bilateral, fairly symmetric, and with normal or mild impairment of the mucosal wave, and they almost always resolve with voice therapy. A vocal fold polyp is more often unilateral than bilateral, is exophytic, and is associated with unorganized gelatinous debris in the subepithelial space. These can be hemorrhagic as is often seen in males secondary to capillary rupture within the mucosa by shearing forces during voice abuse. Hemorrhagic polyps are seen more often in patients on anticoagulants. These lesions usually fail conservative measures (voice rest, voice therapy, smoking cessation, and reflux management) usually requiring microlaryngeal surgery to remove the lesion while preserving normal mucosa. Vocal fold cyst is an encapsulated lesion within the subepithelial or ligamentous space and is associated with reduced mucosal wave. It typically does not resolve with voice therapy. These lesions require microlaryngeal surgery for complete removal of the cyst while preserving the overlying mucosa, and this surgery can be performed with cold steel or carbon dioxide (CO2) laser. A fibrous mass of the vocal fold is amorphous fibrous material within the subepithelial space or

---

**Figure 18-13.** Coronal view of the larynx demonstrate the supraglottic, glottic and subglottis (LEFT) and the layers of the true vocal cord (RIGHT).
ligament often associated with reduced mucosal wave, and it also does not resolve with voice therapy.

Reinke’s edema is characterized by edema in the superficial lamina propria of the vocal cord. Edema is thought to arise from injury to the capillaries that exist in this layer, with subsequent extravasation of fluid. The etiology is multifactorial: smoking, LPR, hypothyroidism, and vocal misuse. This pathology is more common in women (because they present early due to a deep vocal pitch change in their voice) and heavy smokers. The physical examination findings are typically bilateral. Surgery typically involves microlaryngoscopy with removal of the gelatinous debris in Reinke’s space with trimming of the excess mucosa. However, smoking cessation and surgery do not fully reverse the structural abnormalities due to the presence of possible structure alterations in fibroblasts caused by the toxicity of cigarette components, resulting in uncontrolled production of fibrous matrix in the lamina propria, thus preventing complete vocal recovery.

Laryngeal granulomas typically occur in the posterior larynx on the arytenoid mucosa (Fig. 18-14). These lesions are typically multifactorial: chronic throat clearing, phonotrauma, endotracheal intubation, compensatory supraglottic squeeze from vocal fold paralysis, and LPR. The majority of these lesions (82%) disappear within 48 weeks with conservative measures such as voice therapy, vocal rest, oral steroids, inhaled steroids, and proton pump inhibitors. Botulinum toxin of thyroarytenoid and lateral cricoarytenoid muscles can be used as first-line treatment in patients who prefer a chemically activated voice rest regimen. LPR appears to be the most important contributing factor, and when aggressive conservative and medical therapy has failed, a Nissen fundoplication may be indicated. Surgery is rarely required for patients with laryngeal granulomas because it does not address the underlying etiology and is frequently associated with recurrence. Nonetheless, excision is sometimes required in patients with airway obstruction or the suspicion of malignancy. Careful preservation of the arytenoid perichondrium intraoperatively is required to assist with reepithelialization and to decrease the risk of recurrence postoperatively.

Recurrent respiratory papillomatosis (RRP) is pathophysiologically associated with human papillomavirus (HPV) within the mucosa of the upper aerodigestive tract. The glottis and supraglottis are the two most common involved subsites. HPV 6 and 11 are the most often implicated types; however, LPR and herpes simplex virus (HSV) type-2 are risk factors of adult-onset RRP. The disorder typically presents in early childhood (juvenile-onset RR; JoRRP) secondary to HPV acquisition during vaginal delivery; however, children born by caesarean section are also at risk for the disease. JoRRP usually resolves around puberty but can progress into adulthood. Adult-onset RRP is less severe and is more likely to involve extralaryngeal subsites. There is no cure for RRP. Surgery excision is used to improve voice and airway symptoms in a palliative fashion. Surgical excision in the operating room involves microlaryngoscopy with the use of the laser (CO2 for bulky disease or KTP for more superficial disease) or the use of a microdebrider. The microdebrider has been demonstrated to have superior voice outcomes in JoRRP; however, CO2 laser is the most commonly used operative ablative technique used in adults. Recent advances have made it possible to treat a select group of adult RRP patients in the office using the KTP laser, typically for those with a lower disease burden. Several adjuvant treatments are used to increase the intersurgical interval, including intralesional cidofovir injection, oral indole-3-carbinol, oral methotrexate, and retinoic acid. In addition to preventing RRP in some patients, the HPV vaccine has also been demonstrated to increase the intersurgical interval in the most aggressive JoRRP patients.

Leukoplakia is a white patch seen on mucosa that can be wiped off on physical examination. This can be seen anywhere in the upper aerodigestive tract. In the larynx, this is typically seen on the superior surface of the true vocal cords and may represent squamous hyperplasia, dysplasia, and/or carcinoma with an associated risk of malignant transformation of 1% to 3% in hyperplastic lesions and 10% to 30% in dysplastic lesions. Lesions that are not overtly suspicious for malignancy, particularly in patients without a strong smoking or alcohol history, can be managed conservatively (increased hydration, elimination of poor vocal habits, phonotrauma, and management of LPR) for 1 month before reevaluation with fiberoptic laryngoscopy. Any lesions that progress, persist, or recur could have microlaryngoscopy with complete excision. Similarly, because erythroplasia and ulceration are more suggestive of malignancy, these lesions also require an excisional biopsy in the operating room.

The most common cause of unilateral vocal cord paresis is iatrogenic in origin, following surgery to the thyroid, parathyroid, carotid, spine through an anterior approach, or cardiothoracic structures. It is therefore very important that all patients undergoing thyroid surgery receive preoperative visualization of the larynx, usually in the form of fiberoptic nasolaryngoscopy, although an indirect mirror exam can be used if adequate visualization is possible. Postthyroidectomy visualization may also be required to document normal vocal cord movement. Less common causes include malignancy of structures near the recurrent laryngeal nerve (RLN) from the skull base jugular foramen to the mediastinum. In the pediatric population, there can be neurologic causes, the most common of which is the Arnold-Chiari malformation. Overall, the left vocal cord is more commonly involved secondary to the longer course of the RLN on that side. Other rare etiologies include trauma, intubation injury, atypical infections, and neurotoxic medications. Patients typically present with a weak breathy voice and may have aspiration secondary to diminished supraglottic sensation if the proximal vagal nerve or superior laryngeal nerve is involved. RLN injury is also associated with delayed relaxation.

Figure 18-14. Laryngeal granuloma.
of the cricopharyngeus muscle that can lead to dysphagia and decreased sensation in the hypopharynx, which can cause pooling of secretions. In children, stridor, weak cry, and airway compromise may be presenting symptoms, whereas in adults this is rarely the case unless there is bilateral vocal cord paralysis. When an obvious cause is not identified after a thorough history and physical examination including fiberoptic nasolaryngoscopy, then a more comprehensive workup is required. A workup should not include autoimmune serology as a screen because this is low yield, but this can be included if there is a suspicion of autoimmune disorders. Imaging, in the form of a CT scan, is the mainstay of the workup and should include the skull base to the mediastinum. Repeat imaging is beneficial in this population within a 2-year period because many patients have undiagnosed small malignancies as the primary cause of their paralysis that are too small to detect on initial imaging. Laryngeal electromyography can assist with identifying whether the paresis is a result of a paralysis or cricoarytenoid joint fixation/dislocation. It can also help prognosticate a paralysis. This is, however, rarely used in practice. Despite an extensive workup, 20% to 35% of cases are idiopathic.

The management of bilateral vocal cord paralysis almost always requires a tracheotomy because the cords are left in a paramedian position leaving a slit light glottic aperture. If the paralysis is permanent, then a cordectomy with or without arytenoidectomy can be used to open up the airway in an attempt to eventually decannulate the patient. However, this has obvious implications for voice with a weak and breathing voice. Many patients with a unilateral paralysis compensate when the cord is in the paramedian position using supraglottic structure and the contralateral cord on their own or with speech therapy. However, in patients with a less than adequate voice-related quality of life, four techniques have been used to surgically manage patients with a unilateral vocal cord paralysis: injection laryngoplasty, medialization thyroplasty, arytenoid addition, and laryngeal reinnervation. Injection laryngoplasty involves injecting a temporary filler medial to the vocalis into the ligament at the posterior and midmembranous vocal cord. This can be performed in the office or in the operating room, depending on the comfort of the surgeon and patient characteristics. Materials used include autologous (fat, collagen) or alloplastic (hydroxyapatite, hyaluronic acid, micronized cadaveric human collagen) compounds. Early medialization is recommended in patients with mediastinal and thoracic malignancies because it is safe and has been shown to improve quality of life in a palliative setting. Teflon is historic and is no longer used because of its granulomatous side effects on the larynx. A more permanent medialization can be performed using a medialization thyroplasty, during which a small window is created in the inferolateral aspect of the thyroid cartilage and a submucosal-carved silastic block is placed in the operating room with the patient under neurolept anesthetic so that vocalization and flexible laryngoscopic visualization of the larynx can be improved (Fig. 18-15). In some cases, this is not enough of a medialization due to a large posterior glottic chink, and an arytenoid adduction is required to provide better closure of the posterior glottis and supraglottis with ensuing improved vocal outcomes. This is a technically challenging procedure that is rarely required, but in select patients it is associated with significant improvements in voice. Lastly, laryngeal reinnervation, typically with the ansa cervicalis that supplies motor function to the strap muscles, can also be performed. This is the best approach in patients who have had a recurrent laryngeal nerve severed during a central or upper mediastinal neck procedure because it is in the field. Multiple studies demonstrate favorable outcomes; however, no significant differences between treatment arms has been demonstrated based on perceptual, acoustic, quality of life, and laryngoscopic outcomes.

Vascular Lesions

Vascular lesions can be broadly classified into two groups: hemangiomas and vascular malformations.

Hemangiomas. Hemangiomas are the most common vascular lesion present in infancy and early childhood. Infantile hemangiomas present largely within the first few weeks of life. Initially they proliferate (2 weeks to 1 year), and then they begin to involute (1–7 years) until they have fully involuted, leaving the child with redundant skin, scar, or a fatty lesion. Children with large facial infantile hemangiomas benefit from regular neurological examinations and brain MRI to rule out PHACES syndrome (Posterior fossa malformations, Hemangiomas, Arterial lesions, Cardiac abnormalities/aortic coarctation, Eye abnormalities). Only 10% of these lesions require early intervention because of impairment of vision or swallowing, or airway compromise. Early intervention can include medical management, such as systemic steroids, intraleisional steroids, intraleisional interferon α-2a, or photocoagulation therapy, and surgical management, including excision with CO₂ laser/microdebrider and tracheotomy. Systemic steroids assist with rapidly proliferating lesions until the child reaches approximately one year of age; however, it is associated with growth retardation and immune suppression. Intraleisional interferon α-2a has been largely abandoned because it is a daily subcutaneous injection and is associated
with significant neurological side effects, including spastic diplegia. Photocoagulation therapy with either the flashlamp-pumped pulsed-dye laser (FPDL), the potassium titanyl phosphate (KTP) laser, or the neodymium yttrium-aluminum garnet (Nd:YAG) laser, is repeated every 4 to 6 weeks until the lesion disappears. A randomized trial recently demonstrated that propranolol was effective at a dose of 3mg/kg per day for 5 months in the treatment of infantile hemangioma with a very acceptable and low side-effect profile.57 Other groups have had success at discontinuing propranolol at 1 year of age with excellent outcomes.58 For patients who do not require early intervention, the lesion is observed every 3 months for involution after the proliferative phase has ended. Surgery is considered if regression has not occurred by 5 years of age because the cosmetic result is less likely to be satisfactory.

**Congenital hemangiomas** differ from infantile hemangiomas in that they reach their maximal size at birth and do not have a proliferative phase. There are two subtypes: rapidly involuting (RICH), which typically disappears by 1 of age with minimal fatty appearance upon resolution, and noninvoluting (NICM). The management is similar to infantile hemangiomas with the exception that medical management is not typically necessary.

**Vascular Malformations.** Vascular malformations, in contrast to infantile hemangioma, are always present at birth, although they may not be apparent for a few months. Although they do not have a proliferative phase, they grow with the patient, have hormonal growth spurts and do not involute.59 Vascular malformations can be classified as low flow (capillary, venous, lymphatic, and mixed), which comprise approximately two-thirds of all vascular malformations, or high flow (arteria and arteriovenous).

Capillary malformations arise from the cutaneous superficial plexus and are made up of capillary and postcapillary venules with a pink, red, or purple macular-papular appearance. Venous malformations arise from dilated vascular channels lined by normal endothelium; therefore, they are soft, compressible, and nonpalpable. If they are superficial, they will increase in size with Valsalva or dependent positioning. They can grow suddenly with trauma or in association with hormonal changes. Lymphatic malformations typically present at birth with the majority (90%) being identified by 2 years of age. They can be macrocystic (>2 cm), microcystic (<2 cm), or a combination. They are most commonly found in the head and neck, particularly on the neck, and on physical examination they are soft and doughy with normal overlying skin. Infrayoid lesions tend to be macrocystic, well circumscribed, and discrete and can be totally excised, whereas suprayoid lesions are typically microcystic, infiltrative, and excision is usually incomplete. On MRI, the best imaging modality for this malformation, a septated mass with low-intensity signal on T1 and high-intensity signal on T2 is noted. They grow slowly with the patient but can have a sudden increase in size with hemorrhage or infection. Rarely, they cause airway compromise, feeding difficulties, and failure to thrive.

Treatment of vascular malformations is based on depth, size, and growth pattern. Capillary malformations are typically treated with the pulsed dye laser (585 nm). Venous lesions can be treated with the KTP laser (532 nm) or the Nd:YAG laser (1064 nm), sclerotherapy, and, in select cases, complete surgical excision is possible. Arteriovenous malformations are rare but typically require surgical excision with negative margins often after embolization. Lymphatic malformations are typically treated at least in part with surgical excision, although this is less successful for microcystic lesions. OK-432 is lyophilized low virulence *S pyogenes* cultured in penicillin. It is used as a sclerotherapy agent for lymphatic malformations and has a 94% response rate in macrocystic lesions, a 63% response rate in mixed macromicrocystic lesions, and no response in microcystic lesions.60

**TRAUMA OF THE HEAD AND NECK**

**Soft Tissue**

Soft tissue trauma of the head and neck is managed with the same general surgical principles as any other body subsite with a few particularities. Most lacerations can be closed primarily if there is not soft tissue loss; even some devitalized soft tissue should be preserved because of the excellent blood supply to head and neck tissue that allows it to recover at a higher rate. Thus, minimal debridement is usually required. Thorough irrigation to remove foreign bodies and clean the tissue is required. This is followed by a careful layered closure. On the face, the deep layers are usually closed with a 3-0 or 4-0 Vicryl/Polysorb after a minimal amount of undermining, and interrupted 5-0 or 6-0 Prolene or Nylon is used for the skin. These sutures are removed at 5 days on the face. Antibiotics are reserved for through-and-through mucosal lacerations, contaminated wounds, bite injuries, and when delayed closure is performed (>72 hours). The chosen antibiotic should cover *S aureus*. Patients are instructed to avoid sunlight because this can cause pigmentary abnormalities in the suture line as it heals and matures over the first year.

Eyelid lacerations are closed in layers with careful reaproximation of the orbicularis oculi as a separate layer. Another important layer to reaproximate separately is the gray line (conjunctival margin) so as to avoid height mismatch or lid notching. Lip injuries follow the same principle with a three-layer closure involving the orbicularis oris, which is the strength layer, followed by careful reaproximation of the vermilion border to avoid a step-deformity (Fig. 18-16). Of course, a mucosal layer closure may also be required for through-and-through defects. Rarely, locoregional flaps or grafts are required for closure when greater than one-fourth of the eyelid width or one-third of the lip width is missing. Auricular hematoma is managed with prompt incision and drainage followed by bolstering technique; anteriorly and posteriorly placed dental pledgets secured with through-and-through sutures. These are to remain in place for at least 4 days to prevent reaccumulation of the hematoma and to prevent a cauliflower ear deformity. Auricular lacerations are typically closed primarily with perichondrial sutures to preserve the precarious cartilage blood supply followed by a primary closure of the skin, making sure to cover the cartilage to prevent chondritis. Given the rich vascular supply to the face and neck, many soft-tissue components that appear devitalized will indeed survive, and therefore minimal debridement of devitalized tissue is required.

Facial lacerations resulting in facial nerve injury are not explored if they are anterior to a vertical line dropped from the lateral canthus as there is excellent collateral innervation in the anterior midface. Posterior to this line, the nerve should be repaired, primarily if possible, using 8-0 to 10-0 monofilament suture to approximate the epineurium under the operative
microscope. If primary reapproximation is not possible due to a missing segment, cable nerve grafts can be performed using the sural nerve or the greater auricular nerve. If the buccal branch is injured, this raises suspicion regarding injury to the parotid duct, which lies along an imaginary line drawn from the tragus to the midline upper lip. The duct should be repaired over a 22-gauge stent or marsupialized into the oral cavity.

**Facial Fractures**

The most common facial fracture involves the mandible. Fig. 18-17 demonstrates the most common sites of fracture, which include the condyle (36%), body (35%), and angle (20%). In most cases, more than one site is involved due to reciprocating forces. The vector forces from the muscles of mastication, vertical from the masseter and horizontal from the pterygoid muscles, can cause a fracture to be favorable or unfavorable depending on the angle of the fracture line. After taking a history and performing a physical examination, imaging is performed in the form of a Panorex or a CT scan. Where closed reduction can be achieved, patients are placed in maxillomandibular fixation (MMF) with arch bars applied via circumdental wiring, and these are left in place for 4 to 6 weeks depending on patient factors and the fracture location. In elderly patients, this is kept in for 6 to 8 weeks. In children and patients with condylar fractures only 2 to 3 weeks is required, and this is important to prevent condylar ankylosis. During this time, patients are placed on a liquid diet and are provided with wire cutters in case of aspiration or airway emergency. Open reduction and fixation is indicated in patients with open, comminuted, displaced, or unfavorable fractures. In these patients, MMF is usually only temporary with a soft diet starting almost immediately in the postoperative setting. Because the MMF is temporary with rigid fixation, it is performed usually using the 4-point fixation technique, where the maxilla and mandible are held in occlusion by wires attached to intraoral cortical bone screws, with two screws above and below the occlusal line anteriorly. This is a benefit of open reduction and internal fixation because prolonged MMF is associated with gingival and dental disease, as well as with significant weight loss and malnutrition, during the fixation period. After fixation, the fracture is exposed, more commonly from a transcervical compared to a transoral approach. Care is made not to injure the marginal mandibular branch of the facial nerve during this exposure. A rigid, locking, load-bearing mandibular plate is used. In edentulous patients, determining the baseline occlusion is of less significance because dentures may be refashioned once healing is complete.

Midface fractures are rarely isolated and include multiple subsites. However, isolated zygoma fractures are typically displaced inferior inferiorly and medially with disruption of the suture lines between the temporal, frontal, and maxillary bones and the zygoma. If multiple zygoma fractures are present or if the zygomatic arch is significantly displaced, a coronal incision is required to perform the reduction and fixation. However, if it is an isolated depressed fracture, a Gilles reduction can be achieved inferiorly (transorally) or superiorly (along temporalis muscle). The pathophysiology of orbital blow-out fractures is (a) hydraulic from increased intraocular pressure or (b) buckling from direct bone conduction. This requires surgical intervention if there is a defect of $\geq 2 \text{ cm}^2$ or $>50\%$ of the floor with herniation. A forced duction test, where the musculature attachment of the inferior oblique is grasped with forceps and manipulated to determine passive ocular mobility, is performed to ensure that there is no inferior rectus entrapment. If there is entrapment, this would also result in diplopia with upward gaze. Blowout fractures demonstrating significant entrapment or enopthalmos are treated by orbital exploration and reinforcement of the floor with titanium mesh, hydroxyapatite, or split calvarial bone grafts. Sometimes, the anterior maxillary bone that has been fractured and is accessed in the process of repairing other fractures can also be used.

There are three classic patterns of more extensive midface fractures: Le Fort I, II, and III. However, fractures rarely follow this exact pattern, and the two sides of the face may have different Le Fort fractures. Nonetheless, a full understanding of midface buttresses is central in understanding these fractures (Fig. 18-18). There are three vertical buttresses: the nasofrontal-maxillary, the frontozygomaticomaxillary, and
DISORDERS OF THE HEAD AND NECK

CHAPTER 18

There are five horizontal buttresses: the frontal bone, nasal bones, upper alveolus, zygomatic arches, and the infraorbital region. Signs of midface fractures include subconjunctival hemorrhage, ocular signs/symptoms, malocclusion, facial asymmetry, midface hypoesthesia (V2), hematoma, and a mobile maxillary complex. Transverse maxillary alveolus fractures above the teeth are Le Fort I fractures, which may result in a mobile hard palate. When this fracture extends superiorly to include the nasofrontal buttress, medial orbital wall, and even as high as the infraorbital rim and zygomaticomaxillary articulation laterally, it is considered a Le Fort II. Mobility includes the palate, nasal dorsum, which is separated from the upper face, and the inferomedial aspect of the orbital rim. When the fracture disrupts the frontozygomaticomaxillary, frontomaxillary, and frontonasal suture line, craniofacial disjunction, a Le Fort III fracture. Of note, all of the Le Fort fractures involve the pterygoid plates posteriorly (Fig. 18-19).

Temporal Bone Fractures

Temporal bone fractures occur in approximately one fifth of skull fractures. Temporal bone fractures were previously classified as longitudinal or transverse describing the path along the temporal bone of the fracture line, but this has been largely replaced by the more relevant otic capsule sparing or involving classification given that most fractures are oblique. Otic capsule sparing fractures present with conductive hearing loss, ossicular injury, bloody otorrhea, and labyrinthine concussion. The facial nerve is rarely injured nor cerebrospinal fluid (CSF) leak common with this fracture pattern. However, in patients with otic capsule involving temporal bone fractures, typically caused by occipitomastoid impact, sensorineural hearing loss, vestibular dysfunction, facial nerve paralysis, and CSF leak are far more common. Regardless of the fracture pattern, when CSF leak is suspected, it usually resolves with conservative measures including bed rest, elevation of the head of the bed, stool softeners, and avoiding sneezing or straining. In some cases, a CSF drain can be placed if there is a delay in spontaneous resolution. Rarely will surgical repair be required. Unlike CSF leaks with temporal bone fractures, the facial nerve needs to be assessed and managed urgently. An incomplete or delayed facial nerve paralysis almost always resolves spontaneously with conservative measures, including oral steroids. An immediate complete paralysis that does not recover within 1 week should be prognosticated to consider nerve decompression. Electroneurography (ENoG), EMG, and nerve stimulation tests have been used to help determine which patients with delayed-onset complete paralysis will benefit from surgical decompression. The finding of >90% degeneration more than 72 hours after the onset of complete paralysis is considered an indication for surgery. A nerve excitability test, where thresholds are increased to elicit visible muscle contraction on each side, can indicate advanced degeneration when there is a difference of >3.0 to 3.5 mA between sides. Whether surgical intervention is indicated or not for facial nerve paresis, it is crucial to protect the eye because a corneal drying and abrasion can lead to blindness in the absence of eye closure and a blink reflex. This requires application of ocular lubricant at night with the eye taped shut, frequent artificial tears application while awake, and a humidity chapter.

TUMORS OF THE HEAD AND NECK

Squamous cell carcinoma (SCC) comprises >90% of all of the malignant pathology of the mucosal lining of the upper aerodigestive tract. Naturally, a discussion of tumors of the head and neck typically focuses on this pathology presenting from the lips and oral cavity to the larynx and hypopharynx. Management of these tumors requires a systematic approach. The ideal treatment protocol varies by subsite, stage, patient comorbidity, and center preference/experience. Given the relative rarity of these tumors, multidisciplinary management is of the utmost importance to provide the patient with a balanced perspective. This can be performed in the form of a multidisciplinary clinic where radiation and surgical oncologists simultaneously see the patient or through a tumor board where a new patient’s history, physical examination findings, imaging, and prior pathology
specimens are reviewed. This encourages discussion from multiple points of view concerning the most appropriate treatment options available. In addition to radiation and surgical oncology, medical oncology, dentistry, speech language pathologists, radiologists, and pathologists contribute to the decision-making in this patient population. Some of the greatest advances in head and neck oncology over the last several decades include the development of standardized organ preservation protocols, advances in free flap reconstruction with microvascular techniques, and vaccinations. The future of head and neck oncology is bright with advances in molecular biology, immunotherapy, and preventative methods with vaccination. These have the potential of significantly decreasing incidence rates and improving survival and quality of life for those with the disease.

Etiology and Epidemiology
The main etiological factors associated with head and neck cancers are tobacco products and alcohol. Overall, there has been a decline in incidence of head and neck cancers of the oral cavity and larynx/hypopharynx subsites, likely related to public health campaigns and government taxation policies as it relates to cigarette consumption. Similarly, the incidence of head and neck cancer between countries varies widely and is strongly associated with the incidence of cigarette smoking. Cigarette smoking triples the likelihood of developing an oral cavity cancer, while the addition of alcohol synergistically increases the likelihood by 10- to 15-fold. The risk increases as the number of years smoking and number of cigarettes smoked per day increases. Individuals who both smoke (two packs per day) and drink (four units of alcohol per day) had a 35-fold increased risk for the development of a carcinoma compared to controls.

The preoperative and perioperative periods are excellent opportunities for head and neck oncologists to pursue a smoking cessation intervention. Continued smoking after completion of treatment is associated with a 3- to 4-fold increased risk of developing a second primary or recurrent tumor. A study assessing patients diagnosed with a new head and neck cancer demonstrated that of the patients that were smoking at diagnosis, only 54% were able to quit, highlighting the difficulty this population has with smoking cessation.

Betel nut/quid chewing, which is a product of the areca catechu tree, is endemic to some parts of Asia and India, and in these regions oral cavity cancer is one of the most common cancers. Betel nut when chewed acts as a mild stimulant similar to that of coffee but can be associated with submucous fibrosis that adds an additional challenge in the management of patients who present with a concurrent oral cavity cancer. These products are associated with particular subsites secondary to direct contact (e.g., buccal mucosa) as well as subsites with dependent saliva drainage (e.g., floor of mouth, mandibular alveolus, and wet lip). Reverse smoking, where the lighted portion of the tobacco product is placed within the mouth during inhalation is also associated with oral cavity cancer, specifically hard palate carcinoma. The risk for this cancer is 47 times greater in patients that exhibit this behavior compared to nonsmokers.

In Europe and North America there has been an increasing interest in decriminalizing marijuana smoking. There is a strong correlation between this activity and head and neck cancers, particularly those who have received solid organ and bone marrow transplants are at an increased risk of head and neck cancers. Similarly, HIV-infected patients have a higher incidence of head and neck cancers, and despite aggressive treatment have poorer results compared to HIV-negative patients. Other conditions associated with oral cancer include Plummer-Vinson syndrome (iron-deficiency anemia, dysphagia, glossitis, cheilitis, and esophageal webs), dyskeratosis congenita, Bloom’s syndrome, and Fanconi anemia.

HPV is a double stranded DNA virus that is transmitted through sexual contact. Over the last two decades, this virus, specifically the 16 and 18 subtypes, has been associated with an epidemic rise in oropharyngeal squamous cell carcinoma. The p16 protein is a surrogate for HPV positivity. HPV status in oropharynx cancer has prognostic and therefore treatment-related implications.

Anatomy and Histopathology
The upper aerodigestive tract is divided into several distinct sites that include the oral cavity, pharynx, larynx, and nasal cavity/paranasal sinuses. Each of these sites has separate subsites as alluded to earlier with specific etiological, pathological, prognostic, and treatment-related peculiarities. Locoregional tumor spread is determined by weaknesses in the framework, fascial planes, and the course of neurovascular and lymphatic channels.

The oral cavity extends from the vermillion border of the lip to the hard-palate/skin-palate junction superiorly, to circumvallate papillae inferiorly, and to the anterior tonsillar pillars laterally. It is divided into eight subsites including the (a) mucosal lip, (b) the mandibular alveolus, (c) floor of mouth, (d) tongue (anterior two-thirds), (e) buccal mucosa, (f) retromolar trigone, (g) maxillary alveolus, and (h) hard palate (Fig. 18-20). Advanced oral cavity cancer can present with mandibular and/or maxillary invasion requiring resection, at least in part, of these structures. Oral cavity cancers typically metastasize to the submental, submandibular, and upper jugular lymph nodes (levels I-III).

Figure 18-20. Oral cavity landmarks.
The pharynx is divided into three regions: nasopharynx, oropharynx, and hypopharynx (Fig. 18-21). The nasopharynx extends from the posterior nasal septum and choana to the skull base and includes the fossa of Rosenmüller and torus tubarius of the Eustachian tubes laterally. The inferior margin of the nasopharynx is the superior surface of the soft palate. In adults, the adenoids are typically absent secondary to involution during late adolescence, but these can be seen in some adults in the posterior aspect of this subsite. Isolated posterior triangle (level V) lymphadenopathy in an adult should be considered nasopharyngeal carcinoma (NPC) until proven otherwise. Due to its midline location, bilateral regional metastatic spread is common in nasopharyngeal carcinoma. Given the epidemic rise of oropharyngeal cancers, isolated level V adenopathy in an adult may also represent oropharyngeal cancer, although cancers at this site typically drain to the upper and lower cervical nodes (levels II–IV) as well as the retropharyngeal nodes. The oropharynx has a number of subsites including the tonsillar region, base of tongue, soft palate, and posterolateral pharyngeal walls. The hypopharynx extends from the vallecula to the lower border of the cricoid posterior and lateral the larynx. It includes several subsites as well including the pyriform fossa, the postcricoid space, and the posterior pharyngeal wall. Lymphatic drainage is to the mid and lower cervical nodes (levels III–IV); however, usually the upper cervical nodes (level II) are addressed at the same time for tumors at this site.

The larynx is divided into three regions: the supraglottis, glottis, and subglottis (Fig. 18-22). The supraglottis includes several subsites: the epiglottis, false vocal cords, medial surface of the aryepiglottic folds, and the upper half of the laryngeal ventricles. The glottic larynx includes the true vocal cords, the anterior and posterior commissure, and the lower half of the laryngeal ventricles. The subglottis extends from below the true vocal cords to the superior cricoid border from within. The supraglottis has a high rate of bilateral metastatic spread secondary to its rich lymphatic drainage, whereas isolated glottic cancers rarely have lymphatic spread. Laryngeal cancers, in addition to having the propensity for lymphatic spread, particularly in advanced cases, can have preepiglottic and paraglottic invasion as well as invasion of the laryngeal framework (thyroid and cricoid cartilage). Furthermore, glottic and subglottic lesions, in addition to potential spread to the upper and lower cervical nodes (levels II–IV), have the propensity for spread to the central neck (level VI) in the paralaryngeal and paratracheal region.

Second Primary Tumors in the Head and Neck
Patients with head and neck squamous cell carcinoma (HNSCC) are at increased risk for the development of a second primary malignancy (SPM), which is defined as a second malignancy that presents either simultaneously or after the diagnosis of an index tumor. A synchronous SPM is diagnosed simultaneously or within 6 months of the index tumor, while a metachronous SPM is diagnosed >6 months after the index tumor. SPMs need to be distinguished from local recurrences or metastasis of the primary tumor. The incidence of SPM ranges from 2% to 7% per year, and this risk remains constant from the time of initial diagnosis throughout the lifetime of the patient. Secondary primary malignancies represent the second leading cause of death in patients with HNSCC. One-quarter to one-third of deaths in these patients are attributable to SPM, highlighting the importance of SPM in the successful management of HNSCC.

The classic criteria for defining second primary malignancy (SPM) were proposed by Warren and Gates and are: (a) histologic confirmation of malignancy in both the index and secondary tumors; (b) two malignancies that are anatomically...
separated by normal mucosa; and (c) the possibility of the SPM being a metastasis from the index tumor must be excluded. Most investigators use these criteria to define an SPM. However, disagreement exists regarding the application of the second and third criteria. For example, when both tumors appear in the same anatomic subsite, there is no agreement on the distance that should exist between the tumors, with some investigators favoring 1.5 cm and others requiring 2 cm. Furthermore, when the tumors occur in the same anatomic subsite, some investigators add that the SPM must present at least three years after the diagnosis of the index tumor, while others require that the SPM present at least five years after the index tumor. Others suggest that molecular analysis is required to classify a tumor as an SPM.

Treatment of SPMs of the upper aerodigestive tract is site specific. In general, the SPM should be treated as a separate entity, in the same manner as a primary index tumor at the anatomic subsite. In many cases, particularly in metachronous SPMs, patients have already received a full complement of treatment, including primary or adjuvant radiation and/or chemoradiation treatment. In these cases, surgical treatment of the SPM is often indicated when feasible. Reirradiation is an option in carefully selected cases when salvage surgery is not possible. Proper patient selection for reirradiation is critical, and only patients with minimal comorbidity and toxicity of previous radiation treatment should be considered. Patients at high risk for local recurrence after salvage surgery may benefit from increased locoregional control from adjuvant reirradiation, although there is no survival advantage compared with salvage surgery alone. Survival in patients with SPM depends upon the stage and location of the primary site of the SPM. Patients with SPM arising in the head and neck have significantly improved survival when compared with patients with SPM arising in the lung and esophagus.

Staging
Staging for upper aerodigestive tract malignancies is defined by the American Joint Committee on Cancer and follows the TNM (primary tumor, regional nodal metastases, distant metastasis) staging format which was recently updated in the 8th edition in 2017. The T stage for each subsite incorporates relevant anatomy; for instance, T3 lesions of the glottis are associated with vocal cord immobility. Recent changes have incorporated HPV/P16 status for oropharynx cancer (Tables 18-1 and 18-2) and depth of invasion for oral cavity cancers (Table 18-3).

The N classification for head and neck sites is nearly uniform for all sites (Tables 18-4 and 18-5) except for the nasopharynx and for HPV-associated (p16-positive) oropharynx cancer. Recent changes have also incorporated extracapsular extension into this nodal staging to improve the discrimination and prognostication of the classification.

Upper Aerodigestive Tract
There are similarities in the initial assessment and management of all patients with a newly diagnosed upper aerodigestive tract malignancy. The frequently reviewed clinical practice guidelines (National Comprehensive Cancer Network; NCCN) provide valuable information by site and stage with regard to workup and management and should be used to direct care.

After a thorough history that should include assessment of the previously discussed risk factors, a comprehensive physical examination should follow. A full head and neck examination including inspection and palpation is critical for nearly all head and neck cancers. Oral cavity and oropharyngeal cancers should be palpated when possible to provide additional tactile information regarding depth of invasion, mobility, and invasion into adjacent structures. A cranial nerve (CN) examination with a focus on the assessment of trigeminal (V2/V3) parasthesia/paresthesia.

| Table 18-1 | Clinical and pathologic T category for HPV-associated (p16-positive) oropharyngeal cancer |
| T CATEGORY | T CRITERIA |
| T0 | No primary identified |
| T1 | Tumor 2 cm or smaller in greatest dimension |
| T2 | Tumor larger than 2 cm but not larger than 4 cm in greatest dimension |
| T3 | Tumor larger than 4 cm in greatest dimension or extension to lingual surface of epiglottis |
| T4 | Moderately advanced local disease |

"Mucosal extension to lingual surface of epiglottis from primary tumors of the base of the tongue and vallecula does not constitute invasion of the larynx."

T CATEGORY | T CRITERIA
--- | ---
TX | Primary tumor cannot be assessed
Tis | Carcinoma in situ
T1 | Tumor ≤2 cm, ≤5 mm depth of invasion (DOI). DOI is depth of invasion and not tumor thickness.
T2 | Tumor ≤2 cm, DOI >5 mm and ≤10 mm or tumor >2 cm but ≤4 cm, and DOI ≤10 mm
T3 | Tumor >4 cm or any tumor with DOI >10 mm but ≤20 mm
T4 | Moderately advanced or very advanced local disease
T4a | Moderately advanced local disease
Tumor invades adjacent structures only (e.g., through cortical bone of the mandible or maxilla, or involves the maxillary sinus or skin of the face) or extensive tumor with bilateral tongue involvement and/or DOI >20 mm.
Note: Superficial erosion of bone/tooth socket (alone) by a gingival primary is not sufficient to classify a tumor as T4.
T4b | Very advanced local disease
Tumor invades masticator space, pterygoid plates, or skull base and/or encases the internal carotid artery

Table 18-4
Clinical N category for non–HPV-associated (p16-negative) oropharyngeal cancer

N CATEGORY | N CRITERIA
--- | ---
NX | Regional lymph nodes cannot be assessed
N0 | No regional lymph node metastasis
N1 | Metastasis in a single ipsilateral lymph node, 3 cm or smaller in greatest dimension and ENE(−)
N2 | Metastasis in a single ipsilateral node larger than 3 cm but not larger than 6 cm in greatest dimension and ENE(−); or metastases in multiple ipsilateral lymph nodes, none larger than 6 cm in greatest dimension and ENE(−); or in bilateral or contralateral lymph nodes, none larger than 6 cm in greatest dimension and ENE(−)
N2a | Metastasis in a single ipsilateral node larger than 3 cm but not larger than 6 cm in greatest dimension and ENE(−)
N2b | Metastasis in multiple ipsilateral nodes, none larger than 6 cm in greatest dimension and ENE(−)
N2c | Metastasis in bilateral or contralateral lymph nodes, none larger than 6 cm in greatest dimension and ENE(−)
N3 | Metastasis in a lymph node larger than 6 cm in greatest dimension and ENE(−); or metastasis in any node(s) and clinically overt ENE(+)
N3a | Metastasis in a lymph node larger than 6 cm in greatest dimension and ENE(−)
N3b | Metastasis in any node(s) and clinically overt ENE(+)

ENE = extranodal extension.

Lip. The lips starting at the vermillion border represent a transition between external skin to internal mucosa. The sphincter function of the lip is created by activation of the circumferential musculature of the orbicularis oris, a critical structure in lip form and function. Lip cancers are most common in men and are often seen in those with fairer complexions. In addition to tobacco use and immunosuppression, UV exposure is an additional important risk factor unique to this head and neck subsite. The majority (>90%) of lip cancers present on the lower lip due to its increased protrusion and increased sun exposure.

Although the vast majority of lip cancers are SCC, other cutaneous malignancies such as basal cell carcinoma and malignant melanoma are not uncommon at this subsite.

Basal cell carcinoma presents more frequently on the upper lip than lower.

Negative prognostic factors for lip cancers include perineural invasion, invasion into bone (maxilla or mandible), upper
lip or oral commissure involvement, positive regional metastasis, and young age at diagnosis.

The primary management of lip cancer is a surgical resection of the primary site with an adequate margin (1 cm). This provides margin analysis and additional pathologic information that can help stratify which patients may benefit from adjuvant treatment. The primary regional nodal drainage basin for lip cancers is the submandibular, submental, and perifacial nodes (level I), and metastases occur in <10% of patients with a higher incidence in those with upper lip cancers.\(^\text{108}\) When there are clinical evident notes, a neck dissection is indicated. Otherwise, in the clinically and radiographically negative neck involvement, a neck dissection is indicated.

### Table 18-5

<table>
<thead>
<tr>
<th>N CATEGORY</th>
<th>N CRITERIA</th>
</tr>
</thead>
<tbody>
<tr>
<td>NX</td>
<td>Regional lymph nodes cannot be assessed</td>
</tr>
<tr>
<td>N0</td>
<td>No regional lymph node metastasis</td>
</tr>
<tr>
<td>N1</td>
<td>Metastasis in a single ipsilateral lymph node, 3 cm or smaller in greatest dimension ENE(−)</td>
</tr>
<tr>
<td>N2</td>
<td>Metastasis in a single ipsilateral node larger than 3 cm but not larger than 6 cm in greatest dimension and ENE(−); or metastases in multiple ipsilateral lymph nodes, none larger than 6 cm in greatest dimension and ENE(−); or in bilateral or contralateral lymph nodes, none larger than 6 cm in greatest dimension, and ENE(−)</td>
</tr>
<tr>
<td>N2a</td>
<td>Metastasis in a single ipsilateral node larger than 3 cm but not larger than 6 cm in greatest dimension, and ENE(−)</td>
</tr>
<tr>
<td>N2b</td>
<td>Metastasis in multiple ipsilateral nodes, none larger than 6 cm in greatest dimension, and ENE(−)</td>
</tr>
<tr>
<td>N2c</td>
<td>Metastasis in bilateral or contralateral lymph nodes, none larger than 6 cm in greatest dimension, and ENE(−)</td>
</tr>
<tr>
<td>N3</td>
<td>Metastasis in a lymph node larger than 6 cm in greatest dimension and ENE(−); or metastasis in any node(s) and clinically overt ENE(+)</td>
</tr>
<tr>
<td>N3a</td>
<td>Metastasis in a lymph node larger than 6 cm in greatest dimension and ENE(−)</td>
</tr>
<tr>
<td>N3b</td>
<td>Metastasis in any node(s) and clinically overt ENE(+)</td>
</tr>
</tbody>
</table>

ENE = extranodal extension.


The preferred approach to management of these tumors is a surgical resection with adequate (1 cm) surgical margins with management of the regional nodal basin. In general, tumors of the oral cavity metastasize to the submandibular, submental, and upper cervical nodes and are almost always treated with a suprathyroid neck dissection at the time of primary resection with a few rare exceptions (T1 oral tongue lesions that have less than 4 mm depth of invasion). In the “Neck” section of this chapter, we will discuss this in more detail. Adjuvant radiotherapy is indicated in patients with close margins, regional lymphadenopathy, advanced stage tumors (T3/T4), perineural invasion, and lymphovascular invasion, while adjuvant chemoradiotherapy is reserved for those with positive margins or extracapsular invasion.\(^\text{115,114}\)

### Oral Tongue

The oral tongue is a muscular structure composed of intrinsic (longitudinal, vertical, and transverse muscle fibers) and extrinsic (genioglossus, hyoglossus, styloglossus, and palatoglossus) muscles separated by a midline raphe and has overlying nonkeratinizing squamous epithelium. The posterior limit of the oral tongue is the circumvallate papillae beyond which the oropharynx begins while the ventral portion is contiguous with the anterior floor of mouth.
Tumors of the tongue typically start along the epithelial surface and can be endophytic or exophytic with or without ulceration (Fig. 18-26) and are typically seen on the lateral and ventral surfaces of the tongue. Lesions on the dorsal aspect of the tongue, particularly along the midline, are less likely to be malignant. What is seen on the surface is typically the tip of the iceberg, and palpation can provide further information regarding the depth of invasion of the tumor. These tumors can be extensive, and when they cross the midline and start to involve the base of tongue an extensive surgical resection including a total glossectomy may be required. However, most tumors present at an early stage due to significant pain, otalgia, voice change secondary to difficulties with articulation, and dysphagia, which may lead to weight loss. On history and physical examination, ipsilateral paresthesias and deviation of the tongue protrusion with fasciculations or atrophy may indicate lingual nerve and hypoglossal nerve tumor invasion respectively (Fig. 18-27).

Early lesions (T1–T2) can be closed primarily, allowed to heal by secondary intention, or reconstructed with a split thickness flap. Figure 18-23. Estlander flap. A. Intra-operative image of lower lip squamous cell carcinoma with buccal and cutaneous extension pre-excision; B. Intra-operative defect and Estlander flap design. C. Immediate post-operative flap. D. One year post-operative image.

Figure 18-24. A-C. Karapandzic labiaplasty for lower lip carcinoma.
skin graft after partial glossectomy. This procedure allows reasonable speech and swallowing function as long as there is not significant tethering in the floor of the mouth if this has been resected. Articulation is determined by premaxillary contact of the tongue, and dental appliances can be used in the postoperative setting to improve this. Tongue protrusion and lateral movement predicts a patient’s ability to swallow, and this is less difficult to repair secondarily. Therefore, many patients, even with small tongue cancers that require significant floor of mouth resection, receive soft pliable fasciocutaneous free flap reconstruction to improve these functional outcomes. Advanced lesions that require a more radical resection require free flaps, which obliterate the oral cavity dead space while creating bulk in the posterior oropharynx to improve the pharyngeal swallowing phase.

---

**Figure 18-25.** Bernard burrow flap reconstruction for a total lower lip defect involving upper and lip advancement rotation flap and cheek advancement.

**Figure 18-26.** Oral tongue squamous cell carcinoma.

**Figure 18-27.** A and B. Anatomy of the floor of mouth and submandibular space. a. = artery; m. = muscle; n. = nerve.
Floor of Mouth The floor of mouth is a mucosal-covered semilunar area that extends from the anterior tonsillar pillar posteriorly to the frenulum anteriorly, and from the inner surface of the mandible to the ventral surface of the oral tongue. The ostia of the submaxillary and sublingual glands are contained in the anterior floor of mouth. The muscular floor of mouth is composed of the sling-like genioglossus, mylohyoid, and hyoglossus muscles, which serve as a barrier to the spread of disease. Invasion into these muscles can result in decreased tongue mobility and poor articulation.

The floor of mouth begins just below the lingual surface of the mandibular alveolus and ends at the ventral tongue where the frenulum connects the floor of mouth to the tongue along the midline and at the anterior tonsillar pillars posteriorly. Just deep to the floor of mouth mucosa is the submandibular (Wharton’s) duct and sublingual minor salivary glands followed by the genioglossus, hyoglossus, and mylohyoid muscles. Direct invasion of these structures is not uncommon and can result in direct spread to the sublingual and submandibular spaces as well as decreased tongue mobility, leading to articulation complaints. The lingual nerve (a branch of V3) provides sensory innervation to this subsite and is in close proximity to it, often requiring resection of this structure. The contiguity of the floor of mouth mucosa with the lingual surface of the mandible can lead to mandibular invasion. This needs to be carefully examined bimanually on physical examination and using imaging (CT, MRI, or Panorex) because a marginal or segmental mandibulectomy may be required to excise these tumors (Fig. 18-28). If the lesion is not fixed to the mandibular cortex on physical examination, then a mandible-sparing procedure is feasible.117

Extension to the sublingual and submandibular ducts and spaces requires that the neck dissection specimen be removed en bloc with the primary tumor. Invasion of the intrinsic tongue musculature requires a partial glossectomy. In our experience, except for the smallest (T1) very superficial floor of mouth lesions, cancers at this subsite nearly always require a reconstructive procedure to separate the floor of mouth from the neck and to avoid tethering of the tongue using a pliable fasciocutaneous flap. If a segmental resection is performed, the vascularized osteocutaneous free flap is used. Given the anterior location of this tumor, a lip-splitting incision is rarely used unless resection of lip and chin skin is required as part of the resection in a select group of T4a tumors with through-and-through involvement.

Mandibular Alveolus and Gingiva The alveolar mucosa overlies the bone of the mandible and extends from the ginvobuccal sulcus to the mucosa of the floor of mouth to the second and third molar, which is the anterior border of the retromolar trigone subsite. Treatment of these lesions requires at the very least marginal resection of the mandibular bone given the proximity and early invasion of the periosteum in this region. A marginal resection is acceptable if there is only very early bony invasion (Fig. 18-29). If the inferior alveolar canal or the medullary cavity is invaded on physical examination or preoperative imaging, a negative locoregional prognostic factor, a segmental resection is recommended with appropriate reconstruction.118,119

Retromolar Trigone The retromolar trigone (RMT) is bordered medially by the anterior tonsillar pillar, anteriorly by the
second or third molar, posteriorly by the maxillary tuberosity, inferiorly by the posterior mandibular alveolus, superiorly by the coronoid process of the mandible, and laterally by the buccal mucosa. Negative margin resection often requires a marginal shave mandibulectomy, even when there is no evidence of mandibular cortical invasion, because of the close proximity to the mandibular periosteum. This is typically achieved through a transoral approach while carefully protecting the lips and cheek. Extension to adjacent subsites including the buccal mucosa, maxillary tuberosity, floor of mouth, and posterolateral tongue often requires these structures be resected as part of the margin. Trismus at this and other subsites is an advanced indication of involvement of the muscles of mastication in the masticator space, which can extend to the skull base. These tumors are aggressive. Infiltration into the masticator space and bony invasion (maxilla more often than mandible) significantly worsens the prognosis.

**Buccal Mucosa** The buccal mucosa includes all of the mucosal lining from the inner surface of the lips to the line of attachment of mucosa of the alveolar ridges and pterygomandibular raphe. The mucosa includes the parotid (Stenson’s) duct opening adjacent to the first and second maxillary molars. An understanding of the layers of the cheek from medial to lateral is important because these layers are very closely adherent to the buccal mucosa. Therefore, tumors in this region have a high propensity for early deep invasion and early lymphatic spread. The layers of the cheek from medial to lateral are: (a) buccal mucosa, (b) pharyngobasilar fascia, (c) buccinator muscle, (d) buccopharyngeal fascia, (e) buccinator fat pad, (f) masseter muscle, (g) muscles of facial expression and the superficial muscular aponeurotic system (SMAS), (h) subcutaneous tissue, and (i) facial skin. It is not uncommon for tumors with deep invasion into the cheek to require a through-and-through resection. Reconstruction aimed at providing both an internal and external lining may be accomplished with a folded fasciocutaneous free flap or a combination of a local flap for the external component and a free flap for the internal component. Marginal bone resection is often required in tumors that extend to the mandibular or maxillary alveolus.

**Maxillary Alveolus and Hard Palate** The hard palate and maxillary alveolus have classically been considered two separate subsites, but due to their anatomic contiguity and the similarities in their oncologic outcomes these two subsites should be discussed together. The junction between the hard palate and soft palate is the posterior border, while the maxillary tuberosity is the posterolateral border separating the retromolar trigone from the maxillary alveolus. The periosteum is at this subsite is closely adherent to the mucosa, and as such, superficial lesions require resection of the bone to achieve a clear margin. An infrastructure maxillectomy may be required for larger lesions involving the palate or maxillary antrum. The greater palatine nerve and foramen can be a pathway for neuropathic spread, and it is important to identify perineural invasion on these tumors in the biopsy specimen.

Although SCC continues to be the primary malignant pathology at this subsite, minor salivary gland tumors such as adenoid cystic carcinoma, mucoepidermoid carcinoma, and adenocarcinoma can also present in this location. Minor salivary gland tumors tend to arise at the junction of the hard and soft palate.

Nonmalignant pathology includes necrotizing sialometaplasia, which appears as a butterfly-shaped ulcer on the hard palate that otherwise looks like a neoplasm. Treatment is symptomatic as these lesions typical disappear with time; however, a biopsy is warranted to confirm the diagnosis. A torus palatinus is a benign bony outgrowth seen on midline of the hard palate. This does not require biopsy to confirm the diagnosis and only requires treatment to relieve symptoms.

Reconstruction of the maxillectomy defect depends on a number of variables, including patient preference, dentition, patient comorbidity, and extent of defect. A partial palatoplasty or partial infrastructure palatoplasty can often be reconstructed with a dental obturator or a soft tissue flap alone to separate the oral cavity from the nasal cavity and maxillary sinus. More extensive suprastructure maxillectomies can be reconstructed with a free flap composed only of soft tissue, although this will leave the patient with a significant malar asymmetry over an osseous free flap. The layered fibular free flap and the scapular tip have been recently popularized to reconstruct more extensive orbitomaxillary reconstruction. Supporting the orbital floor when it is resected is critical in supporting the orbital contents and avoiding eventual diploplia because there can be a drop in these contents when they are not supported.

**Oropharynx** The borders of the oropharynx start at the soft palate superiorly, the hyoid (vallecular root) inferiorly, the anterior tonsillar pillar anterolaterally, and the cricuvallate papilla at the junction between the anterior two-thirds and posterior third of the tongue. There are five subsites in the oropharynx: the tonsillar region that includes the anterior and posterior tonsillar pillars, the soft palate, the posterior pharyngeal wall, the lateral pharyngeal wall, and the base of tongue. Tumors at this subsite can have direct extension laterally in the parapharyngeal space, posteriorly into the retropharyngeal space, anteriorly into the oral cavity, superiorly into the nasopharynx, or inferiorly into...
the supraglottic larynx. Laterally, through the superior constrictor, invasion of the jugular vein, carotid artery, and cranial nerves IX to XII, as well as the sympathetic chain, is possible. The pharyngobasilar fascia (resectable) deep to the constrictor muscles is a natural barrier from invasion into the prevertebral fascia (unresectable). The ascending ramus of the mandible can be involved when tumors invade the medial pterygoid muscle.

Although SCC is the predominant pathology, minor salivary gland tumors can present as submucosal lesions in the soft palate or tongue base, and lymphoma can present in the tonsils as an asymmetric enlargement, underlying the importance of a tissue diagnosis before treatment.

Oropharyngeal cancers, other than those on the soft palate or tonsils, are often not obvious on oral cavity exam inspection; therefore, a high degree of suspicion should exist in patients with a muffled voice as would be experienced in tongue base tumors, patients with dysphagia and weight loss, or referred otalgia from the tympanic branches of CN IX and X. Trismus may indicate advanced disease with pterygoid involvement. As previously mentioned, because of the epidemic rise in incidence of oropharyngeal cancers, secondary to HPV-associated tumors, and the high regional metastatic rate for these tumors, the presenting symptom is often a nontender cervical lymphadenopathy, which should be investigated with a fine-needle aspiration (FNA) biopsy. Approximately 50% of patients have metastases at the time of diagnosis. Bilateral metastases are common in patients with soft palate and base of tongue tumors. Treatment of the neck should include the upper jugulodigastric nodes to which these tumors most commonly metastasize to, followed by levels II, IV, V, and the retropharyngeal lymph nodes.

A discussion about oropharyngeal cancer cannot be had without discussing the important prognostic information provided by the HPV status of these tumors. The incidence of oropharyngeal squamous cell carcinoma has increased significantly over the last four decades secondary to HPV-16 related development of this tumor.125 HPV infection can induce the production of two viral oncoproteins, E6 and E7, which inactivate tumor suppressors p53 and Rb leading to tumor promotion.126 HPV-positive tumors are more common in younger male patients and are associated with a history of a higher lifetime number of sexual partners and oral sex.127 Ang et al demonstrated that oropharyngeal cancers can be stratified on overall survival into low risk (HPV-positive tumors in patients with ≤10 pack years of smoking or >10 pack years of smoking but N0-N2a), intermediate risk (HPV-positive tumors with >10 pack years of smoking and N2b-N3 or HPV-negative tumors in patients with ≤10 pack years of smoking and T2-T3 tumors), and high risk (HPV-negative tumors in patients with ≤10 pack years of smoking and T4 tumors or HPV-negative tumors in patients with >10 pack years of smoking).12 The rate of distant metastases in the HPV-positive and HPV-negative tumors does not differ, and therefore the survival benefit in the HPV-positive group is due to improved locoregional control.

Management of squamous cell cancers of this region includes single modality (surgery or radiotherapy alone) treatment for early stage disease (stage I/II) and multimodality treatment for advanced stage (stage III/IV) disease (surgery followed by postoperative radiotherapy or concurrent chemoradiotherapy).108 Historically, from 1971 to 2000, oropharyngeal cancers, at the time mostly HPV-negative, were treated heterogeneously with surgery followed by radiotherapy or primary radiotherapy similar survival until Parsons et al demonstrated in a meta-analysis similar survival rates between the two treatment groups with improved locoregional control in the radiation-alone group and much higher complication rates in the surgery group (32% severe complications, 3.5% mortality) compared to the radiotherapy group (3.8% severe complications, 0.4% mortality).129 For this reason, for many years, advanced-stage tumors were treated with primary concurrent chemoradiotherapy. However, this is now a moving target given the excellent results in early and some intermediate-stage HPV-positive disease regardless of treatment. More recently, there has been a push to study de-escalation, particularly in the aforementioned low and intermediate risk groups given the excellent survival rates. The standard of care, regardless of HPV status, for advanced tumors (T3/T4 or N2b-N3 or evidence of gross ECE) continues to be concurrent chemoradiotherapy.129

The high complication and mortality rate in the surgical group analyzed by Parsons et al was associated not just with HPV-negative tumors but also with open resections for advanced tumors that necessitated a lip-splitting mandibulotomy approach. More recently, particularly for early stage tumors (T1, T2, N0-N2a), there has been a push towards minimally invasive transoral robotic surgery (TORS) using the da Vinci Surgical System. Oncologic outcomes are similar between surgery and radiotherapy in this group, and TORS has been demonstrated to be cost-effective in this setting.130-132 Functional outcomes related to swallowing (G-tube dependency) and airway (tracheotomy dependency) are also similar between the groups.130 These outcomes are heavily dependent on the surgeon’s ability to achieve negative margins, which can be challenging, and on good preoperative predictive value of imaging to stage the neck, given that advanced nodal disease, particularly with ECE, continues to benefit from adjuvant chemoradiotherapy. Positive margins or ECE ultimately leads to adjuvant chemoradiotherapy. This results in triple modality treatment with its associated higher morbidity. Therefore, clinical recommendations based on these favorable early retrospective poorly controlled studies with small sample sizes is not yet possible. Meanwhile, clinical trial evidence is pending to help elucidate in which settings and patients this new approach may be beneficial.133

Extensive oropharyngeal cancers that fail concurrent chemoradiotherapy are treated with resection. If the mandible is involved, a marginal mandibulectomy or segmental mandibulectomy may be required depending on the extent of bony involvement. Tongue base resection may necessitate total glossectomy depending on the contralateral extent of the tumor and the ability to save the lingual artery and to a lesser extent the hypoglossal nerve on that side. When the larynx is preserved many patients, if careful reconstruction is performed, 90% of patients can be decannulated and have acceptable voice outcomes.134 However, it is not uncommon to have to perform a total laryngectomy at the same time as the total glossectomy for tumors with supraglottic extent, and this is associated with poor quality of life. Generally, these patients also have poorer survival.135-137

The primary goal of oropharyngeal reconstruction is swallowing rehabilitation. For soft palate defects, palatal obturators may assist in providing a seal between the nasopharynx and the posterior pharyngeal wall. The modified Gehanno technique sutures the posterior wall of the remaining soft palate to the remaining incised pharyngeal mucosa to close off the ipsilateral hemi-nasopharyngeal port.138,139 A flap can then be inset overlying this defect, which has effectively separated the nasopharynx from the oropharynx. This prevents nasal regurgitation of air
and liquids, therefore impacting both speech and swallowing. Similarly, total glossectomy reconstruction has several goals, including filling the oral cavity dead space, allowing the neotongue to reach the premaxilla to assist with articulation, and, most importantly, creating posterior bulk to allow the base of tongue to touch the posterior pharyngeal wall, which assists with the pharyngeal phase of swallowing. This is often achieved with a large rectus abdominis or anterolateral thigh free flap. If the neotongue does not successfully touch the premaxilla and hard palate and speech is impeded, a palatal obturator can be used to bring down the palate and achieve better contact.

**Hypopharynx and Cervical Esophagus** The hypopharynx, which extends from the vallecular to the lower border of the cricoid cartilage (Fig. 18-30), has three subsites; the piriform sinuses, the lateral and posterior pharyngeal walls, and the post cricoid space. SCC of the hypopharynx typically presents with progressive dysphagia, first to solids then to liquids, followed by weight loss. Similar to oropharyngeal tumors, patients can also present with voice change, referred otalgia or a neck mass. Rarely, when the larynx is involved, patients may present with stridor and airway compromise necessitating an urgent tracheotomy.

Unfortunately, there is significant delay in diagnosis of patients with hypopharyngeal cancer and late presentation is common. Routine physical examination will not typically detect the tumor. Fiberoptic nasolaryngoscopy is important in assessing the extent of the tumor and laryngeal function. Vocal cord paralysis is a poor prognostic factor and indicates fixation of the cricoarytenoid joint from direct extension of the tumor or recurrent laryngeal nerve invasion. A Valsalva maneuver during laryngoscopy allows for a better evaluation of the opened pyriform sinuses and postcricoid space. Functional endoscopic evaluation of swallowing (FEES) can be useful to assess laryngeal penetration and aspiration, but a modified barium swallow (MBS) is better at assessing inferior extent of the disease, multifocality within the esophagus, and aspiration. A thorough metastatic workup is required, with special attention paid to paratracheal and upper mediastinal metastases.

This site has the poorest survival outcomes of all head and neck subsites. There is no difference in survival when surgery is used as the primary modality of treatment followed by radiotherapy or chemoradiotherapy compared to primary radiotherapy or concurrent chemoradiotherapy followed by surgery. Concurrent chemoradiotherapy appears to be the modality of choice for laryngeal preservation; however, when surgical salvage is required, there is a low cure rate and increased wound complications. Early T1 lesions without clinical or radiographic evidence of adenopathy can be treated with primary radiotherapy, but this is relatively rare for this subsite due to a high rate of adenopathy and an advanced T stage at presentation.

Surgical resection, typically in the salvage setting, involves a total laryngopharyngectomy typically with a circumferential defect or a very small strip of mucosa preserved in continuity with the cervical esophagus. A total thyroidectomy and central neck dissection (level VI) is simultaneously performed and removed en bloc with the specimen. Bilateral neck dissection of levels II to IV is indicated. Careful dissection of the central neck, and in some cases the upper mediastinum (level VII), is required to clear regional disease, and this is critical in preventing a peristomal recurrence.

Given the circumferential or near circumferential defect, reconstruction is required to prevent saliva from accumulating in the wound and to create a neopharynx. A pedicled pectoralis major flap sutured to the prevertebral fascia has been described, but advances in free flap reconstruction has popularized a number of fasciocutaneous flaps for reconstruction of this defect, namely the radial forearm flap and the anterolateral thigh free flap. When total laryngopharyngoesophagectomy is required, a gastric pull-up may be performed for the pharyngeal reconstruction.

**Larynx** Laryngeal carcinoma typical presents with a progressive voice complaint in a long-time smoker. A thorough understanding of laryngeal anatomy is critical in the proper diagnosis, staging, and treatment of laryngeal cancers. The larynx is divided into the supraglottis, glottis, and subglottis as previously described (Fig. 18-32). The larynx starts superiority at the epiglottis and ends inferiorly at the inferior border of the cricoid cartilage of the larynx span from the epiglottis superiority to the cricoid cartilage inferiorly. Laterally, it is separated from the hypopharynx by the aryepiglottic folds.

The supraglottis includes all of the laryngeal structures above the inferior half of the ventricle, and this includes the upper half of the ventricle, the false vocal cords, the arytenoids, the aryepiglottic folds, and the epiglottis. The membranes and cartilages of the larynx act as barriers to laryngeal spread: the thyroid and cricoid cartilage, conus elasticus, the quadrangular membrane, the ventricle, the hyoepiglottic ligament, thyrohyoid membrane, and cricothyroid membrane. Although the majority of tumors of the larynx are SCC, minor salivary glands, and their associated malignancies, can be found in the supraglottis and subglottis. Other rarer pathologies include granular cell

---

**Figure 18-30.** Relationship of nasopharynx, oropharynx, and hypopharynx.
tumors and laryngeal framework tumors, typically arising from the cricoid, such as chondroma and chondrosarcoma.

The larynx functions to (a) phonate, (b) protect the airway during swallowing, and (c) maintain airway patency. This is a fine balance. For instance, if the glottic aperture is enlarged and/or supraglottic structures are excised, phonation and airway protection suffer while airway patency is improved. It is therefore not surprising that patients with laryngeal tumors can present with dysphonia (hot potato voice in supraglottic tumors and hoarseness in glottic tumors), dysphagia, and airway concerns. These patients can also present with dysphagia, weight loss, referred otalgia, and a neck mass. Vocal cord fixation can be a result of a mass effect from large obstructing masses, secondary to direct extension into the paraglottic space or through direct invasion of the cricoarytenoid joint involving either the muscle or the recurrent laryngeal nerve (RLN). Although subglottic tumors represent <1% of laryngeal cancers, they can also present with vocal cord paralysis and/or airway compromise.

Direct laryngoscopy is beneficial in the assessment of laryngeal tumors to assess the local extent of tumor spread. This is particularly important in assessing vallecula and base of tongue as there can be direct extension to the oropharynx. Similarly, glottic cancers can have subglottic extension, which necessitates a wider radiation field and/or a more extensive resection.

Esophagoscopy and bronchoscopy are also recommended to assess second primary tumors. Furthermore, when a laryngectomy is planned, the direct laryngoscopy provides information about the best possible site of entry into the pharynx. Entry can be achieved through (a) a suprahyoid pharyngotomy, (b) lateral pharyngotomy (lateral to the thyroid cartilage), or (c) inferiorly through a postcricoid or hypopharyngeal pharyngotomy.

Appropriate preoperative staging with a CT scan with contrast is critical in assessing cervical lymphadenopathy and extralaryngeal spread. Erosion or invasion of the thyroid and cricoid cartilage can significantly impact outcomes and treatment as can extension into the preepiglottic or paraglottic spaces. The supraglottic and subglottic sites are lymphatic rich, and bilateral lymphadenopathy is not uncommon, whereas the glottic site has relatively poor lymphatic drainage (1%–4% regional metastasis for isolated larynx cancer). The supraglottic drains through the neurovascular bundle to the thyrohyoid membrane, mainly draining to the upper and lateral cervical nodes (levels II–IV), whereas the glottis and subglottis drain through the cricothyroid membrane and can have spread to the prelaryngeal (Delphian nodes), paratracheal, and lower cervical nodes (levels IV and VI), although in these cases we still treat levels II to IV surgically because of the significant occult nodes in this region.

The primary management of laryngeal cancer depends on a variety of factors, including tumor extent, patient comorbidities, and surgeon/center experience. In general, similar to other subsites, early-stage disease can be treated with single modality treatment (surgery or radiotherapy) while advanced stage disease is treated with at least two modalities, typically either surgery followed by radiotherapy (with or without chemotherapy) or concurrent chemoradiotherapy. Supraglottic and subglottic lesions are typically treated with primary concurrent chemoradiotherapy in an attempt to preserve the organ; however, in patients where the primary functions of the larynx are not being fulfilled preoperatively (tracheotomy– and gastrostomy tube–dependent), primary surgical management with a total laryngectomy (Fig. 18-33) can be considered. The original trials that popularized organ preservation techniques with concurrent chemoradiotherapy either excluded or had a very small sample size of large (T4) tumors.147,148 Similarly, advanced glottic cancers (T3/T4a), even when there is no evidence of nodal disease or supraglottic tumors of all stages, have superior survival outcomes when surgery is used as the primary treatment modality.149,150 This is particularly true for tumors that extend beyond the endolarynx or with cartilage destruction, for which total
laryngectomy followed by postoperative radiotherapy continues to be the standard of care. When primary chemoradiotherapy is used, surgical salvage is available if there is treatment failure or recurrent disease.

The early glottic and supraglottic lesions can be safely treated with CO\textsubscript{2} laser transoral microlaryngoscopic resection with excellent oncologic outcomes and laryngeal preservation rates.\textsuperscript{151,152} Patients with limited involvement of the arytenoid or anterior commissure are the best candidates for a good posttreatment vocal quality result with this approach. One of the benefits of this approach is that it does not burn any bridges to more invasive treatment. Often, multiple procedures are required to control the disease. Nonetheless, for early stage cancers of the glottis and the supraglottis, radiation therapy is equally as effective as surgery in controlling disease with excellent voice outcomes.

**Laryngeal Preservation Techniques** Beyond CO\textsubscript{2} laser transoral microlaryngoscopic resection for the most early of lesions, more advanced open laryngeal preservation techniques have been developed for the resection of select, moderately advanced supraglottic and glottic tumors. These techniques can be divided into vertical and horizontal partial laryngeal procedures.

Vertical partial laryngectomy (VPL) (Fig. 18-34) involves a midline thyrotomy followed by dissection of the inner perichondrium off of the thyroid cartilage with resection of the entire true cord and a portion of the false cords, followed by reconstruction with pedicle strap muscles and bipedicled outer perichondrial flaps. A temporoparietal fascial free flap has also been used to reconstruct these defects with excellent voice outcomes.\textsuperscript{153} This can be extended to include a frontal vertical VPL where the excision crosses the midline as far laterally as to leave only the posterior commissure and one functional cricoarytenoid unit. This procedure is best reserved for recurrent glottic T1/T2 lesions involving only one vocal cord (although anterior commissure involvement is not a contraindication), <5 mm sublottic extension, with a mobile cord, and no cricoid cartilage or extralaryngeal extension. This technique leads to excellent locoregional control with improvements in voice related quality of life with advanced reconstructive techniques.\textsuperscript{153}

Supraglottic and supracricoid partial laryngectomies are horizontally oriented resections. In a supraglottic laryngectomy, a laryngectomy is performed below the hyoid and includes the upper portion of the thyroid cartilage while preserving a lower portion approximately the height of the cricoid cartilage. This is reserved for lesions not involving the vocal cords, false cords, or the arytenoids. Cartilage invasion and extensive base of tongue involvement are contraindications. Most lesions amenable for resection using this procedure are typically small enough that a laser or TORS procedure is adequate for resection, and therefore this procedure is rarely performed. For T3 glottic lesions without preepiglottic space or cricoarytenoid joint involvement, a supracricoid laryngectomy with a cricohyoidopexy or cricothyroidopiglottopexy (CHEP) are options. A single cricoarytenoid unit is preserved to allow for phonation through apposition with the remnant epiglottis or base of tongue. The procedure is associated with excellent oncologic outcomes, tracheostomy decannulation rates, and swallowing function.\textsuperscript{154} Phonation is reasonable after this procedure but can be characterized as breathy and coarse. Many surgeons prefer not to decannulate patients until the patient has had a significant period of time with good oral intake to allow for pulmonary toilet given the high initial rate of aspiration with this procedure.

All partial laryngeal procedures are associated with a high risk of aspiration. Therefore, patients should have excellent pulmonary reserve through pulmonary function tests. When this is not possible, a simple measure includes whether patients can climb two flights of stairs without stopping.
Speech and Swallowing Rehabilitation

Speech and language pathology (SLP) assessment is critical in the management of patients with laryngeal and hypopharyngeal cancer. It is a critical part of the preoperative assessment and counseling and postoperative therapy. In the elderly larynx cancer population, Starmer et al demonstrated that SLP care is underutilized and is largely reserved for select patients in anticipation of total laryngectomy or after the onset of impaired airway and swallowing function. SLP care was, however, strongly associated with improved outcomes (lower rates of dysphagia, stricture, weight loss, and pneumonia).  

SLP often discusses with the patient speech rehabilitation options after total laryngectomy, which include esophageal speech, tracheoesophageal puncture, and use of an electrolarynx. Esophageal speech is produced by actively swallowing and releasing air from the esophagus, resulting in vibrations of the esophageal walls and pharynx that can then be articulated into words. This requires a very motivated patient, and unfortunately, <20% of postlaryngectomy patients develop fluent esophageal speech.

The electrolarynx is a device that creates vibratory electric type sounds when held against the neck or cheek that the patient can articulate into speech. This device is typically used in the postoperative inpatient setting, but it can also be used by patients who are not able to create esophageal speech.

The ultimate speech rehabilitation for patients with laryngectomy is a tracheoesophageal puncture (TEP) with insertion of a voice prosthesis. This prosthesis is a one-way valve that allows air from the trachea to enter the upper esophagus while preventing retrograde passage of food or saliva into the trachea. Patients who undergo placement of a tracheoesophageal puncture have a success rate of >90% in achieving functional speech. Many surgeons do not like to place a TEP at the time of the primary laryngectomy, particularly in the salvage setting after radiotherapy due to wound complication concerns. However, primary and secondary TEP patients experience similarly high complication rates, and the extent of the pharyngeal reconstruction rather than preoperative exposure to radiotherapy appear to be more important factors in selection of TEP timing. Free flap patients used their TEP more commonly for primary communication after secondary versus primary TEP.

Postoperative swallowing rehabilitation is another important task performed by SLPs. Modified barium swallows where the consistency and amount of food provided is varied to minimize aspiration can be critical particularly in the management of patients with partial laryngeal procedures. This is performed under fluoroscopy in the radiology suite to allow for the assessment of all phases of swallowing. A more limited examination in FEES utilizes the fiberoptic nasolaryngoscope to visualize the larynx during swallow and directly visualize whether there is any laryngeal penetration.

Unknown Primary Tumors

Patients with cervical nodal metastases confirmed to be carcinoma without clinical or radiologic evidence of an upper aerodigestive tract primary tumor are referred to as having carcinoma of unknown primary (CUP). CUP comprise 2% to 5% of all head and neck cancers, although the true incidence is probably lower given advances in surgical visualization and radiological imaging to identify the primary site. Recently, there has been a rise in CUP likely related to the increase in HPV-associated oropharyngeal cancer, although CUP could also be from a primary thyroid or skin malignancy. After a thorough history and physical examination including fiberoptic nasolaryngoscopy, an FNA biopsy is used to confirm carcinoma in the cervical metastases. This is preferred over an open biopsy to avoid the risk of tumor spillage, challenging revision surgery secondary to disruption of fascial planes, and increased risk of recurrence and distant metastases. If the primary is not identified on physical examination, patients should undergo a PET-CT scan. A recent systematic review of 7 studies (246 patients) demonstrates an overall sensitivity of 44% and specificity of 97% with this technique, which can often detect tumors >1 cm in size. This should be followed by thorough diagnostic operative endoscopy (nasopharyngoscopy, direct laryngoscopy, esophagoscopy, and bronchoscopy). Operative manipulation of the tissues in the upper aerodigestive tract specifically with biopsy may lead to false positive results on the PET-CT scan, and therefore PET-CT should be performed before endoscopy. Furthermore, having the PET-CT results prior to operative endoscopy allows the surgeon to focus on specific high-risk sites for biopsy, particularly as it relates to the base of tongue. When the primary site is not evident, bilateral tonsillectomies and bilateral base of tongue biopsies can be performed to try to identify the primary site. Patients in whom a primary is identified proceed to receive appropriate treatment, and if radiotherapy is part of this treatment regimen, a more limited radiation field is administered, highlighting the importance of identifying a primary site. When the primary site is not identified, primary chemoradiotherapy is advocated, treating all of the mucosal sources of the upper aerodigestive tract at risk (from nasopharynx to hypopharynx) and the cervical regional basin bilaterally. For patients with advanced neck disease (N2a or greater) or with persistent lymphadenopathy after radiation, a neck dissection may be necessary. In the preradiation setting, a neck dissection is preferred over radiotherapy for patients with N1 disease, according to the NCCN guidelines, because some of these patients will be upstaged, ECE is not accurately diagnosed on imaging alone, and because some patients without ECE and a pathologically N1 node benefit from radiation alone without chemotherapy. The additional prognostic information provided by a neck dissection can significantly impact treatment algorithms and is also associated with lower morbidity compared to postoperative neck dissection.

Nose and Paranasal Sinuses

Cancers of the nasal cavity and paranasal sinuses are exceedingly rare, and pathology in this anatomic subsite is dominated by infectious and inflammatory sources as previously discussed in the “Sinonasal Inflammatory Disease” section of this chapter. Malignant pathology at this site is often diagnosed after failed repeated treatment of suspected benign inflammatory sinonasal pathology. Concerning preoperative imaging findings (unilateral disease; extensive disease; bony, orbital or intracranial invasion) and unusual clinical features may raise concerns about malignancy, and in these cases referral to a tertiary head and neck oncology center is preferred. A concerning history is one that involves a slow progression and worsening of symptoms, which may include nasal obstruction, facial pain, headache, epistaxis, and facial numbness. Most tumors at this site present with advanced stage given the inevitable delay in diagnosis. Numbness in the V2 distribution suggests invasion of pterygopalatine fossa, and V3 distribution numbness can be an indication of extension to the infratemporal fossa and skull base invasion to foramen ovale. Proposis, epiphora, diplopia, and change in vision (typically starting with loss of color vision) are
SPECIFIC CONSIDERATIONS

all signs of advanced orbital invasion. Maxillary sinus tumors, the most common site for cancers of this site, can be prognos-ticated simply using Ohgren’s line (Fig. 18-35), an imaginary line from medial canthus to the angle of the mandible, which divides maxillary sinus into anterior-inferior and posterior-superior parts. Tumors from the anterior-inferior are more prognostically favorable.

Although the most common pathology at this site continues to be squamous cell carcinoma, a brief discussion of other histo-pathology is warranted given significant variety, prognostic, and treatment-related differences between these at this subsite. Benign pathology at this site includes inverted papilloma, hemangiomas, hemangiopericytomas, angiofibromas, minor salivary tumors, and benign fibrous histiocytomas. Fibro-osseous and osseous lesions, such as fibrous dysplasias, ossifying fibromas, osteomas, and myxomas, can also arise in this region. Additionally, encephaloceles and meningo-encephaloceles with herniation of intracranial content into the nasal cavity can present as sinonasal lesions; therefore, imaging, typically with an MRI, is warranted before biopsy of any sinonasal mass to prevent an iatrogenic CSF leak. In the evaluation of sinonasal malignant pathology, both CT and MRI are required because they provide complimentary information. MRI provides improved skull base, intracranial, and orbital invasion assessment, while CT provides better assessment of bony anatomy and invasion.

Beyond squamous cell carcinoma, the next two most common malignancies at this site include adenoid cystic carcinoma and adenocarcinoma. Other pathologies include sinonasal undiffer-enitized carcinoma (SNUC), mucosal melanoma, lymphoma, esthesioneuroblastoma (previously known as olfactory neuro-blastoma), rhabdomyosarcoma, and angiosarcoma. Unlike other head and neck cancers, metastases to the regional lymphatic basis are extremely rare, and rarely will patients require or receive primary or adjuvant treatment to the neck unless there is clinical or radiographic evidence of neck disease (approximately 15%).

The standard treatment for malignant tumors of the paranasal sinuses is driven by the primary pathology; however, for most pathology, including SCC, the standard of care includes surgical resection followed by adjuvant radiotherapy. Advances in endoscopic approaches has led to a shift in management of these tumors with minimally invasive approaches that are associated with significantly lower complication and morbidity rates with comparable oncologic outcomes. Open approaches are, however, indicated when there is tumor abutting the anterior wall of the frontal sinus, anterior extension into nasal bones, anterior maxillary wall invasion, facial skin or soft tissue invasion, dural involvement above the orbit or periorbital invasion, tumors with significant lateral temporal fossa invasion, and extension into the oral cavity, including the hard palate or the floor of the maxillary sinus. Many tumors can be treated with an endoscopic approach such as a medial maxillectomy when the tumor arises from the medial wall of the maxilla. Multidisciplinary assessment and treatment should include a skull base tumor board discussion with a head and neck oncologist/surgeon, a neurosurgeon, ophthalmologist including oculoplastics surgeons, prosthodontists, and reconstructive surgeons. Preoperative embolization within 24 hours of tumor excision can be useful for vascular tumors.

Extent of surgery and prognosis is dependent on the tumor location and extent. For tumors limited to the hard palate and lower maxillary sinus, an infrastructure maxillectomy is sufficient. A total maxillectomy without removal of the orbital floor may be warranted for more extensive tumors limited to the maxillary sinus. When the orbital periosteum is not invaded but tumor abuts this region, removal of the orbital floor with appropriate reconstruction is warranted. When there is invasion of periorbita, an orbital exenteration is warranted for most pathology. Tumors originating in the ethmoid sinuses may require excision of the cribriform plate and repair of subsequent skull base defect if the tumor originates or invades through the bony skull base. This is performed through an anterior craniofacial resection, where a neurosurgeon performs a frontal craniotomy for exposure of the anterior cranial fossa floor, while the head and neck surgeon performs a transfacial or endoscopic resection of the inferior bony and soft tissue structures. This approach often requires resection of dura and a dural repair to achieve negative margins. A less extensive surgery including a sphenoethmoidectomy or medial maxillectomy can be entertained for smaller tumors originating in the lateral nasal wall through endoscopic or open approaches.

Tumors are deemed to be unresectable if both optic nerves are involved, if there is carotid artery invasion, or if there is extensive intracranial extension. Chemotherapy has a limited application in the management of tumors at this subsite with two exceptions: rhabdomyosarcoma, which is primarily treated with chemotherapy followed by radiation therapy with surgery reserved for the salvage setting, and SNUC, where triple modal-ity treatment is required given tumor aggressiveness. Chemotherapy in this setting may help to reduce the tumor bulk and allow for orbital preservation.

**Nasopharynx**

The anatomic borders of the nasopharynx are superiorly the adenoid patch, superolaterally the fossa of Rosenmüller and the Eustachian tube orifices (torus tubarius), inferiorly the plane of the hard palate from the choana, anteriorly the posterior nasal cavity, and posteriorly the pharyngeal wall. Malignant
tumors of the nasopharynx are typically well differentiated or lymphoepithelial SCC. However, other tumors can present in this region including lymphoma, chordoma, chondroma, nasopharyngeal cyst (Tornwaldt’s cyst), angiofibroma, minor salivary gland tumor, paraganglioma, rhabdomyosarcoma, extramedullary plasmacytoma, and, rarely, sarcoma.

Unlike other head and neck cancers, the nasopharynx site has unique ethnic and geographic predilection, namely, a higher incidence in southern China, Africa, Alaska, and in Greenland Eskimos. EBV is also more commonly seen in patients with NPC, and EBV titers are helpful in following treatment response.

As previously discussed, a posterior (level V) neck mass should be considered NPC until proven otherwise. Other signs and symptoms include nasal obstruction, epistaxis, unilateral serous otitis media in an adult, and otalgia. Advanced disease can present with cranial neuropathies, particularly of the cranial nerves, which run in the cavernous sinus (CN V1, V2, III, IV, VI). Bilateral regional disease spread is common, and the lymphatic level involved include the posterior neck (level V), as well as the upper (level II) cervical and retropharyngeal nodes. Distant metastatic disease is present in 5% of patients at diagnosis, highlighting the importance of a thorough staging workup.

Staging includes a thorough physical examination using either a flexible or rigid endoscope to assess the mucosal extent of the disease. CT and MRI are complimentary as in the assessment of nasal cavity and paranasal sinus tumors with CT providing better assessment of bony invasion and the MRI providing better soft tissue delineation, skull base invasion, and perineural spread with cranial nerve enhancement. Multimodality therapy with chemoradiotherapy is superior to radiotherapy alone in the management of nasopharyngeal carcinoma.169 Recurrent tumors are treated typically with reirradiation; however, there has been recent success with surgical salvage procedures, particular in those patients in which a negative margin can be achieved.170

When resection is contemplated for recurrent nasopharyngeal carcinoma or for low-grade tumors such as some minor salivary gland tumors, a number of surgical approaches can be utilized for resection. These include endoscopic, transpalatal, transfacial via a maxillary swing procedure, and transcervical. In many cases, a combination of these techniques is required to achieve a negative margin. The transcervical approach provides the added benefit of early access and control of the carotid artery. For benign and low-grade tumors, advances in EEA have made use of the open approaches less common.

**Ear and Temporal Bone**

Temporal bone and ear tumors are rare account for <0.5% of all head and neck cancers. Subsites in this head and neck site from lateral to medial include the pinna (external ear), external auditory canal, middle ear, mastoid, and petrous portion of the temporal bone. Although the typical pathology at this site is squamous cell carcinoma, minor salivary gland tumors such as adenocarcinoma and adenoid cystic carcinoma can also present here. Given that the ear is in the high-risk region for aggressive skin cancers due to its unique exposure to ultraviolet light, cutaneous malignancies such as basal cell carcinoma and melanoma can also present here. In the pediatric population, soft tissue sarcomas, most commonly rhabdomyosarcoma, can present at this site. These tumors typically present with an advanced stage,171 and resection with clear margins and functional preservation is challenging because of the close proximity of vital structures, namely the facial nerve and the external auditory canal.172

Tumors involving the petrous apex or intracranial structures may present with headache and palsies of CN V and VI as well.

Patients can present with ulceration, granulation, or bleedings from the external ear and auditory canal. This is often mistaken for an infectious or inflammatory process given the rarity of malignancy at this subsite; however, persistent granulation tissue in the ear should be biopsied and imaged to rule out malignancy. Patients can then present with otitis media, otalgia, hearing loss, vertigo, and facial nerve paralysis. Appropriate imaging with CT and MRI is often required to appropriately delineate the lesion and stage and assist with the appropriate management plan.

Cutaneous malignancies of the pinna and tragus can usually be locally excised. However, at this subsite, spread into the perichondrium and cartilage can lead to rapid spread along the tissue plane. The importance of negative margins cannot be overstated at this subsite. Mohs microsurgery has been advocated for select tumors at this subsite for this reason; however, some tumors are so extensive that a total auriculectomy provides the best oncologic and cosmetic result. When there is extension of tumor to the bony cartilaginous EAC junction, spread to parotid, temporomandibular joint, and skull base is possible. Advanced tumors anterior to a vertical line along the EAC from a sagittal view benefit from a parotidectomy as well as a supraomohyoid neck dissection (levels I–III), whereas those behind this line benefit from a posterosilateral neck dissection (levels II–V). As with other cutaneous malignancies, adjuvant radiotherapy is indicated for positive margins, perineural spread, or multiple involved lymph nodes.

Tumors involving the EAC and middle ear require different management, including a sleeve resection of the external auditory canal, a lateral temporal bone resection, or a subtotal temporal bone resection (Fig. 18-36). A sleeve resection of the EAC skin and cartilage is rarely enough to achieve negative margins with the exception of some basal cell carcinomas of the skin overlying the cartilaginous EAC. For more extensive

![Figure 18-36.](image-url) Levels of the neck denoting lymph node bearing regions.
tumors and more aggressive pathology, a lateral temporal bone resection may be required removing the cartilaginous and bony external auditory canal as well as the middle ear en bloc. A subtotal temporal bone resection also removes the inner ear and facial nerve as part of the resection and is indicated when the tumor extends into the middle ear and a deeper resection margin is required. Both of these procedures are followed by postoperative radiotherapy, which provides improved locoregional control. The neck is managed in a similar fashion to pinna and external auditory canal malignancies typically requiring a supraomohyoid (levels I–III) neck dissection. Survival outcomes are poor with a 5-year overall survival of <40%. Important predictors of disease free survival include margin status, perineural invasion, and regional lymphatic spread; the most important of these on multivariate analysis being lymphatic spread of disease.

Lateral temporal bone resections often require reconstruction to close the wound, provide bulk, and vascularize tissue. If dura is encountered and even resected, a watertight dural closure is required to prevent a CSF leak and meningitis. Vascularized tissue has the added benefit of preparing the surgical bed for postoperative radiotherapy. These defects can be reconstructed with regional pedicled flaps (e.g., submental) or free flaps. The most common free flaps used are the anterolateral thigh, although depending on body habitus and the depth of the defect, the radial forearm, lateral arm, and rectus abdominus may also be used. The deformity resulting from a total auriculotomy is often not reconstructed primarily, but an auricular prosthesis can be designed for further rehabilitation. Facial nerve reconstruction when sacrifice is required is typically performed with cable grafts from the proximal facial nerve to select distal facial nerve branches. Because of the long distance between the proximal and distal branches, facial movement is typically delayed 6 to 12 months. However, if the masseteric nerve is connected through a cable graft to select distal facial nerve branches (typically the zygomatic branch), a shorter cable graft is required, and facial movement can be achieved earlier. A variety of other static and dynamic procedures can be performed secondarily. The most important of these procedures are related to preserving eye closure to avoid corneal abrasions or desiccation, which can ultimately lead to blindness. In the immediate postoperative period, taping of the eyelids and generous application of eye lubrication is required to prevent exposure keratitis. Upper lid gold weight implants, lower lid shortening procedures, and tarsorrhaphy can be performed secondarily to assist with eye closure.

**Neck**

An undiagnosed neck mass needs to be carefully evaluated and worked up so as to not interfere with the definitive management of the patient and future treatment options. The patient’s age, social history, including alcohol and smoking history, preceding illness history, and synchronous upper aerodigestive tract physical examination findings can significantly impact the differential diagnosis and the investigation to work up a neck mass. A thorough history and head and neck examination, including fiberoptic nasolaryngoscopy, are therefore paramount to complete evaluation. With regard to age, in children, a neck mass is far more likely to be congenital, inflammatory, or infectious, whereas in adults, neck masses >2 cm have a >80% probability of being malignant. Typically, the first investigation is an FNA biopsy, which can be performed with ultrasound or CT guidance when the mass is not easily palpable or largely cystic with a small solid component. Imaging is critical in characterizing the neck mass, particularly assessing the borders, consistency, and location which then impacts the differential diagnosis. For instance, a cystic neck mass can be a branchial cleft cyst or a regional metastasis from an oropharynx cancer or metastatic papillary thyroid cancer. Therefore, the imaging findings also significantly impact the differential diagnosis.

When the imaging and FNA does not provide adequate information for a diagnosis, a core biopsy can be considered, particularly if the diagnosis of lymphoma is suspected and an open biopsy wants to be avoided. For a suspected carcinoma, an open biopsy may be required; however, in that case, the incision needs to be planned such that the procedure can be converted to a neck dissection, and a frozen section can be sent. If the diagnosis of squamous cell carcinoma is confirmed on frozen section, then a neck dissection should be performed to further prognosticate the disease. In the case of lymphoma, biopsy does not need to remove the entire lymphoma, particularly if there is an added risk of injuring normal anatomical structures.

**Patterns of Lymph Node Metastasis.** The lymphatic drainage into the neck is divided into seven levels with standardized reporting within and across specialties, particularly as radiologists, pathologists, surgeons, radiation oncologists, and radiologists share the findings (Fig. 18-37). The levels include:

- **Level I**—the submental and submandibular nodes
- **Level Ia**—the submental nodes; medial to the anterior belly of the digastric muscle bilaterally, symphysis of mandible superiorly, and hyoid inferiorly; this level does not have any laterality as it includes both right and left sides
- **Level Ib**—the submandibular nodes and gland; posterior to the anterior belly of digastric, anterior to the posterior belly of digastric, and inferior to the body of the mandible

![Figure 18-37. Shaded region indicates the region included in a supraomohyoid neck dissection.](image-url)
There is a well-established pattern of regional spread from upper aerodigestive tract primary tumors. Lesions of the lip and oral cavity typically metastasize to levels I to III and skip metastasizes to the lower basin (levels III–IV) without involvement of the upper level (levels I–II). Oropharyngeal, laryngeal, and hypopharyngeal tumors most commonly spread to the lateral neck (levels II–IV). It is rare for any of these tumors to have isolated regional metastases to level V; however, nasopharyngeal, thyroid, and head and neck malignant melanoma can metastasize to this level. Other sites for metastasis include the retropharyngeal nodes (oropharyngeal, nasopharyngeal, and hypopharyngeal tumors), paratracheal and level VII nodes (thyroid, hypopharynx, and cervical esophageal tumors), and pretracheal (Delphian) nodes (thyroid and advanced glottic tumors with subglottic extension).

Historically, a radical neck dissection (RND) was performed for all upper aerodigestive tract malignancies with sacrifice of the SCM, internal jugular vein (IJV), and accessory nerve (CN XI) and removal of all lymphatic level (levels I–V). This was because cervical metastasis decreased the 5-year overall survival rate by approximately 50%. However, growing evidence demonstrated that this was not necessary, and now a neck dissection is only recommended for upper aerodigestive tract malignancies when the risk of occult disease is >20% in the clinically negative neck. When the neck is clinically positive, the level discussed in the previous paragraph for each site are excised with every attempt to preserve the SCM, IJV, and CN XI (selective neck dissection; SND). When there is direct extension of the tumor or extralymphatic spread into these structures, sacrifice may be necessary in a modified radical neck dissection (MRND). The RND has been largely abandoned because the SND and MRND have been demonstrated to be equally effective when it comes to oncologic outcomes with far improved functional outcomes.  

SND has become the standard of care for most patients who are clinically node negative (cN0) and in those with limited nodal extension (cN1) disease. Patients with oral cavity cancer typically receive a supraomohyoid (Fig. 18-38) neck dissection (levels I–III). Many surgeons will include a portion of level IV just below the omohyoid muscle given the rate of skip metastases previously discussed. Approximately 80% of patients with oral cavity cancer present cN0; however, the rate of occult metastatic disease is approximately 30% and differs by subsite. This rate is further impacted by tumor thickness at the tongue subsite, with tumors 4 mm or thicker having a higher rate of occult disease. A recent prospective, randomized trial demonstrated the oncologic benefit of an elective neck dissection in cN0 oral cavity patients regardless of tumor thickness over an observation followed by therapeutic neck dissection in those with regional failures. An additional role of SND is as a staging tool to determine the need for postoperative radiation therapy. The lateral (Fig. 18-39) neck dissection (levels II–IV) is typically used in laryngeal and hypopharyngeal cancers. The posterolateral (Fig. 18-40) neck dissection (levels II–V) is typically recommended in thyroid cancers, although recent evidence has demonstrated that a partial level V dissection may be all that is necessary for equivalent outcomes to a full level II to V neck dissection. 

Despite advances in the surgical management of neck disease, in clinically advanced nodal disease (with the exception of uncomplicated N1 disease), an MRND remains the treatment of choice. When the neck disease is advanced with extranodal extension (ENE), perineural invasion (PNI), lymphovascular invasion (LVI), and the presence of multiple involved nodes, postoperative radiotherapy improves locoregional control. If there is a positive margin or ENE, then the addition of adjuvant chemotherapy to radiotherapy provides a survival benefit. In patients receiving primary radiotherapy with advanced N stage disease (N2a or greater) or only a partial response to
treatment, a planned postradiotherapy neck dissection can be performed 6 to 8 weeks after completion of radiotherapy. This is to consolidate the treatment and provide prognostic information.

Tumor factors that preclude surgery include prevertebral fascia invasion, skull base invasion, and >270° circumferential encasement of the internal carotid artery. These factors are associated with very poor 5-year survival (<20%). In such cases, sacrifice of the carotid is not indicated given the risk of stroke and death. Surgical debulking is also not associated with improved survival. However, there is a role for neoadjuvant chemotherapy, and in those that respond and if the disease becomes resectable, survival benefit has been demonstrated.189 Recurrent neck metastasis after radiotherapy to the neck or a comprehensive neck dissection is associated with very poor survival.190

Parapharyngeal Space Masses. The parapharyngeal space is a potential inverted pyramidal space bordered superiorly at the skull base along the sphenoid and inferiorly at the greater cornu of the hyoid. Medially it is bordered by the buccopharyngeal fascia covering the superior constrictor, anteriorly the pterygomandibular raphe, posteriorly the prevertebral fascia, and laterally by the deep surface of the parotid gland and ramus of the mandible. The differential diagnosis for parapharyngeal masses is very much dependent on the anatomy and contents of this space which is divided into the pre- and poststyloid spaces by the tensor-styloid fascia. This fascia attaches the tensor veli palatini muscle to the styloid. The contents of the prestyloid parapharyngeal space include fat, the deep lobe of the parotid, and lymph nodes, and branches of V3 (lingual, inferior alveolus, and auriculotemporal nerves), whereas the contents of the poststyloid space including cranial nerves IX to XII, the internal jugular vein, the internal carotid artery, and the sympathetic chain. Nearly half of all parapharyngeal masses are of parotid origin, while 20% to 25% are of neurogenic origin, such as paragangliomas (glomus vagale, carotid body tumor), schwannomas, and neurofibromas. Lymphatic origin masses such as lymphoma and lymph node metastases represent 15% of tumors at this subsite. Therefore, most prestyloid lesions are considered of salivary gland origin, whereas poststyloid lesions are typically vascular or neurogenic.

Tumors of the parapharyngeal space can displace the lateral pharyngeal wall medially into the oropharynx (Fig. 18-41) and can thus cause obstructive sleep apnea, voice change, and dysphagia in addition to cranial neuropathies, Horner’s syndrome, or vascular compression. In addition to CT and MRI, poststyloid lesions should be investigated with a 24-hour urinary catecholamine collection because some paragangliomas are functional and this should be managed preoperatively.

Surgical access to these tumors can be performed using a purely transcervical approach with the excision of the submandibular gland for access. A transfacial or transparotid approach can be used as an adjunct for certain tumors by removing the parotid gland. This ensures identification of the facial nerve.

**Figure 18-39.** Shaded region indicates the region included in a posterolateral neck dissection.

**Figure 18-40.** Parapharyngeal mass—prestyloid with prominent oropharyngeal presentation typical of a dumbbell tumor.
prior to removal of the mass, which is just deep to it. Rarely, a transmandibular approach is required by performing a midline or parasymphyseal mandibulotomy with a lateral swing. Transoral approaches have been described, but they are not recommended and are largely contraindicated due to poor exposure and control of the associated vasculature.

Benign Neck Masses. Many benign neck masses require surgical intervention for diagnostic, cosmetic, and symptomatic relief. This is particularly true for lesions that are prone to recurrent infections, especially in the pediatric population. Such masses include thyroglossal duct cyst, branchial cleft cyst, lymphangioma (cystic hygroma), hemangioma, and dermoid cyst. Lymphangioma and hemangioma were previously discussed and will not be discussed in this section.

During fetal growth, the thyroid gland descends along a tract from the foramen cecum at the base of tongue into the anterior low neck. A vestigial remainder of this tract is called a thyroglossal duct cyst, which typically presents as a subcutaneous swelling near the hyoid in the midline or slightly paramedian. Patients may complain of recurrent infections of this mass after an upper respiratory tract infection. Investigations include thyroid function tests and a neck and thyroid ultrasound to confirm that the patient has thyroid tissue in the lower neck. Treatment involves removal of the cyst, the tract, and the central portion of the hyoid (Sistrunk procedure), often with a small portion of the base of tongue if the tract extends above the hyoid.

During fetal growth, the branchial cleft apparatus may persist, forming a branchial cleft remnant (cyst, sinus, or tract), numbered to their corresponding embryologic branchial cleft. First branchial cleft anomalies parallel the EAC (Work Type I; preauricular) or go through the parotid gland ending at the bony-cartilaginous EAC junction (Work Type II; angle of the mandible). Second branchial anomalies (Fig. 18–42), the most common type, start at the anterior border of the SCM and head toward the tonsillar fossa traveling deep to second arch structures (CN VII and external carotid artery) and superficial to third arch structures (stylopharyngeus, IX, and internal carotid artery). Third and fourth branchial anomalies are difficult to distinguish clinically and frequently open into the pyriform sinus often presenting with recurrent thyroid infections. These anomalies ascend posterior the internal carotid artery and deep to CN IX but superficial to CN XI and XII. Dermoid cysts tend to present as midline masses and represent trapped epithelium originating from the embryonic closure of the midline. These can be reliably diagnosed and distinguished from thyroglossal duct cysts using an ultrasound predictive model.

Cervical Fascial Planes. The fascial planes often predict the pathway and extent of infectious spread in the neck and are therefore clinically important. The deep fascial layers of the neck

![Figure 18-41](image1.png)

Figure 18-41. Computed tomography scan demonstrating a branchial cleft cyst with operative specimen.

![Figure 18-42](image2.png)

Figure 18-42. Example of a tumor in the parotid with the pattern of the facial nerve and associated anatomy. m. = muscle; n. = nerve; v. = vein.
include three separate layers: the superficial deep (investing) layer, the pretracheal (visceral) layer, and the prevertebral layer. The investing layer forms a cone around the neck and surrounds the SCM muscle and the anterior and posterior neck. It spans from the mandible to the clavicle and manubrium. The visceral layer surrounds the trachea, thyroid, and esophagus and blends laterally with the carotid sheath extending inferiorly to the upper mediastinum. Between this layer and the prevertebral fascia is the retropharyngeal space. The prevertebral fascia covers the prevertebral musculature and space and extends down to the thoracic vertebral and diaphragm. Infections of the prevertebral space between this fascia and the prevertebral musculature are considered to be in the prevertebral space and can extend all the way down to the sacrum. Therefore, neck infections can extend to the mediastinum or beyond and need to be treated aggressively.

Salivary Gland Tumors

Primary malignant tumors of the salivary glands are relatively rare and account for <2% of all head and neck malignancies. As previously mentioned, minor salivary gland malignancies can present anywhere in the upper aerodigestive tract, particularly on the palate; however, the major salivary glands are the parotid, submandibular, and sublingual glands. The majority of tumors (80%) arise in the parotid gland (Fig. 18-44); however, 80% of these are benign, most commonly, pleomorphic adenomas (benign mixed tumors). As the salivary gland gets smaller, the proportion of tumors that are malignant increases; 50% of submandibular/sublingual tumors and 80% of minor salivary gland tumors are malignant.

Patients typically present with a mass because these tumors are well circumscribed and slow growing. However, certain signs and symptoms, such as pain, paresthesia, facial nerve weakness, or rapid growth, raise the concern for malignancy. If there is facial nerve weakness (10%–15% of cases), this usually represents tumor invading the facial nerve. Submandibular and sublingual tumors present with a mass or swelling in the neck or floor of the mouth, respectively. Tumors in this region can invade the lingual nerve leading to tongue paresthesia or the hypoglossal nerve invasion leading to paralysis. The close proximity to the mandible and tongue necessitates a thorough bimanual palpation to assess for fixation to these structures.

The decision to dissect the neck in parotid cancers is fraught with uncertainty. However, parotid malignancies, particularly carcinomas, have a propensity for regional lymphatic spread, first to the intra- and periglandular nodes followed by the upper cervical chain (levels I–III). Occult nodal metastases are present in 30% of cases and are predicted by intra- or periglandular nodes, high-risk histology (high histological grade), and extraparotid extension. Patients with advanced tumor stage (T3/T4a), perineural invasion, high risk histology, or clinically involved adenopathy should have their neck dissected. Submandibular gland cancers metastasize to the submental (Ia) and submandibular triangle lymph nodes followed by the upper cervical chain (levels II–III). Extraglandular extension and regional metastases are poor prognostic factors.

Following a thorough history and physical examination, an FNA biopsy should be performed to provide an accurate preoperative diagnosis in 70% to 80% of cases when reviewed by an experienced cytopathologist. If the biopsy is nondiagnostic, a repeat biopsy should be performed under image-guidance, typically with an ultrasound. An open or incisional biopsy should be avoided because of the risk of tumor spillage and cutaneous spread. Also, this approach is fraught with risk to the facial nerve. Salivary gland tumors are worked up with appropriate imaging, typically with an MRI because of the increased soft tissue definition. FNA and imaging results are critical in guiding the surgeon to the extent of surgery. The minimal extent of surgery for salivary gland tumors is a superficial parotidectomy, removing all of the salivary gland tissue superficial to CN VII, which is meticulously dissected during this procedure.

The final histopathologic diagnosis in salivary gland tumors can be challenging. Nonetheless, there is a well-outlined histological classification used by pathologists. Benign and malignant tumors of the salivary glands are divided into epithelial, nonepithelial, and metastatic neoplasms. Benign epithelial tumors are most commonly pleomorphic adenoma (85%), monomorphic adenoma, Warthin’s tumor (papillary cystadenoma lymphomatosum), oncocytoma, or sebaceous neoplasm. Nonepithelial benign lesions include lipoma and hemangioma. Treatment of benign neoplasms is surgical excision for diagnostic and therapeutic purposes. The parotid superficial lobe is usually dissected off of the facial nerve, which is preserved. For pleomorphic adenoma, an extracapsular dissection is favored over enucleation due to tumor pseudopods, incomplete excision, and a higher risk of tumor spillage, all of which are associated with higher recurrence rates. Recurrence is associated with a high degree of morbidity.

Malignant epithelial tumors range in aggressiveness based on tumor histology, grade, perineural invasion, and regional metastases. Mucoepidermoid carcinoma is the most common primary malignancy of the salivary glands and can be high grade (more epidermoid) or low grade (more mucinous). High grade mucoepidermoid carcinoma can be hard to differentiated from squamous cell carcinoma, particularly on FNA. Adenoid cystic is the second most common primary salivary gland malignancy and has three histological subtypes: tubular, cribriform, and solid. Higher grade/risk tumors have a higher degree of solid differentiation. Adenoid cystic cancers are known for perineural invasion and late recurrences and distant metastases. Carcinoma ex pleomorphic adenoma is an aggressive malignancy that arises from a preexisting benign mixed tumor highlighting the importance of removing these benign masses before malignant transformation.

Surgical excision remains the standard of care, typically with facial nerve preservation unless the nerve is directly invaded by tumor. For tumors that extend beyond the superficial lobe, nerve branches can be spared, and a total parotid can be performed by removing parotid tissue deep to the nerve while preserving the integrity and function of the nerve. Whenever possible, the nerve is preserved even if microscopic disease is left on the nerve, so long as gross tumor is not left behind (i.e., the nerve is not encased). If this is not possible or if the nerve is not working preoperatively, nerve sacrifice is usually recommended.

Elective neck dissection is warranted in high-grade mucoepidermoid carcinomas and other high-risk pathology and grade where the risk of occult disease is greater than 15% to 20%. Therapeutic neck dissection is recommended in patients with clinically or radiographically evident disease. Postoperative radiotherapy is indicated in patients with perineural invasion, advanced local disease (T4a), extraglandular disease including regional metastases, and high-grade histology.
Local Flaps and Skin Grafts

Local flaps are commonly used for cutaneous reconstruction in the head and neck. Local flaps are most commonly utilized for reconstruction after Mohs micrographic surgery for cutaneous malignancy, or for reconstruction of melanoma defects. Skin grafts are also commonly used for reconstruction of scalp defects after surgical resection of cutaneous malignancies. Skin grafts may also be utilized in the oral cavity for resurfacing of superficial defects of the tongue, floor of mouth, and buccal mucosa.

Regional Flaps

Three regional flaps deserve mention as potential flaps for head and neck reconstruction. The first is the pectoralis major myocutaneous flap, based upon the thoracoacromial artery.196 This flap may be used as a primary option for hypopharyngeal reconstruction and/or parotid and temporal bone reconstruction. Care must be taken during the neck dissection in order to preserve the submental vessels that supply this flap. Finally, the supraclavicular flap is based upon the suprascapular artery, arising from the transverse cervical artery.198 This is a thin, fasciocutaneous flap that is commonly used for external neck and facial reconstruction in which thin tissue is desired.

Free Tissue Transfer

The majority of major defects of the head and neck require free tissue transfer for optimal reconstruction.199 A full discussion of head and neck reconstructive microsurgery is beyond the scope of this chapter; however, a brief overview of free tissue transfer is provided in this section. Free tissue transfer allows the surgeon to transplant tissue from a wide array of donor sites, each of which have distinct advantages.200 For example, for floor of mouth reconstruction, where thin tissue is desired, the surgeon may select the radial forearm as the donor site. On the other hand, when presented with a total glossectomy defect, where thick tissue is desired for adequate volume reconstruction, the rectus may be the optimal donor site. Considering osseous defects, for reconstruction of a segmental mandible defect with minimal soft tissue deficit, the fibula osseocutaneous free tissue transfer may be the optimal choice.201 On the other hand, reconstruction of an osseous mandible defect with a large mucosal and external soft tissue deficit may be best served by the scapula donor site, where vascularized bone can be combined with a large skin paddle, and an additional latissimus dorsi myocutaneous free tissue transfer, if needed.202 The ability to harvest tissue from multiple donor sites is critical to obtaining the optimal reconstructive result. Table 18-6 lists the commonly utilized donor sites and their reconstructive advantages and disadvantages.

<table>
<thead>
<tr>
<th>FLAP</th>
<th>BLOOD SUPPLY</th>
<th>CHARACTERISTICS</th>
<th>COMMON DEFECTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radial forearm</td>
<td>Radial artery</td>
<td>Thin, pliable, long pedicle</td>
<td>Partial and hemiglossectomy, floor of mouth, buccal defects</td>
</tr>
<tr>
<td>Anterolateral thigh</td>
<td>Descending branch</td>
<td>Thicker adipose than radial forearm, can have myocutaneous (most common) or</td>
<td>Hypopharynx, external neck/facial skin, extended hemiglossectomy/total</td>
</tr>
<tr>
<td></td>
<td>of lateral femoral</td>
<td>septocutaneous perforators</td>
<td>glossectomy</td>
</tr>
<tr>
<td></td>
<td>circumflex artery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral arm</td>
<td>Posterior radial</td>
<td>Outstanding color match for facial skin, resists ptosis, diminutive pedicle</td>
<td>Parotid, temporal bone, external face and neck skin</td>
</tr>
<tr>
<td></td>
<td>collateral artery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rectus</td>
<td>Deep inferior</td>
<td>Thick adipose tissue for large volume defects, long pedicle, poor external skin</td>
<td>Total glossectomy, skull base</td>
</tr>
<tr>
<td></td>
<td>epigastric artery</td>
<td>color match</td>
<td></td>
</tr>
<tr>
<td>Latissimus dorsi</td>
<td>Thoracodorsal artery</td>
<td>Large surface area of muscle, requires semi-lateral position, can be difficult</td>
<td>Extensive scalp and skull base defects</td>
</tr>
<tr>
<td>Fibula osseocutaneous</td>
<td>Peroneal artery</td>
<td>Excellent bone stock and length, long pedicle, thin skin paddle</td>
<td>Segmental mandible and maxilla</td>
</tr>
<tr>
<td>Scapula osseocutaneous</td>
<td>Circumflex scapular artery</td>
<td>Less bone length compared to fibula, large scapular or parascapular skin</td>
<td>Segmental mandible and maxilla defects with extensive soft tissue components</td>
</tr>
<tr>
<td>Radial forearm osseocutaneous</td>
<td>Radial artery</td>
<td>Long pedicle, diminutive bone stock</td>
<td>Partial mandible defects, orbit</td>
</tr>
<tr>
<td>Iliac crest</td>
<td>Deep circumflex iliac artery</td>
<td>Up to 16 cm of bone available, limited soft tissue, significant donor site</td>
<td>Segmental mandible defects with small intraoral component and large external skin component</td>
</tr>
</tbody>
</table>

Table 18-6

Free tissue transfer donor sites for head and neck reconstruction
Figure 18-43 shows a prototypical hemiglossectomy defect from a T2 N0 oral tongue cancer that was reconstructed with a rectangle template radial forearm free tissue transfer. The radial forearm free tissue transfer provides thin, pliable tissue, with a long pedicle, and is a staple for hemiglossectomy and partial glossectomy reconstruction.

Figure 18-44 shows a composite mandible defect from a T4a N0 mandibular alveolus cancer, after segmental mandibulectomy, reconstructed with a fibula osseocutaneous free tissue transfer. The 2.5-mm titanium reconstruction plate was bent to a mandible model. A template of the osseous defect is made and transferred to the fibula, and wedge ostectomies are made in the bone so that it can be snug fit into the bone defect.

Figure 18-45 shows a palate defect after an infrastructure maxillectomy for a T2 N0 maxillary alveolus cancer. The defect resulted in direct communication with the buccal space, nasal cavity, and maxillary sinus. A radial forearm free tissue transfer was utilized to achieve oronasal separation.

TRACHEOTOMY

Indications and Timing
The most common cause for tracheotomy is prolonged intubation typically in critically ill intensive care unit patients. Prolonged intubation increases the risk of laryngeal and subglottic injury, which may lead to stenosis. In the critically ill patient, it has been hypothesized that early tracheotomy may improve inpatient survival and decreased intensive care unit length of stay while increasing patient comfort. However, a large randomized clinical trial demonstrated no benefit from early tracheotomy on short- or long-term survival and other important secondary outcomes. Furthermore, clinicians are poor predictors of which patients require extended ventilatory support. Another study demonstrated no evidence that early tracheostomy reduced mortality, duration of mechanical ventilation, intensive care unit stay, or ventilatory associated pneumonia. It did, however, provide a shorter duration of sedation. Beyond prolonged intubation, tracheotomy is also indicated in patients who require frequent pulmonary toilet, in patients with neurologic deficits that impair protective airway reflexes, and in head and neck upper aerodigestive tract surgery as a temporary airway in the perioperative period to bypass airway obstruction.

Technique and Complications
The procedure can be performed using an open or a percutaneous technique. Complications of tracheostomy include pneumothorax, tracheal stenosis, wound infection/stomatitis with large-vessel erosion, and failure to close after decannulation. A meta-analysis of 15 randomized studies assessing nearly 1000 patients demonstrated no difference between the open and percutaneous techniques, although there was a trend toward fewer complications in the percutaneous approach. The percutaneous approach was also found to be cheaper and had the added benefit of being performed at the bedside outside of the operating room. A Cochrane review on the topic lower wound infection/stomatitis and unfavorable scarring rates with the percutaneous approach. Mortality and serious adverse events did not differ between the two techniques.

The use of cricothyroidotomy, typically in the emergency setting, is inferior to a tracheotomy due to higher incidence of vocal cord dysfunction and subglottic stenosis. Therefore, soon after a cricothyroidotomy is performed, a formal
tracheotomy should be used with decannulation of the cricothyroidotomy site. Most tracheostomies are not permanent and can be reversed simply by removing the tube and applying a pressure dressing. The stoma usually spontaneously heals within 2 to 3 weeks.

**Speech with Tracheotomy and Decannulation**

When a large cuffed tracheostomy is initially placed, speech is not possible, particularly when the cuff is up. However, when the tube is downsized to a cuffless tracheostomy tube,
intermittent finger occlusion or placement of Passy-Muir valve can allow the patient to voice while still bypassing the upper airway obstruction in inspiration. Prior to decannulation, the patient has to tolerate capping for 24 to 48 hours, but this period can be extended in patients with concerns for pulmonary toilet and an inability to clear secretions.

**LONG TERM MANAGEMENT AND REHABILITATION**

**Palliative Care**

For patients with unresectable disease (greater than 180° of encasement around the carotid artery, prevertebral fascia invasion, and skull base invasion) or distant metastases, palliative care options exist. The NCCN guidelines recommend clinical trials for patients in this category because there is not a single accepted regimen for patients with incurable disease but the goal of treatment is to control symptoms and maintain quality of life while minimizing the side effects of treatment. This may include a combination of radiotherapy, usually in a hypofractionated pattern with high dose per fraction regimen, chemotherapy, or simply pain management. A recent trial demonstrated the utility of immunotherapy, specifically, Nivolumab, in the management of recurrent unresectable head and neck cancer, showing a higher response rate (13.3%) compared to standard therapy (5.8%) with lower treatment-related adverse events (13.1% vs. 35.1%, respectively). From a surgical perspective, some patients require tracheostomy or gastrostomy tube placement to manage airway compromise and dysphagia, respectively. Palliative care facilities and hospice care allow patients to retain dignity when they have a limited short-term outlook.

**Follow-Up Care**

Patients diagnosed and treated for a head and neck tumor require follow-up care aimed at monitoring for recurrence and the side effects of therapy. The NCCN guidelines recommend follow-up assessment every 3 months for the first year after treatment, every 4 months during the following year, and then every 6 months until year 4, with an annual follow-up at 5 years post treatment and thereafter. This regimen is not well followed in North America, and further investigation is required to assess why this might be and to improve adherence rates. Follow-up should consist of a thorough history to assess for any emerging symptoms such as pain, otalgia, or dysphagia as these are often the first sign of a recurrence. Assessment by speech language pathology and a dietician is often beneficial to ascertain swallowing function and nutritional intake, respectively. Some patients require dilation or reinsertion of a gastrostomy tube if they develop pharyngeal strictures and are unable to maintain their weight. The history should be followed with a thorough head and neck examination, including fiberoptic nasolaryngoscopy, because of the significant risk of developing a second primary in the upper aerodigestive tract. Patients should have their thyroid stimulating hormone (TSH) checked once a year, especially in those that have radiation as they may develop hypothyroidism at an earlier age than the general population. Shoulder dysfunction after neck dissection with extensive accessory nerve dissection or in patients who have had a scapular system free flap should be managed with physiotherapy to minimize the long-term effects and improve function. Chronic pain can occur in head and neck cancer patients, and this is often assessed and managed by a pain specialist. Ongoing dental evaluation is needed in some patients to treat caries and prevent osteoradionecrosis.

**REFERENCES**

Entries highlighted in bright blue are key references.


206. Szakmany T, Russell P, Wilkes AR, Hall JE. Effect of early tracheostomy on resource utilization and clinical outcomes in...


TRACHEA

Anatomy
The trachea is composed of cartilaginous and membranous portions, beginning with the cricoid cartilage, the first complete cartilaginous ring of the airway. The cricoid cartilage consists of an anterior arch and a posterior broad-based plate. Articulating with the posterior cricoid plate are the arytenoid cartilages. The vocal cords originate from the arytenoid cartilages and then attach to the thyroid cartilage. The subglottic space, the narrowest part of the trachea with an internal diameter of approximately 2 cm, begins at the inferior surface of the vocal cords and extends to the first tracheal ring. The remainder of the distal trachea is 10.0 to 13.0 cm long, consists of 18 to 22 rings, and has an internal diameter of 2.3 cm (Fig. 19-1).

Bronchoscopically, the tracheal rings are visible as C-shaped hyaline cartilaginous structures that provide rigidity to the anterior and lateral tracheal walls. The open ends of the C-rings are connected by the trachealis smooth muscle and encased in a dense band of connective tissue called perichondrium. The first tracheal ring is attached directly to the cricoid cartilage; there are approximately two rings for every 1 cm of tracheal length.

The tracheal blood supply, which includes the inferior thyroid, subclavian, supreme intercostal, internal thoracic, innominate, and superior and middle bronchial arteries, enters the airway near the junction of the membranous and cartilaginous portions (Fig. 19-2). Each arterial branch supplies a segment of 1.0 to 2.0 cm, thereby limiting circumferential mobilization to that same distance. The vessels are interconnected along the lateral surface of the trachea by an important longitudinal vascular anastomosis that feeds transverse segmental vessels to the soft tissues between the cartilages.

Tracheal Injury
Tracheal injury can result from a variety of causes, including inhalation of smoke or toxic fumes, aspiration of liquids or solid objects, endotracheal intubation, blunt and penetrating trauma, and iatrogenic injury during operative procedures. Early diagnosis is critical to avoid subsequent complications, including respiratory infection and tracheal stenosis. Management of smoke or toxic fume inhalation and liquid aspiration is commonly supportive; use of antibiotics, respiratory support, and airway clearance with flexible bronchoscopy is dictated by the patient’s condition. In rare circumstances, extracorporeal membrane oxygenation is required if there is associated injury to the more distal airways and lung parenchyma.

Despite ubiquitous use of high-volume–low-pressure cuffs, overinflation of the endotracheal cuff is the most common cause of injury secondary to endotracheal intubation. High cuff pressures can cause ischemia of the contiguous airway wall in as short as 4 hours. Prolonged overinflation can lead to scarring...
Key Points

1. Lung cancer continues to be a highly lethal and extremely common cancer, with 57% of patients presenting with distant metastasis and 5-year survival of 18%. Lung cancer incidence is second only to the incidence of prostate cancer in men and breast cancer in women, with 222,500 estimated new cases in 2017. Squamous cell carcinoma and adenocarcinoma of the lung are the most common subtypes and are rarely found in the absence of a smoking history. Nonsmokers who live with smokers have a 24% increased risk of lung cancer compared to nonsmokers who do not live with smokers.

2. A multidisciplinary approach to evaluation of NSCLC, with standardized criteria and terminology for diagnosis in cytologic and small biopsy specimens, and routine molecular testing for known mutations, such as EGFR mutations and EML4-ALK fusion oncogenes is now recommended for the evaluation and management of lung nodules due to major advances in targeted therapy. Adequate tissue acquisition at the time of diagnostic workup is critical and facilitates patient care while minimizing the number of procedures to which the patient is subjected.

3. The terms bronchioloalveolar carcinoma and mixed subtype adenocarcinoma have been eliminated from the classification of lung adenocarcinoma as a result of increased understanding of important clinical, radiologic, pathologic, and genetic differences between mucinous and nonmucinous adenocarcinomas. The classification system delineates a stepwise pathologic progression, from AAH to nonmucinous adenocarcinomas. The classification system includes many tumor cell types, including large cell, squamous cell, and adenocarcinoma. The approach to diagnosis and management and the terminology used in describing these tumors are evolving rapidly. In particular, the evaluation and management of adenocarcinoma of the lung has shifted dramatically and firm establishment of NSCLC cell type prior to chemotherapy for advanced stage lung cancer is essential.

4. The U.S. Preventive Services Task Force now recommends annual screening for lung cancer with low-dose computed tomography screening in high risk patients. Annual screening averted 14% of lung cancer deaths when applied to a population of asymptomatic adults age 55 to 80 years who have a 30 pack-year smoking history and are either currently smoking or have quit within the past 15 years. Patients should be healthy enough to tolerate curative treatment, specifically surgery per guidelines, and screening should be discontinued once the patient has not smoked for 15 years or develops a life-limiting health condition, becomes unable to tolerate lung surgery, or is unwilling to undergo curative lung resection. With this approach, it is expected that 50% of diagnosed cancers will be early stage. Screening of patients age 50 years or older with a 20 pack-year or greater history and additional risk factors (as determined by the Tammemagi lung cancer risk calculator or other validated risk scores) that increase the risk of lung cancer to 1.3% or greater should also be considered as part of lung cancer screening programs. In all cases, patient–physician shared decision-making should be undertaken, with a discussion of the risks and benefits of screening.

5. Assessment of patient risk before thoracic resection is based on clinical judgment and systematic assessment of cardiopulmonary status using established algorithms.

6. Maximum oxygen consumption ($$\text{VO}_{2\text{max}}$$) values provide important additional information in those patients with severely impaired DLco and forced expiratory volume in 1 second. Values of <10 mL/kg per minute generally prohibit any major pulmonary resection because the mortality in patients with these levels is 26% compared with only 8.3% in patients whose $$\text{VO}_{2\text{max}}$$ is ≥10 mL/kg per minute; values of >15 mL/kg per minute generally indicate the patient’s ability to tolerate pneumonectomy.

7. Tumor ablative strategies are viable alternatives to surgical resection for early stage lung cancer in inoperable patients. While premature, ablative techniques may ultimately be shown to have efficacy equivalent to lobectomy for the primary treatment of very small peripheral early-stage lung cancers and become primary therapy, even in operable patients, although limited resection with wedge (at least 2 cm margin and at least 1:1 tumor/margin ratio) and segmentectomy provide better margins of treatment and nodal sampling ensures occult nodal metastasis are identified. Multidisciplinary collaboration among thoracic surgery, interventional radiology/pulmonology, and radiation oncology is required to ensure that development of these ablative techniques occurs through properly designed and well-controlled prospective studies and will ensure that patients receive the best available therapy, regardless of whether it is surgical resection or ablative therapy.

8. The term non–small cell lung carcinoma (NSCLC) includes many tumor cell types, including large cell, squamous cell, and adenocarcinoma. The approach to diagnosis and management and the terminology used in describing these tumors are evolving rapidly. In particular, the evaluation and management of adenocarcinoma of the lung has shifted dramatically and firm establishment of NSCLC cell type prior to chemotherapy for advanced stage lung cancer is essential.

9. Increasing evidence suggests a significant role for gastroesophageal reflux disease in the pathogenesis of chronic lung diseases such as bronchiectasis and idiopathic pulmonary fibrosis, and it may also contribute to bronchiolitis obliterans syndrome in lung transplant patients.

10. Treatment of pulmonary aspergillosis/aspergilloma is individualized. Following colonization of a lung cavity or area of bronchiectasis, fungal growth within the cavity appears as an irregular cavity lining, progressing over time as a late finding in chronic pulmonary aspergillosis to a fungal ball called an aspergilloma. Asymptomatic patients can be observed without any additional treatment. Similarly, mild hemoptysis, which is not life-threatening, can be managed with medical therapy, including antifungals and cough suppressants. Oral triazole therapy is now considered the standard of care for chronic, cavitary pulmonary aspergillosis. Massive hemoptysis had traditionally been an indication for urgent or emergent operative intervention. However, with the advancement of endovascular techniques, bronchial artery embolization in select centers with experience in these techniques has been effective.

11. In patients with malignant pleural effusion, poor expansion of the lung (because of entrapment by tumor or adhesions) generally predicts a poor result with pleurodesis and is the primary indication for placement of indwelling pleural catheters. These catheters have dramatically changed the management of end-stage cancer treatment because they substantially shorten the amount of time patients spend in the hospital during their final weeks of life.
and stenosis; full-thickness injury can result in fistulae between the innominate artery anteriorly and the esophagus posteriorly. Avoidance requires careful cuff management to keep pressures as low as possible; in circumstances of prolonged ventilatory support and high airway pressure, cuff pressure monitoring (to maintain pressures <20 mmHg) is advisable.

Historically, clinically significant tracheal stenosis after tracheostomy occurred in 3% to 12% of cases, with severe stenosis in 1% to 2%. With the use of low-pressure cuffs, the estimated incidence has decreased to 4.9 cases per million patients per year. Intubation-related risk factors include prolonged intubation; high tracheostomy through the first tracheal ring or cricothyroid membrane; transverse rather than vertical incision on the trachea; oversized tracheostomy tube; prior tracheostomy or intubation; and traumatic intubation. Stenosis is also more common in older patients, in women, after radiation, or after excessive corticosteroid therapy, and in the setting of concomitant diseases such as autoimmune disorders, severe reflux disease, or obstructive sleep apnea and the setting of severe respiratory failure. However, even a properly placed tracheostomy can lead to tracheal stenosis because of scarring and local injury. Mild ulceration and stenosis are frequently seen after tracheostomy removal. Use of the smallest tracheostomy tube possible, rapid downsizing, and a vertical tracheal incision minimize the risk for posttracheostomy stenosis.

Stridor and dyspnea on exertion are the primary symptoms of tracheal stenosis. In the setting of postintubation injury, a significant portion of the cartilaginous structural support to the airway is destroyed by regional ischemic necrosis; during healing, a weblike fibrous growth develops and narrows the airway (Fig. 19-3). In contrast, stenosis caused by tracheostomy is most commonly due to an excess of granulation tissue formation around the tracheal stoma site. Time to onset of symptoms after extubation or tracheostomy decannulation usually ranges from 2 to 12 weeks, but symptoms can appear immediately or as long as 1 to 2 years later. Frequently, patients are misdiagnosed as having asthma or bronchitis, and treatment for such illnesses can persist for some time before the correct diagnosis is discovered. Generally, symptom intensity is related to the degree of stenosis and to the patient’s underlying pulmonary disease.
Figure 19-3. Diagram of the principal postintubation lesions. A. A circumferential lesion at the cuff site after the use of an endotracheal tube. B. Potential lesions after the use of tracheostomy tubes. Anterolateral stenosis can be seen at the sternal level. Circumferential stenosis can be seen at the cuff level (lower than with an endotracheal tube). The segment in between is often inflamed and malacic. C. Damage to the subglottic larynx. D. Tracheoesophageal fistula occurring at the level of the tracheostomy cuff; circumferential damage is usual at this level. E. Tracheoinnominate artery fistula. (Adapted with permission from Grillo H. Surgical treatment of postintubation tracheal injuries. J Thorac Cardiovasc Surg. 1979 Dec;78(6):860-875.)

**Acute Management.** A comprehensive bronchoscopic evaluation is critical in the initial phase of evaluation. Stenosis length, location, distance between the vocal cords and proximal stenosis, and distance from the distal aspect to the major carina must be documented. In patients with severe stenosis and respiratory compromise, rigid bronchoscopy can be used to dilate the stenosis; this provides immediate relief of the airway obstruction and facilitates thorough evaluation of the stenosis. Rarely, if ever, is tracheostomy necessary.

Most intubation injuries are located in the upper third of the trachea and can be accessed for resection through a collar incision. Resection typically involves 2 to 4 cm of trachea for benign stenosis. It is critical to fully resect all inflamed and scarred tissue. However, a primary anastomosis can still be performed without undue tension, even if up to one half of the trachea requires resection. Ideally, the patient is extubated in the operating room or shortly thereafter. For patients in whom tracheal resection is not possible, such as patients with significant comorbidities or with an excessively long stenosis, endotracheal stenting, typically silicone T-tubes, can provide palliation. Wire mesh stents should not be used, given their known propensity to erode through the wall of the airway. Balloon dilation, laser ablation, and tracheoplasty have also been described, though the efficacy is marginal.

Tracheal replacement is evolving as an option for management of tracheal stenosis as bioengineering techniques for decellularizing donor trachea have been developed. This removes all antigens against which the recipient immune system might react and enables use of the donor trachea scaffold without risk of rejection. Following decellularization, the donor tracheal scaffold is seeded with recipient chondrocytes, to restore tracheal rigidity, and with recipient epithelial cells, to recreate the inner epithelial lining. Several case reports of successful allogeneic tracheal transplantation have been published. The technique continues to be limited to a few highly specialized centers, due, in part, to the scarcity of donor trachea and the need for tissue bioengineering expertise as well as the lack of established efficacy for the approach. Current efforts are focused on creation of biosynthetic scaffolding that can be used instead of donor trachea. This would substantially increase the availability of the tracheal replacement material and enable widespread use of the technique, but early results have been contested, including three case reports called into question as containing multiple data fabrications and omissions.

**Tracheal Fistulas**

**Tracheoinnominate Artery Fistula.** Tracheoinnominate artery fistula has two main causes: low placement of a tracheostomy and hyperinflation of the tracheal cuff. Tracheostomy placement should be through the second to fourth tracheal rings without reference to the location of the sternal notch. When placed below the fourth tracheal ring, the inner curve of the tracheostomy cannula will be positioned to exert pressure on the posterior aspect of the innominate artery, leading to arterial erosion. Similarly, the tracheal cuff, when hyperinflated, will cause ischemic injury to the anterior airway and subsequent erosion into the artery. Most cuff-induced fistulas will develop within 2 weeks after placement of the tracheostomy. Clinically, tracheoinnominate artery fistulas present with bleeding. A premonitory hemorrhage often occurs and, although it is usually not massive, must not be ignored or simply attributed to general airway irritation or wound bleeding. With significant bleeding, the tracheostomy cuff can be hyperinflated to temporarily occlude the arterial injury. If such an effort is unsuccessful, the tracheostomy incision should be immediately opened widely and a finger inserted to compress the artery.
against the manubrium (Fig. 19-4). The patient can then be orally intubated, and the airway suctioned free of blood. Emergent surgical resection of the involved segment of artery is performed, usually without reconstruction.

**Tracheoesophageal Fistula.** Tracheoesophageal fistulas (TEFs) occur primarily in patients receiving prolonged mechanical ventilator support concomitant with an indwelling nasogastric tube. Cuff compression of the membranous trachea against the nasogastric tube leads to airway and esophageal injury and fistula development. Clinically, airway suctioning reveals saliva, gastric contents, or tube feedings. Gastric insufflation, secondary to positive pressure ventilation, can occur. Bronchoscopy is diagnostic; with the bronchoscope inserted, the endotracheal tube is withdrawn, and the fistula at the cuff site is exposed. Alternatively, esophagoscopy demonstrates the cuff of the endotracheal tube in the esophagus.

Treatment, first and foremost, requires removing tubes from the esophagus and weaning the patient from the ventilator. The cuff of the endotracheal tube should be placed below the fistula, avoiding overinflation. To minimize aspiration, a gastrostomy tube should be placed for gastric decompression (to prevent reflux) and a jejunostomy tube for feeding. If aspiration persists, esophageal diversion with cervical esophagostomy can be performed. Once weaned from the ventilator, tracheal resection and primary anastomosis, repair of the esophageal defect, and interposition of a muscle flap between the trachea and esophagus can be performed (Fig. 19-5).

**Tracheal Neoplasms**

Although extremely rare, the most common primary tracheal neoplasms are squamous cell carcinomas (related to smoking) and adenoid cystic carcinomas. Clinically, tracheal tumors present with cough, dyspnea, hemoptysis, stridor, or symptoms of invasion of contiguous structures (such as the recurrent laryngeal nerve or the esophagus). The most common radiologic finding of tracheal malignancy is tracheal stenosis, but it is found in only 50% of cases. With tumors other than squamous cell carcinomas, symptoms may persist for months because of slow tumor growth rates. Stage of presentation is advanced, with approximately 50% of patients presenting with stage IV disease. Five-year survival for all tracheal neoplasms is 40% but falls to 15% for those with stage IV disease.

Squamous cell carcinomas often present with regional lymph node metastases and are frequently unresectable at presentation. Their biologic behavior is similar to that of squamous cell carcinoma of the lung. Adenoid cystic carcinomas, a type of...
salivary gland tumor, are generally slow growing, spread submucosally, and tend to infiltrate along nerve sheaths and within the tracheal wall. Although indolent in nature, adenoid cystic carcinomas are malignant and can spread to regional lymph nodes, lung, and bone. Squamous cell carcinoma and adenoid cystic carcinomas represent approximately 65% of all tracheal neoplasms. The remaining 35% is comprised of small cell carcinomas, mucoepidermoid carcinomas, adenocarcinomas, lymphomas, and others.7

**Therapy.** Evaluation and treatment of patients with tracheal tumors should include neck and chest computed tomography (CT) and rigid bronchoscopy. Rigid bronchoscopy permits general assessment of the airway and tumor; it also allows debridement or laser ablation of the tumor to provide relief of dyspnea. If the tumor is judged to be completely resectable, primary resection and anastomosis is the treatment of choice for these tumors (Fig. 19-6). Up to 50% of the length of the trachea can be resected with primary anastomosis. In most tracheal resections, anterolateral tracheal mobilization and suturing of the chin to the sternum for 7 days are done routinely. Use of laryngeal and hilar release is determined at the time of surgery, based on the surgeon’s judgment of the degree of tension present. For longer resections, specialized maneuvers are necessary such as laryngeal release and right hilar release to minimize tension on the anastomosis.

Postoperative mortality, which occurs in up to 10% of patients, is associated with the length of tracheal resection, use of laryngeal release, the type of resection, and the histologic type of the cancer. Factors associated with improved long-term survival include complete resection and use of radiation as adjuvant therapy in the setting of incomplete resection.9 Due to their radiosensitivity, radiotherapy is frequently given postoperatively after resection of both adenoid cystic carcinomas and squamous cell carcinomas. A dose of 50 Gy or greater is usual. Nodal positivity does not seem to be associated with worse survival. Survival at 5 and 10 years is much better for adenoid...
the nodes of the lymphatic sump lie around the bronchus intermedius (bounded above by the right upper lobe bronchus and below by the middle lobe and superior segmental bronchi). On the left, the lymphatic sump is confined to the interlobar fissure, with the lymph nodes in the angle between the lingular and lower lobe bronchi. These nodes are always in close proximity to pulmonary arterial branches and typically must be carefully dissected to identify the pulmonary arterial segments for division during lung resection.

The N2 lymph nodes consist of four main groups. (a) The anterior mediastinal nodes are located in association with the upper surface of the pericardium, the phrenic nerves, the ligamentum arteriosum, and the left aspect of the innominate vein. (b) The posterior mediastinal group includes paraeosophageal lymph nodes within the inferior pulmonary ligament and, more superiorly, between the esophagus and trachea near the arch of the azygos vein. (c) The tracheobronchial lymph nodes are made up of three subgroups that are located near the bifurcation of the trachea. These include the subcarinal nodes, which lie in the obtuse angle between the trachea and each main stem bronchus, and the nodes that lay anterior to the lower end of the trachea. (d) Paratracheal lymph nodes are located in proximity to the trachea in the superior mediastinum. Those on the right side form a chain with the tracheobronchial nodes inferiorly and with some of the deep cervical nodes above (scalene lymph nodes).

Lymphatic drainage to the mediastinal lymph nodes from the right lung is ipsilateral, except for occasional bilateral drainage to the superior mediastinum. In contrast, in the left lung, particularly the left lower lobe, lymphatic drainage occurs with equal frequency to ipsilateral and contralateral superior mediastinal nodes.

cystic (73% and 57%, respectively) than for tracheal cancers (47% and 36%, respectively; \( P < .05 \)). For patients with unresectable tumors, radiation may be given as the primary therapy to improve local control, but it is rarely curative. For recurrent airway compromise, stenting or laser therapies should be considered as part of the treatment algorithm.

**LUNG**

**Anatomy**

**Segmental Anatomy.** The segmental bronchial and vascular anatomy of the lungs allows subsegmental and segmental resections, if the clinical situation requires or if lung tissue can be preserved\(^a\) (Fig. 19-7). Note the continuity of the pulmonary parenchyma between adjacent segments of each lobe.

**Lymphatic Drainage.** Lymph nodes that drain the lungs are divided into two groups according to the tumor-node-metastasis (TNM) staging system for lung cancer: the pulmonary lymph nodes (N1) and the mediastinal nodes (N2) (Fig. 19-8).

The N1 lymph nodes constitute the following: (a) intrapulmonary or segmental nodes that lie at points of division of segmental bronchi or in the bifurcations of the pulmonary artery; (b) lobar nodes that lie along the upper, middle, and lower lobe bronchi; (c) interlobar nodes located in the angles formed by the main bronchi bifurcating into the lobar bronchi; and (d) hilar nodes along the main bronchi. The interlobar lymph nodes lie in the depths of the interlobar fissure on each side and constitute a lymphatic sump for each lung, referred to as the **lymphatic sump of Borrie**; all of the pulmonary lobes of the corresponding lung drain into this group of nodes (Fig. 19-9). On the right,
Normal Lung Histology

The lung can be conveniently viewed as two linked components: the tracheobronchial tree (or conducting airways component) and the alveolar spaces (or gas exchange component). The tracheobronchial tree consists of approximately 23 airway divisions to the level of the alveoli. It includes the main bronchi, lobar bronchi, segmental bronchi (to designated bronchopulmonary segments), and terminal bronchioles (i.e., the smallest airways still lined by bronchial epithelium and without alveoli). The tracheobronchial tree is normally lined by pseudostratified ciliated columnar cells and mucous (or goblet) cells, which both derive from basal cells (Fig. 19-10). Ciliated cells predominate. Goblet cells, which release mucus, can significantly increase in number in acute bronchial injury, such as exposure to cigarette smoke. The normal bronchial epithelium also contains bronchial submucosal glands, which are mixed salivary-type glands.
containing mucous cells, serous cells, and neuroendocrine cells called Kulchitsky cells, which are also found within the surface epithelium. The bronchial submucosal glands can give rise to salivary gland–type tumors, including mucoepidermoid carcinomas and adenoid cystic carcinomas.

Two cell types, called type I and type II pneumocytes, make up the alveolar epithelium. Type I pneumocytes comprise 40% of the total number of alveolar epithelial cells, but cover 95% of the surface area of the alveolar wall. These cells are not capable of regeneration because they have no mitotic potential. Type II pneumocytes cover only 3% of the alveolar surface, but comprise 60% of the alveolar epithelial cells. In addition, clusters of neuroendocrine cells are seen in the alveolar spaces.

**Preinvasive Lesions**

The term “precancerous” does not mean that an inevitable progression to invasive carcinoma will occur, but such lesions, particularly those with high-grade dysplasia,\(^{11,12}\) do constitute a clear marker for potential development of invasive cancer.

Three precancerous lesions of the respiratory tract are currently recognized.

1. **Squamous dysplasia and carcinoma in situ.** Cigarette smoke can induce a transformation of the tracheobronchial pseudostratified epithelium to metaplastic squamous mucosa, with subsequent evolution to dysplasia as cellular abnormalities accumulate. Dysplastic changes include altered cellular polarity and increased cell size, number of cell layers, nuclear-to-cytoplasmic ratio, and number of mitoses. Gradations are considered mild, moderate, or severe. Carcinoma in situ represents carcinoma still confined by the basement membrane.

2. **Atypical adenomatous hyperplasia (AAH).** AAH is a lesion smaller than 5.0 mm, comprising epithelial cells lining the alveoli that are similar to type II pneumocytes. Histologically, AAH is similar to adenocarcinoma in situ; it represents the beginning stage of a stepwise evolution to adenocarcinoma in situ and then to adenocarcinoma. With the availability of thin-section CT, it is possible to detect
Table 19-1

<table>
<thead>
<tr>
<th>Difference between invasive mucinous adenocarcinoma and nonmucinous adenocarcinoma in situ/minimally invasive adenocarcinoma/lepidic predominant adenocarcinoma</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INVASIVE MUCINOUS ADENOCARCINOMA</strong> (FORMERLY MUCINOUS BAC)</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Smoker</td>
</tr>
<tr>
<td>Radiographic appearance</td>
</tr>
<tr>
<td>Cell type</td>
</tr>
<tr>
<td>Phenotype</td>
</tr>
<tr>
<td>CK7</td>
</tr>
<tr>
<td>CK20</td>
</tr>
<tr>
<td>TTF-1</td>
</tr>
<tr>
<td>Genotype</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Note: *Numbers represent the percentage of cases that are reported to be positive.

Abbreviations: BAC = bronchioloalveolar carcinoma; AIS = adenocarcinoma in situ; MIA = minimally invasive adenocarcinoma; LPA = lepidic predominant adenocarcinoma; EGFR = epidermal growth factor receptor; TTF = thyroid transcription factor.

Table 19-2

New classification system for lung adenocarcinoma

<table>
<thead>
<tr>
<th>Preinvasive lesions</th>
<th>Invasion predominant tumor with $\leq 5$ mm invasion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atypical adenomatous hyperplasia</td>
<td>Lepidic predominant adenocarcinoma (formerly nonmucinous BAC pattern)</td>
</tr>
<tr>
<td>Adenocarcinoma in situ ($\leq 3$ cm formerly BAC)</td>
<td>Micropapillary predominant</td>
</tr>
<tr>
<td>Nonmucinous</td>
<td>Mixed mucinous/nonmucinous</td>
</tr>
<tr>
<td>Mucinous</td>
<td>Minimally invasive adenocarcinoma ($\leq 3$ cm lepidic predominant tumor with $\leq 5$ mm invasion)</td>
</tr>
<tr>
<td>Mixed mucinous/nonmucinous</td>
<td>Nonmucinous</td>
</tr>
<tr>
<td>Invasive adenocarcinoma</td>
<td>Mixed mucinous/nonmucinous</td>
</tr>
<tr>
<td>Lepidic predominant adenocarcinoma</td>
<td>Adventitious lesions</td>
</tr>
<tr>
<td>(formerly nonmucinous BAC pattern)</td>
<td>Colloid</td>
</tr>
<tr>
<td>with $&gt;5$ mm invasion</td>
<td>Fetal (low and high grade)</td>
</tr>
<tr>
<td>Acinar predominant</td>
<td>Enteric</td>
</tr>
<tr>
<td>Papillary predominant</td>
<td></td>
</tr>
<tr>
<td>Micropapillary predominant</td>
<td></td>
</tr>
<tr>
<td>Solid predominant with mucin production</td>
<td></td>
</tr>
</tbody>
</table>

Variants of invasive adenocarcinoma

- Invasive mucinous adenocarcinoma (formerly mucinous BAC)
- Colloid
- Fetal (low and high grade)
- Enteric

Abbreviations: BAC = bronchioloalveolar carcinoma; IASLC = International Association for the Study of Lung Cancer; ATS = American Thoracic Society; ERS = European Respiratory Society.

but occasionally it will present as part of a solid or part-solid nodule. Mucinous AIS is more likely to appear solid or to have the appearance of consolidation. As with AAH, the lesions can be single or multiple; the ground-glass changes in AIS, however, tend to have a higher attenuation compared to AAH. The 8th edition American Joint Committee on Cancer (AJCC) staging manual t-stage for AIS is tumor in situ (Tis).

2. Minimally invasive adenocarcinoma (MIA). In the same size solitary lesion, if less than 5 mm of invasion are noted within a predominantly lepidic growth pattern, the lesion is termed minimally invasive adenocarcinoma (MIA) to indicate a patient group with near 100% survival when the lesion is completely resected. This differentiates patients with AIS, but recognizes the fact that the presence of invasion becomes prognostically significant when the size of the invasive component reaches 5 mm or greater in size.16 If multiple areas of microscopic invasion are found within the lepidic growth, the size of the largest invasive area, measured in the largest dimension, is used; this area must be $\leq 5$ mm to be considered MIA. As with AIS, MIA is very rarely mucinous. The invasive component histologically is acinar, papillary, micropapillary, and/or solid and shows tumor cells infiltrating into the surrounding myofibroblastic stroma. On CT scan, the appearance of MIA is often a part-solid nodule ($\leq 5$ mm) with a predominant ground-glass component, but can be highly variable. The 8th edition American Joint Committee on Cancer (AJCC) staging manual t-stage for MIA is T1mi.

3. Lepidic predominant adenocarcinoma (LPA). If lymphovascular invasion, pleural invasion, tumor necrosis, or more than 5 mm of invasion are noted in a lesion that has lepidic growth as its predominant component, MIA is excluded, the lesion is called lepidic predominant adenocarcinoma (LPA), and the size of the invasive component is recorded for the T stage.

4. Invasive adenocarcinoma. The new classification system now recommends classifying invasive adenocarcinoma by the most predominant subtype after histologic evaluation of the resection specimen. To determine the predominant subtype, histologic sections are evaluated, and the patterns are determined, in 5% increments, throughout the specimen. This semiquantitative method encourages the viewer to identify and quantify all patterns present, rather than focusing on a single pattern. In the pathology report, the tumor is classified by the predominant pattern, with percentages of the subtypes also reported (Fig. 19-11). Subtypes include:

   a. Lepidic predominant
   b. Acinar predominant
   c. Papillary predominant
   d. Micropapillary predominant
   e. Solid predominant

Adenocarcinoma is often peripherally located and frequently discovered incidentally on routine chest radiographs, unlike squamous cell cancers. When symptoms occur, they are due to pleural or chest wall invasion (pleuritic or chest wall pain) or pleural seeding with malignant pleural effusion. Invasive adenocarcinoma is usually solid by CT scan, but can also be part-solid and even a ground-glass nodule. Occasionally, a lobular ground-glass opacification may be present, which is often associated with significant respiratory compromise and can be mistaken for lobar pneumonia. Bubble-like or cystic lucency on CT scan in small ($\leq 2$ cm) adenocarcinomas or extensive associated ground-glass components correlate with slow growth and well-differentiated tumors and a more favorable prognosis. Intratumoral air bronchograms are usually indicative of well-differentiated tumor, whereas spiculations that are coarse and thick ($\geq 2$ mm) portend vascular invasion and nodal metastasis and are associated with decreased survival following complete surgical resection. Pleural retraction is also a poor prognostic indicator.

5. Additional histologic variants include colloid adenocarcinoma (formerly mucinous cystadenocarcinoma), fetal adenocarcinoma, and enteric adenocarcinoma. Clear cell and signet ring cell types are no longer considered to be distinct subtypes as they are found in association with most of the five dominant histologic patterns (lepidic, acinar, papillary, micropapillary, and solid). However, they are still notable, as they can signal clinically relevant molecular changes, such as the presence of the EML4-ALK fusion gene in solid tumors with signet ring features.

Squamous Cell Carcinoma Representing 30% to 40% of lung cancers, squamous cell carcinoma is the most frequent cancer in men and highly correlated with cigarette smoking. They arise primarily in the main, lobar, or first segmental bronchi, which are collectively referred to as the central airways. Symptoms of airway irritation or obstruction are common, and include cough,
hemoptysis, wheezing (due to high-grade airway obstruction), dyspnea (due to bronchial obstruction with or without postobstructive atelectasis), and pneumonia (caused by airway obstruction with secretion retention and atelectasis).

Occasionally a more peripherally based squamous cell carcinoma will develop in a tuberculosis scar or in the wall of a bronchiectatic cavity. Histologically, cells develop a pattern of clusters with intracellular bridges and keratin pearls. Central necrosis is frequent and may lead to the radiographic findings of a cavity (possibly with an air-fluid level). Such cavities may become infected, with resultant abscess formation.

**Large Cell Carcinoma** Large cell carcinoma accounts for 10% to 20% of lung cancers and may be located centrally or peripherally. These tumors have cell diameters of 30 to 50 µm, which are often admixed with various other malignant cell types. Large cell carcinoma can be confused with a large cell variant of neuroendocrine carcinoma, but can be differentiated by special immunohistochemical stains.

**Salivary Gland–Type Neoplasms.** Salivary-type submucosal bronchial glands throughout the tracheobronchial tree can give rise to tumors that are histologically identical to those seen in the salivary glands. The two most common are adenoid cystic carcinoma and mucoepidermoid carcinoma. Both occur centrally due to their site of origin. Adenoid cystic carcinoma is a slow-growing tumor that is locally and systemically invasive, growing submucosally and infiltrating along perineural sheaths. Mucoepidermoid carcinoma consists of squamous and mucous cells and is graded as low or high grade, depending on mitotic rate and degree of necrosis.

**Neuroendocrine Neoplasms.** Neuroendocrine lung tumors are classified into neuroendocrine hyperplasia and three grades of neuroendocrine carcinoma (NEC). Immunohistochemical staining for neuroendocrine markers (including chromogranins, synaptophysin, CD57, and neuron-specific enolase) is essential to accurate diagnosis.17

Grade I NEC (classic or typical carcinoid) is a low-grade NEC; 80% arise in the central airway epithelium and occur primarily in younger patients. Because it is a central lesion, hemoptysis, with or without airway obstruction and pneumonia is the most common presentation. Histologically, tumor cells

---

Figure 19-11. Major histologic patterns of invasive adenocarcinoma. A. Lepidic predominant pattern with mostly lepidic growth (right) and a smaller area of invasive acinar adenocarcinoma (left). B. Lepidic pattern consists of a proliferation type II pneumocytes and Clara cells along the surface alveolar walls. C. Area of invasive acinar adenocarcinoma (same tumor as in A and B). D. Acinar adenocarcinoma consists of round to oval-shaped malignant glands invading a fibrous stroma. E. Papillary adenocarcinoma consists of malignant cuboidal to columnar tumor cells growing on the surface of fibrovascular cores. F. Micropapillary adenocarcinoma consists of small papillary clusters of glandular cells growing within this airspace, most of which do not show fibrovascular cores. G. Solid adenocarcinoma with mucin consisting of sheets of tumor cells with abundant cytoplasm and mostly vesicular nuclei with several conspicuous nucleoli. No acinar, papillary, or lepidic patterns are seen, but multiple cells have intracytoplasmic basophilic globules that suggest intracytoplasmic mucin. H. Solid adenocarcinoma with mucin. Numerous intracytoplasmic droplets of mucin are highlighted with this diastase-periodic acid Schiff stain. (Reproduced with permission from Travis WD, Brambilla E, Noguchi M, et al: International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma, J Thorac Oncol. 2011 Feb;6(2):244-285.)
are arranged in cords and clusters with a rich vascular stroma, which can lead to life-threatening hemorrhage with even simple bronchoscopic biopsy maneuvers. Regional lymph node metastases are seen in 15% of patients, but systemic spread and death from Grade I NEC is rare.

Grade II NECs (atypical carcinoid) have a much higher malignant potential and, unlike grade I NEC, are etiologically linked to cigarette smoking and more likely to be peripherally located. Histologic findings may include areas of necrosis, nuclear pleomorphism, and higher mitotic rates. Lymph node metastases are found in 30% to 50% of patients. At diagnosis, 25% of patients already have remote metastases.

Grade III NEC large cell–type tumors occur primarily in heavy smokers and in the mid to peripheral lung fields. Often large with central necrosis and a high mitotic rate, their neuroendocrine nature is revealed by positive immunohistochemical staining for at least one neuroendocrine marker.

Grade IV NEC (small cell lung carcinoma [SCLC]) is the most malignant NEC and accounts for 25% of all lung cancers; these NECs often have early, widespread metastases. These cancers also arise primarily in the central airways. As with squamous cell cancers, symptoms include cough, hemoptysis, wheezing (due to high-grade airway obstruction), dyspnea (due to bronchial obstruction with or without postobstructive atelectasis), and pneumonia (caused by airway obstruction with secretion retention and atelectasis). Evaluation includes expert pathology review and comprehensive evaluation for metastatic disease. Three groups of grade IV NEC are recognized: pure small cell carcinoma (previously referred to as oat cell carcinoma), small cell carcinoma with a large cell component, and combined (mixed) tumors.

Grade IV NECs consist of smaller cells (diameter 10 to 20 µm) with little cytoplasm and very dark nuclei; they can be difficult to distinguish from lymphoproliferative lesions and atypical carcinoid tumors. Histologically, a high mitotic rate with easily visualized multiple mitoses and areas of extensive necrosis are characteristic. Importantly, very small bronchoscopic biopsies can distinguish NSCLC from SCLC, but crush artifact may make NSCLC appear similar to SCLC. If uncertainty exists, special immunohistochemical stains or rebiopsy (or both) will be necessary. These tumors are the leading producer of paraneoplastic syndromes.

Lung Cancer Epidemiology
Lung cancer is the leading cancer killer and second most frequently diagnosed cancer in the United States, accounting for 26% of all cancer deaths in 2017—more than cancers of the breast, prostate, ovary, and colon and rectum combined (Fig. 19-12).14 Lung cancer incidence continues to decline, though at twice the rate for men compared to women (Fig. 19-13A, B). It is encouraging, however, that the average annual death rate declined by 3.5% per year for men and 2% per year for women from 2010 to 2014, representing a 43% decline in mortality for men and a 17% decline for women from 1990 to 2014.

![Estimated new cases](image1)

**Estimated new cases**

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prostate</td>
<td>161,360</td>
<td>19%</td>
</tr>
<tr>
<td>Lung &amp; bronchus</td>
<td>116,990</td>
<td>14%</td>
</tr>
<tr>
<td>Colon &amp; rectum</td>
<td>71,420</td>
<td>9%</td>
</tr>
<tr>
<td>Urinary bladder</td>
<td>60,490</td>
<td>7%</td>
</tr>
<tr>
<td>Melanoma of the skin</td>
<td>52,170</td>
<td>6%</td>
</tr>
<tr>
<td>Kidney &amp; renal pelvis</td>
<td>40,610</td>
<td>5%</td>
</tr>
<tr>
<td>Non-Hodgkin lymphoma</td>
<td>40,080</td>
<td>5%</td>
</tr>
<tr>
<td>Leukemia</td>
<td>36,290</td>
<td>4%</td>
</tr>
<tr>
<td>Oral cavity &amp; pharynx</td>
<td>35,720</td>
<td>4%</td>
</tr>
<tr>
<td>Liver &amp; intrahepatic bile duct</td>
<td>29,200</td>
<td>3%</td>
</tr>
<tr>
<td><strong>All Sites</strong></td>
<td><strong>836,150</strong></td>
<td>100%</td>
</tr>
</tbody>
</table>

![Estimated deaths](image2)

**Estimated deaths**

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung &amp; bronchus</td>
<td>84,590</td>
<td>27%</td>
</tr>
<tr>
<td>Colon &amp; rectum</td>
<td>27,150</td>
<td>9%</td>
</tr>
<tr>
<td>Prostate</td>
<td>26,730</td>
<td>8%</td>
</tr>
<tr>
<td>Pancreas</td>
<td>22,300</td>
<td>7%</td>
</tr>
<tr>
<td>Liver &amp; intrahepatic bile duct</td>
<td>19,610</td>
<td>6%</td>
</tr>
<tr>
<td>Leukemia</td>
<td>14,300</td>
<td>4%</td>
</tr>
<tr>
<td>Esophagus</td>
<td>12,720</td>
<td>4%</td>
</tr>
<tr>
<td>Urinary bladder</td>
<td>12,240</td>
<td>4%</td>
</tr>
<tr>
<td>Non-Hodgkin lymphoma</td>
<td>11,450</td>
<td>4%</td>
</tr>
<tr>
<td>Brain &amp; other nervous system</td>
<td>9,620</td>
<td>3%</td>
</tr>
<tr>
<td><strong>All Sites</strong></td>
<td><strong>318,420</strong></td>
<td>100%</td>
</tr>
</tbody>
</table>

![Image 1](image1)

Figure 19-13. Trends in death rates by sex for select cancers, United States, 1930 to 2014. **A.** Males. **B.** Female rates are age-adjusted to the 2000 U.S. standard population. Due to improvements in International Classification of Diseases (ICD) coding over time, numerator data for cancers of the lung and bronchus, colon and rectum, liver, and uterus differ from the contemporary time period. For example, rates for lung and bronchus include pleura, trachea, mediastinum, and other respiratory organs. *(Adapted with permission from Siegel RL, Miller KD, Jemal A: Cancer Statistics, 2017, CA Cancer J Clin. 2017 Jan;67(1):7-30.)*
to 2014. Unfortunately, most patients are still diagnosed at an advanced stage of disease (22% with regional metastasis and 57% with distant metastasis), so therapy is rarely curative.

Prognostic markers for lung cancer survival include female sex (5-year survival of 18.3% for women vs. 13.8% for men), younger age (5-year survival of 22.8% for those <45 years vs. 13.7% for those >65 years), and white race (5-year survival of 16.1% for whites vs. 12.2% for blacks). When access to advanced medical care is unrestricted, as for the military population, the racial difference in survival disappears, suggesting that, at least in part, differences in survival may be explained by less access to advanced medical care and later diagnosis.

**Risk Factors for Lung Cancer.** Cigarette smoking is the leading preventable cause of cancer death, accounting for 29% of the population attributable fraction in 2010, and is implicated as a causal factor in approximately 90% of lung cancers in men and nearly 80% in women. Two lung cancer types—squamous cell and small cell carcinoma—are extraordinarily rare in the absence of cigarette smoking. The risk of developing lung cancer escalates with the number of cigarettes smoked, the number of years of smoking, and the use of unfiltered cigarettes. Conversely, the risk of lung cancer declines with smoking cessation, but never drops to that of never smokers, regardless of the length of abstinence (Table 19-3).

Radon exposure accounts for the vast majority of the remaining cancers. Approximately 25% of all lung cancers worldwide and 53% of cancers in women are not related to smoking, and most of them (62%) are adenocarcinomas. Table 19-4 summarizes the existing data regarding the etiology of lung cancer in nonsmokers.

Nearly 3500 deaths from lung cancer each year are attributable to secondhand (environmental) smoke exposure, which confers an excess risk for lung cancer of 24% when a nonsmoker lives with a smoker. Risk is conferred by exposure to any burning tobacco, including cigars. The amount of secondhand exposure from one large cigar is equivalent to the exposure from 21 cigarettes. As with active smoking, risk of developing lung cancer increases with longer duration and higher level of exposure to environmental tobacco.

Over 7000 chemicals have been identified in tobacco smoke, and more than 70 of the compounds are known to be carcinogens. The main chemical carcinogens are polycyclic aromatic hydrocarbons, which are actively or passively inhaled in the tobacco smoke and absorbed; these compounds are activated by specific enzymes and become mutagenic, bind to macromolecules such as deoxyribonucleic acid (DNA), and induce genetic mutations. In treating any patient with a previous smoking history, it is important to remember that there has been field cancerization of the entire aerodigestive tract. The patient’s risk is increased for cancers of the oral cavity, pharynx, larynx, tracheobronchial tree and lung, and esophagus. In examining such patients, a detailed history and physical examination of these organ systems must be performed.

Other causes of lung cancer include exposure to a number of industrial compounds, including asbestos, arsenic, and chromium compounds. In fact, the combination of asbestos and cigarette smoke exposure has a multiplicative effect on risk. Pre-existing lung disease confers an increased risk of lung cancer—up to 13%—for individuals who have never smoked. Patients with chronic obstructive pulmonary disease are at higher risk for lung cancer than would be predicted based on smoking risk alone. Patients with secondary scar formation related to a history of tuberculosis or other lung infections also have a higher risk of primary lung carcinoma. This increase is thought to be related to poor clearance of inhaled carcinogens and/or to the effects of chronic inflammation.

**Screening for Lung Cancer in High-Risk Populations**

In 2002, the National Lung Screening Trial (NLST) was launched to determine whether screening with CT in high-risk populations would reduce mortality from lung cancer. The study randomized 53,353 eligible patients age 55 to 74 years to either three annual low-dose helical CT scans (LDCT; aka spiral CT) or posteroanterior view chest radiograph. Patients were eligible for the trial if they had a greater than 30 pack-year history of cigarette smoking; had smoked within the past 15 years if a former smoker; had no prior history of lung cancer; had no history of other life-threatening cancers in the prior 5 years; did not have symptoms suggestive of an undiagnosed lung cancer (such as hemoptysis or weight loss); and had not had a chest CT scan in the prior 18 months. Accrual to the study was excellent, and the primary endpoint of a 20% relative reduction in mortality was achieved in 2010. An absolute risk reduction of lung cancer death of four per 1000 individuals screened by LDCT was realized. Interestingly, all-cause mortality was also reduced by nearly 7% in the LDCT group, further emphasizing the impact of lung cancer on the mortality of smokers and former smokers.

The U.S. Preventive Services Task Force (USPSTF) now recommends annual screening for lung cancer with low-dose computed tomography screening in high risk patients. Annual screening averted 14% of lung cancer deaths when applied to a population of asymptomatic adults age 55 to 80 years who have a 30 pack-year smoking history and are either currently smoking or have quit within the past 15 years. Patients should be healthy enough to tolerate curative treatment, specifically surgery per guidelines, and screening should be discontinued once the patient has not smoked for 15 years or develops a life-limiting health condition, becomes unable to tolerate lung surgery or is unwilling to undergo curative lung resection. Requirements for coverage differ between Medicare and private insurers, with private insurers following the USPSTF guidelines, while Medicare uses age 55 to 77 years with the same smoking history but does not define comorbid conditions. Use of standardized reporting with criteria for lung nodule identification and classification is required by the Center for Medicaid & Medicare Services (CMS) but is only recommended by private insurers. Shared decision-making is also required by Medicare, but it is only recommended by private insurers. Medicare also differs from private

---

**Table 19-3**

Relative risk of lung cancer in smokers

<table>
<thead>
<tr>
<th>SMOKING CATEGORY</th>
<th>RELATIVE RISK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never smoked</td>
<td>1.0</td>
</tr>
<tr>
<td>Currently smoke</td>
<td>15.8–16.3</td>
</tr>
<tr>
<td>Formerly smoked</td>
<td></td>
</tr>
<tr>
<td>Years of abstinence</td>
<td></td>
</tr>
<tr>
<td>1–9</td>
<td>5.9–19.5</td>
</tr>
<tr>
<td>10–19</td>
<td>2.0–6.1</td>
</tr>
<tr>
<td>&gt;20</td>
<td>1.9–3.7</td>
</tr>
</tbody>
</table>

### Table 19-4
Summary of selected studies of risk factors for lung cancer in individuals who never smoked

<table>
<thead>
<tr>
<th>RISK FACTOR</th>
<th>RISK ESTIMATE (95% CI)</th>
<th>COMMENTS</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environmental tobacco smoke</td>
<td>1.19 (90% CI: 1.04–1.35)</td>
<td>Meta-analysis of 11 U.S. studies of spousal exposure (females only)</td>
<td>225</td>
</tr>
<tr>
<td></td>
<td>1.21 (1.13–1.30)</td>
<td>Meta-analysis of 44 case-control studies worldwide of spousal exposure</td>
<td>226</td>
</tr>
<tr>
<td></td>
<td>1.22 (1.13–1.33)</td>
<td>Meta-analysis of 25 studies worldwide of workplace exposure</td>
<td>226</td>
</tr>
<tr>
<td></td>
<td>1.24 (1.18–1.29)</td>
<td>Meta-analysis of 22 studies worldwide of workplace exposure</td>
<td>227</td>
</tr>
<tr>
<td>Residential radon</td>
<td>8.4% (3.0%–15.8%) per 100 Bq m⁻³ increase in measured radon</td>
<td>Meta-analysis of 13 European studies</td>
<td>228</td>
</tr>
<tr>
<td></td>
<td>11% (0%–28%) per 100 Bq m⁻³</td>
<td>Meta-analysis of 7 North American studies</td>
<td>229</td>
</tr>
<tr>
<td>Cooking oil vapors</td>
<td>2.12 (1.81–2.47)</td>
<td>Meta-analysis of 7 studies from China and Taiwan (females who never smoked)</td>
<td>230</td>
</tr>
<tr>
<td>Indoor coal and wood burning</td>
<td>2.66 (1.39–5.07)</td>
<td>Meta-analysis of 7 studies from China and Taiwan (both sexes)</td>
<td>230</td>
</tr>
<tr>
<td></td>
<td>1.22 (1.04–1.44)</td>
<td>Large case-control study (2861 cases and 3118 controls)</td>
<td>231</td>
</tr>
<tr>
<td></td>
<td>2.5 (1.5–3.6)</td>
<td>Large case-control study (1205 cases and 1541 controls)</td>
<td>232</td>
</tr>
<tr>
<td>Genetic factors: family history,</td>
<td>1.51 (1.11–2.06)</td>
<td>Meta-analysis of 28 case-control, 17 cohort, and 7 twin studies</td>
<td>233</td>
</tr>
<tr>
<td>CYP1A1 Ile462Val polymorphism, XRCC1</td>
<td>2.99 (1.51–5.91)</td>
<td>Meta-analysis of 14 case-control studies of Caucasian never smokers</td>
<td>234</td>
</tr>
<tr>
<td>variants</td>
<td>2.04 (1.17–3.54)</td>
<td>Meta-analysis of 21 case-control studies of Caucasian and Asian never smokers (significant for Caucasians only)</td>
<td>235</td>
</tr>
<tr>
<td>No association</td>
<td></td>
<td>Meta-analysis of 13 case-control studies</td>
<td>236</td>
</tr>
<tr>
<td>No association overall; reduced risk</td>
<td>0.65 (0.46–0.83) with Arg194Trp polymorphism and 0.56 (0.36–0.86) with Arg280His for heavy smokers</td>
<td>Large case-control study from Europe (2188 cases and 2198 controls)</td>
<td>237</td>
</tr>
<tr>
<td>Viral factors: HPV 16 and 18</td>
<td>10.12 (3.88–26.4) for never smoking women &gt;60 y</td>
<td>Case-control study (141 cases, 60 controls) from Taiwan of never smoking women</td>
<td>239</td>
</tr>
</tbody>
</table>

**Abbreviations:** Bq = becquerels; CI = confidence interval; CYP1A1 = cytochrome P450 enzyme 1A1; HPV = human papilloma virus.

Solitary Pulmonary Nodule
A solitary pulmonary nodule is typically described as a single, well-circumscribed, spherical lesion that is 3 cm or less cm in diameter and completely surrounded by normal aerated lung parenchyma.24 Lung atelectasis, hilar enlargement, and pleural effusion are absent. The majority are detected incidentally on chest radiographs (CXR) or CT scans obtained for some other purpose. About 150,000 solitary nodules are found incidentally each year, with increasing numbers as low-dose computed tomography screening in high-risk populations is adopted. The clinical significance of such a lesion depends on whether or not it represents a malignancy.

The differential diagnosis of a solitary pulmonary nodule should include a broad variety of congenital, neoplastic, inflammatory, vascular, and traumatic disorders. The probability of cancer in a solitary pulmonary nodule increases if the patient has a history of smoking (50% or higher for smokers compared to 20% to 40% in never smokers). It is also more likely to be malignant if it is symptomatic or the patient is older, male, or has had occupational exposures.

Solitary pulmonary nodules were defined by findings on CXR, but with the increased sensitivity of low-dose screening CT, up to 50% of solitary lesions are found to be associated with multiple (one to six) other, usually subcentimeter, nodules. In the Early Lung Cancer Action project, almost 7% of healthy volunteers were found to have between one and three nodules, and 25% had up to six nodules. CT scanning is necessary to characterize nodule number, location, size, margin morphology, calcification pattern, and growth rate.25 Spiral (helical) CT allows continuous scanning as the patient is moved through a scanning gantry, allowing the entire thorax to be imaged during a single breath hold (Fig. 19-14). Compared to conventional CT, this provides a superior image quality, because motion artifacts are eliminated, and improves detection of pulmonary nodules and central airway abnormalities. The shorter acquisition time of spiral CT also allows for consistent contrast filling of the great vessels, resulting in markedly improved visualization of pathologic states and anatomic variation contiguous to vascular structures. In addition, three-dimensional spiral CT images can be reconstructed for enhanced visualization of spatial anatomic relationships. Thin sections (1 to 2 mm collimation) at 1-cm intervals should be used to evaluate pulmonary parenchyma and peripheral bronchi. If the goal is to find any pulmonary metastases, thin sections at intervals of 5 to 7 mm collimation are recommended. For assessing the trachea and central bronchi, collimation of 3 to 5 mm is recommended. Providing accurate clinical history and data is of paramount importance to obtaining appropriate imaging.

CT findings characteristic of benign lesions include small size, calcification within the nodule, and stability over time.

Figure 19-14. Spiral computed tomography scan showing normal transverse chest anatomy at four levels. A. At the level of the tracheal bifurcation, the aorticopulmonary window can be seen. B. The origin of the left pulmonary artery can be seen at a level 1 cm inferior to A. C. The origin and course of the right pulmonary artery can be seen at this next most cephalad level. The left upper lobe bronchus can be seen at its origin from the left main bronchus. D. Cardiac chambers and pulmonary veins are seen in the lower thorax. AA = ascending aorta; APW = aorticopulmonary window; DA = descending aorta; LA = left ventricle; LMB = left main bronchus; LPA = left pulmonary artery; MPA = main pulmonary artery; RA = right atrium; RPA = right pulmonary artery; RV = right ventricle; SVC = superior vena cava; T = trachea.
Four patterns of benign calcification are common: diffuse, solid, central, and laminated or “popcorn.” Granulomatous infections such as tuberculosis can demonstrate the first three patterns, whereas the popcorn pattern is most common in hamartomas. In areas of endemic granulomatous disease, differentiating benign versus malignant can be challenging. Infectious granulomas arising from a variety of organisms account for 70% to 80% of this type of benign solitary nodules; hamartomas are the next most common single cause, accounting for about 10%.

CT findings characteristic of malignancy include growth over time; increasing density on CT scan (40% to 50% of partial solid lesions are malignant compared to only 15% of subcentimeter solid or nonsolid nodules); size >3 cm; irregular, lobulated, or spiculated edges; and the finding of the corona radiata sign (consisting of fine linear strands extending 4 to 5 mm outward and appearing spiculated on radiographs) (Fig. 19-15). Calcification that is stippled, amorphous, or eccentric is usually associated with cancer.

Growth over time is an important characteristic for differentiating benign and malignant lesions. Lung cancers have volume-doubling times from 20 to 400 days; lesions with shorter doubling times are likely due to infection, and longer doubling times suggest benign tumors, but can represent slower-growing lung cancer. Positron emission tomography (PET) scanning can differentiate benign from malignant nodules28; most lung tumors have increased signatures of glucose uptake, as compared with healthy tissues, and thus glucose metabolism can be measured using radio-labeled 18F-fluorodeoxyglucose (FDG). Meta-analysis estimates 97% sensitivity and 78% specificity for predicting malignancy in a nodule. False-negative results can occur (especially in patients who have AIS, MIA, or LPA, carcinoids, and tumors <1 cm in diameter), as well as false-positive results (because of confusion with other infectious or inflammatory processes).

**Metastatic Lesions to the Lung**

The cause of a new pulmonary nodule(s) in a patient with a previous malignancy can be difficult to discern.29 Features suggestive of metastatic disease are multiplicity; smooth, round borders on CT scan; and temporal proximity to the original primary lesion. One must always entertain the possibility that a single new lesion is a primary lung cancer. The probability of a new primary cancer vs. metastasis in patients presenting with solitary lesions depends on the type of initial neoplasm. The highest likelihood of a new primary lung cancer is in patients with a history of uterine (74%), bladder (89%), lung (92%), and head and neck (94%) carcinomas.

Surgical resection of pulmonary metastases has a role in properly selected patients.30 The best data regarding outcomes of resection of pulmonary metastases come from the International Registry of Lung Metastases (IRLM). The registry was established in 1991 by 18 thoracic surgery departments in Europe, the United States, and Canada and included data on 5206 patients. About 88% of patients underwent complete resection. Survival analysis at 5, 10, and 15 years (grouping all primary tumor types) was performed (Table 19-5). Multivariate analysis showed a better prognosis for patients with germ cell tumors, osteosarcomas, a disease-free interval over 36 months, and a single metastasis.31 Depicted in Fig. 19-16, survival after metastasectomy in a variety of cancers is optimal when metastatic disease is resectable, solitary, and identified 36 or more months after initial treatment. When any or all of these optimal characteristics are absent, survival progressively declines.

---

**Figure 19-15.** Computed tomography scan images of solitary pulmonary nodules. A. The corona radiata sign demonstrated by a solitary nodule. Multiple fine striations extend perpendicularly from the surface of the nodule like the spokes of a wheel. B. A biopsy-proven adenocarcinoma demonstrating spiculation. C. A lesion with a scalloped border, an indeterminate finding suggesting an intermediate probability for malignancy.

---

**Table 19-5**  
**Actuarial survival data from the International Registry of Lung Metastases**

<table>
<thead>
<tr>
<th>SURVIVAL</th>
<th>COMPLETE RESECTION (%)</th>
<th>INCOMPLETE RESECTION (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 years</td>
<td>36</td>
<td>13</td>
</tr>
<tr>
<td>10 years</td>
<td>26</td>
<td>7</td>
</tr>
<tr>
<td>15 years</td>
<td>22</td>
<td>—</td>
</tr>
</tbody>
</table>
Figure 19-16. The actuarial survival after metastasectomy is depicted for patients with various tumor types (A-F) further categorized into four groups according to resectability, solitary or multiple, the interval between primary resection and metastasectomy, and a combination of factors known in our work and in others, as follows: (1) resectable, solitary, and disease-free interval (DFI) greater than or equal to 36 months; (2) resectable, solitary, and DFI 36+ months; (3) resectable, multiple metastases, and DFI <36 months; and (4) unresectable. (Reproduced with permission from Pastorino U: The development of an international registry, J Thorac Oncol. 2010 Jun; 5(6 Suppl 2):S196-S197.)

The general principles of patient selection for metastasectomy are listed in Table 19-6. The technical aim of pulmonary metastasectomy is complete resection of all macroscopic tumors. In addition, any involved adjacent structures should be resected en bloc (i.e., chest wall, diaphragm, and pericardium). Multiple lesions and/or hilar lesions may require lobectomy. Pneumonectomy is rarely justified or employed.

Pulmonary metastasectomy can be approached through a thoracotomy or via video-assisted thoracic surgery (VATS) techniques. McCormack and colleagues reported their experience at Memorial Sloan-Kettering in a prospective study of 18 patients who presented with no more than two pulmonary metastatic lesions and underwent VATS resection.32 A thoracotomy was performed during the same operation; if palpation...
identified any additional lesions, they were resected. The study concluded that the probability that a metastatic lesion will be missed by VATS excision is 56%. Patients in the Memorial study were evaluated before the advent of spiral CT scanning, however, and it remains controversial whether metastasis resection should be performed via VATS. Proponents of VATS argue that the resolution of spiral CT scanning is so superior that prior studies using standard CT scanners are no longer relevant. Indeed, a recent study suggested that only 18% of malignant nodules would be missed using a VATS approach in the current era while another study from the United Kingdom found equivalent outcomes with regard to missed lesions and pulmonary progression comparing open and VATS approaches. To date, no prospective study using spiral CT scan has been performed to resolve this clinical dilemma.

**Primary Lung Cancer-Associated Signs and Symptoms**

Lung cancer displays one of the most diverse presentation patterns of all human maladies (Table 19-7). The wide range of symptoms and signs is related to (a) histologic features, which often help determine the anatomic site of origin in the lung; (b) the specific tumor location in the lung and its relationship to surrounding structures; (c) biologic features and the production of a variety of paraneoplastic syndromes; and (d) the presence or absence of metastatic disease. Symptoms related to the local intrathoracic effect of the primary tumor can be conveniently divided into two groups: pulmonary and nonpulmonary thoracic.

**Pulmonary Symptoms.** Pulmonary symptoms result from the direct effect of the tumor on the bronchus or lung tissue. Symptoms (in order of frequency) include cough (secondary to irritation or compression of a bronchus), dyspnea (usually due to central airway obstruction or compression, with or without atelectasis), wheezing (with narrowing of a central airway of >50%), hemoptysis (typically, blood streaking of mucus that is rarely massive; indicates a central airway location), pneumonia (usually due to airway obstruction by the tumor), and lung abscess (due to necrosis and cavitation, with subsequent pneumonia (due to airway obstruction or compression), and lung abscess (due to necrosis and cavitation, with subsequent infection).

**Nonpulmonary Thoracic Symptoms.** Nonpulmonary thoracic symptoms result from invasion of the primary tumor directly into a contiguous structure (e.g., chest wall, diaphragm, pericardium, phrenic nerve, recurrent laryngeal nerve, superior vena cava, and esophagus), or from mechanical compression of a structure (e.g., esophagus or superior vena cava) by enlarged tumor-bearing lymph nodes.

**Table 19-6**

<table>
<thead>
<tr>
<th>General principles governing appropriate selection of patients for pulmonary metastasectomy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Primary tumor must already be controlled.</td>
</tr>
<tr>
<td>2. Patient must be able to tolerate general anesthesia, potential single-lung ventilation, and the planned pulmonary resection.</td>
</tr>
<tr>
<td>3. Metastases must be completely resectable based on computed tomographic imaging.</td>
</tr>
<tr>
<td>4. There is no evidence of extrapulmonary tumor burden.</td>
</tr>
<tr>
<td>5. Alternative superior therapy must not be available.</td>
</tr>
</tbody>
</table>

**Table 19-7**

| Clinical presentation of lung cancer |
|---|---|---|
| CATEGORY | SYMPTOM | CAUSE |
| Pulmonary symptoms | Cough | Bronchus irritation or compression |
| | Dyspnea | Airway obstruction or compression |
| | Wheezing | >50% airway obstruction |
| | Hemoptysis | Tumor erosion or irritation |
| | Pneumonia | Airway obstruction |
| Nonpulmonary thoracic symptoms | Pleuritic pain | Parietal pleural irritation or invasion |
| | Local chest wall pain | Rib and/or muscle involvement |
| | Radicular chest pain | Intercostal nerve involvement |
| | Pancoast’s syndrome | Stellate ganglion, chest wall, brachial plexus involvement |
| | Hoarseness | Recurrent laryngeal nerve involvement |
| | Swelling of head and arms | Bulky involved mediastinal lymph nodes |
| | | Medially based right upper lobe tumor |

Peripherally located tumors (often adenocarcinomas) extending through the visceral pleura lead to irritation or growth into the parietal pleura and potentially to continued growth into the chest wall structures. Three types of symptoms, depending on the extent of chest wall involvement, are possible: (a) **pleuritic pain**, from noninvasive contact of the parietal pleura with inflammatory irritation or direct parietal pleural invasion; (b) **localized chest wall pain**, from deeper invasion and involvement of the rib and/or intercostal muscles; and (c) **radicular pain**, from involvement of the intercostal nerve(s). Radicular pain may be mistaken for renal colic in the case of tumors invading the inferoposterior chest wall.

Other specific nonpulmonary thoracic symptoms include:

1. **Pancoast’s syndrome.** Tumors originating in the superior sulcus (posterior apex) elicit: apical chest wall and/or shoulder pain (from involvement of the first rib and chest wall); Horner’s syndrome (unilateral enophthalmos, ptosis, miosis, and facial anhidrosis from invasion of the stellate sympathetic ganglion); and radicular arm pain (from invasion of T1, and occasionally C8, brachial plexus nerve roots).

2. **Phrenic nerve palsy.** The phrenic nerve traverses the hemithorax along the mediastinum, parallel and posterior to the superior vena cava and anterior to the pulmonary hilum. Tumors at the medial lung surface or anterior hilum can directly invade the nerve; symptoms include shoulder pain (referred), hiccups, and dyspnea with exertion because of
diaphragm paralysis. Radiographically, unilateral diaphragm elevation on chest radiograph is present; the diagnosis is confirmed by fluoroscopic examination of the diaphragm with paradoxical motion with breathing and sniffing (the “sniff” test).

3. **Recurrent laryngeal nerve palsy.** Recurrent laryngeal nerve (RLN) involvement most commonly occurs on the left side, given the hilar location of the left RLN as it passes under the aortic arch. Paralysis results from: (a) invasion of the vagus nerve above the aortic arch by a medially based left upper lobe tumor; or (b) direct invasion of the RLN by hilar tumor and/or hilar or aortopulmonary lymph node metastases. Symptoms include voice change, often referred to as hoarseness, but more typically a loss of tone associated with a breathy quality, and coughing, particularly when drinking liquids.

4. **Superior vena cava (SVC) syndrome.** As a result of bulky enlargement of involved mediastinal lymph nodes compressing or a medially based right upper lobe involving the SVC, SVC syndrome symptoms include variable degrees of swelling of the head, neck, and arms; headache; and conjunctival edema. It is seen most commonly with NEC grade IV (small cell) lung cancer.

5. **Pericardial tamponade.** Pericardial effusions (benign or malignant), associated with increasing levels of dyspnea and/or arrhythmias, and pericardial tamponade occur with direct pericardial invasion. Diagnosis requires a high index of suspicion in the setting of a medially based tumor with symptoms of dyspnea and is confirmed by CT scan or echocardiography.

6. **Back pain.** Results from direct invasion of a vertebral body and is often localized and severe. If the neural foramina are involved, radicular pain may also be present.

7. **Other local symptoms.** Dysphagia is usually secondary to external esophageal compression by enlarged lymph nodes involved with metastatic disease, usually with lower lobe tumors. Finally, dyspnea, pleural effusion, or referred shoulder pain can result from invasion of the diaphragm by a tumor at the base of a lower lobe.

**Associated Paraneoplastic Syndromes.** All lung cancer histologies are capable of producing a variety of paraneoplastic syndromes, most often from systemic release of tumor-derived biologically active materials (Table 19-8). Paraneoplastic syndromes may produce symptoms even before any local symptoms are produced by the primary tumor, thereby aiding in early diagnosis. Their presence does not influence resectability or treatment options. Symptoms often abate with successful treatment; paraneoplastic symptom recurrence may herald tumor recurrence. The majority of such syndromes are associated with grade IV NEC (small cell carcinoma), including many endocrinopathies.

<table>
<thead>
<tr>
<th>Table 19-8</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Paraneoplastic syndromes in patients with lung cancer</strong></td>
</tr>
<tr>
<td><strong>Endocrine</strong></td>
</tr>
<tr>
<td>Hypercalcemia (ectopic parathyroid hormone)</td>
</tr>
<tr>
<td>Cushing’s syndrome</td>
</tr>
<tr>
<td>Syndrome of inappropriate secretion of antidiuretic hormone</td>
</tr>
<tr>
<td>Carcinoid syndrome</td>
</tr>
<tr>
<td>Gynecomastia</td>
</tr>
<tr>
<td>Hypercalcitoninemia</td>
</tr>
<tr>
<td>Elevated growth hormone level</td>
</tr>
<tr>
<td>Elevated levels of prolactin, follicle-stimulating hormone, luteinizing hormone</td>
</tr>
<tr>
<td>Hypoglycemia</td>
</tr>
<tr>
<td>Hyperthyroidism</td>
</tr>
<tr>
<td><strong>Neurologic</strong></td>
</tr>
<tr>
<td>Encephalopathy</td>
</tr>
<tr>
<td>Subacute cerebellar degeneration</td>
</tr>
<tr>
<td>Progressive multifocal leukoencephalopathy</td>
</tr>
<tr>
<td>Peripheral neuropathy</td>
</tr>
<tr>
<td>Polymyositis</td>
</tr>
<tr>
<td>Autonomic neuropathy</td>
</tr>
<tr>
<td>Eaton-Lambert syndrome</td>
</tr>
<tr>
<td>Optic neuritis</td>
</tr>
<tr>
<td><strong>Skeletal</strong></td>
</tr>
<tr>
<td>Clubbing</td>
</tr>
<tr>
<td>Pulmonary hypertrophic osteoarthropathy</td>
</tr>
<tr>
<td><strong>Hematologic</strong></td>
</tr>
<tr>
<td>Anemia</td>
</tr>
<tr>
<td>Leukemoid reactions</td>
</tr>
<tr>
<td>Thrombocytosis</td>
</tr>
<tr>
<td>Thrombocytopenia</td>
</tr>
<tr>
<td>Eosinophilia</td>
</tr>
<tr>
<td>Pure red cell aplasia</td>
</tr>
<tr>
<td>Leukoerythroblastosis</td>
</tr>
<tr>
<td>Disseminated intravascular coagulation</td>
</tr>
<tr>
<td><strong>Cutaneous</strong></td>
</tr>
<tr>
<td>Hyperkeratosis</td>
</tr>
<tr>
<td>Dermatomyositis</td>
</tr>
<tr>
<td>Acanthosis nigricans</td>
</tr>
<tr>
<td>Hyperpigmentation</td>
</tr>
<tr>
<td>Erythema gyratum repens</td>
</tr>
<tr>
<td>Hypertrichosis lanuginosa acquista</td>
</tr>
<tr>
<td><strong>Other</strong></td>
</tr>
<tr>
<td>Nephrotic syndrome</td>
</tr>
<tr>
<td>Hypouricemia</td>
</tr>
<tr>
<td>Secretion of vasoactive intestinal peptide with diarrhea</td>
</tr>
<tr>
<td>Hyperamylasemia</td>
</tr>
<tr>
<td>Anorexia or cachexia</td>
</tr>
</tbody>
</table>

**Hypertrichosis lanuginosa acquista** may antedate the diagnosis of cancer by months. Clinically, ankle, feet, forearm, and hand tenderness and swelling are characteristic, resulting from periostitis of the fibula, tibia, radius, metacarpals, and metatarsals. Clubbing of the digits may occur in up to 30% of patients with grade IV NEC (Fig. 19-17). Plain radiographs show periosteal inflammation and elevation, while bone scans demonstrate intense but symmetric uptake in the long bones. Aspirin or nonsteroidal anti-inflammatory agents provide temporary relief; treatment requires successful tumor eradication.

2. **Hypercalcemia.** Up to 10% of patients with lung cancer will have hypercalcemia, most often due to metastatic disease. Ectopic parathyroid hormone secretion by the tumor, most often squamous cell carcinoma, is causative in up to 15%, however, and should be suspected if metastatic bone disease is not present. Symptoms of hypercalcemia include lethargy, depressed level of consciousness, nausea, vomiting,
Figure 19-17. Hypertrophic pulmonary osteoarthropathy associated with small cell carcinoma. A. Painful clubbing of the fingers. B. Painful clubbing of the toes (close-up). C. The arrows point to new bone formation on the femur.

and dehydration. Following complete tumor eradication, the calcium level will normalize. Unfortunately, tumor recurrence is extremely common and may manifest as recurrent hypercalcemia.

3. **Hyponatremia.** Characterized by confusion, lethargy, and possible seizures, hyponatremia can result from the inappropriate secretion of antidiuretic hormone from the tumor into the systemic circulation (syndrome of inappropriate secretion of antidiuretic hormone [SIADH]) in 10% to 45% of patients with grade IV NEC (small cell). It is diagnosed by the presence of hyponatremia, low serum osmolality, and high urinary sodium and osmolality. Another cause of hyponatremia can be the ectopic secretion of atrial natriuretic peptide (ANP).

4. **Cushing’s syndrome.** Autonomous tumor production of an adrenocorticotropic hormone (ACTH)-like molecule leads to rapid serum elevation of ACTH and subsequent severe hypokalemia, metabolic alkalosis, and hyperglycemia. Symptoms are primarily related to the metabolic changes while the physical signs of Cushing’s syndrome (e.g., truncal obesity, buffalo hump, striae) are unusual due to the rapidity of ACTH elevation. Diagnosis is made by demonstrating hypokalemia (<3.0 mmol/L); non-suppressible elevated plasma cortisol levels that lack the normal diurnal variation; elevated blood ACTH levels; or elevated urinary 17-hydroxycorticosteroids, all of which are not suppressible by administration of exogenous dexamethasone. Immunoreactive ACTH is present in nearly all extracts of SCLC, and a high percentage of patients with SCLC have elevated ACTH levels by radioimmunoassay, yet fewer than 5% have symptoms of Cushing’s syndrome.

5. **Peripheral and central neuropathies.** Unlike other paraneoplastic syndromes, which are usually due to ectopic secretion of an active substance, these syndromes are felt to be immune mediated. Cancer cells are thought to secrete antigens normally expressed only by the nervous system, generating antibodies leading either to interference with neurologic function or to immune neurologic destruction. Up to 16% of lung cancer patients have neuromuscular disability, and, of these, half have grade IV NEC (small cell) and 25% have squamous cell carcinomas. In patients with neurologic or muscular symptoms, central nervous system (CNS) metastases must be ruled out with CT or magnetic resonance imaging (MRI) of the head. Other metastatic disease leading to disability must also be excluded.

6. **Lambert-Eaton syndrome.** This myasthenia-like syndrome is caused by tumor secretion of immunoglobulin G (IgG) antibodies targeting voltage-gated calcium channels, which causes a neuromuscular conduction defect by decreasing the amount of acetylcholine released from presynaptic sites at the motor end plate. Symptoms, including gait abnormalities from proximal muscle weakness and impaired coordination, may actually precede radiographic evidence of the tumor. Therapy is directed at the primary tumor with resection, radiation, and/or chemotherapy. Many patients have dramatic improvement after successful therapy. For patients with refractory symptoms, treatment consists of guanidine
hydrochloride, immunosuppressive agents such as predni-
sone and azathioprine, and occasionally plasma exchange.
Unlike with myasthenia gravis patients, neostigmine is usu-
ally ineffective.

**Symptoms Associated with Metastatic Lung Cancer.** Lung
cancer metastasizes most commonly to the CNS, vertebral bod-
ies, bone, liver, adrenal glands, lungs, skin, and soft tissues.
CNS metastases are present at diagnosis in 10% of patients;
another 10% to 15% will develop CNS metastases following
diagnosis. Focal symptoms, including headache, nausea, vom-
iting, seizures, hemiplegia, and dysarthria, are common. Lung
cancer is the most common cause of spinal cord compression,
either by primary tumor invasion of an intervertebral foramen
or direct extension of vertebral metastases. Bony metastases
are identified in 25% of lung cancer patients. They are primar-
ily lytic and produce pain locally; thus, any new and localized
skeletal symptoms must be evaluated radiographically. Liver
metastases and adrenal metastases are typically asymptomatic
and usually discovered by routine CT scan. Adrenal metastasis
may lead to adrenal hypofunction. Skin and soft tissue meta-
tases occur in 8% of patients dying of lung cancer and gen-
erally present as painless subcutaneous or intramuscular masses.
Occasionally, tumor erodes through overlying skin; excision
may then be necessary for both mental and physical palliation.

**Nonspecific Cancer-Related Symptoms.** Lung cancer often
produces a variety of nonspecific symptoms such as anorexia,
weight loss, fatigue, and malaise and their presence raises con-
cern for metastatic disease.

**Lung Cancer Management**

**Role of Histologic Diagnosis and Molecular Testing.**
Establishing a clear histologic diagnosis early in the evalua-
tion and management of lung cancer is critical to effective treat-
ment. Molecular signatures are also key determinants of treat-
ment algorithms for adenocarcinoma and will likely become
important for squamous cell carcinoma as well. Currently,
differentiation between adenocarcinoma and squamous cell
carcinoma in cytologic specimens or small biopsy specimens
is imperative in patients with advanced stage disease, as treat-
ment with pemetrexed or bevacizumab-based chemotherapy is
associated with improved progression-free survival in patients
with adenocarcinoma but not squamous cell cancer. Further-
more, life-threatening hemorrhage has occurred in patients with
squamous cell carcinoma who were treated with bevacizumab.
Finally, *EGFR* mutation predicts response to *EGFR* tumor
kinase inhibitors and is now recommended as first-line therapy in
advanced adenocarcinoma. Because adequate tissue is required
for histologic assessment and molecular testing, each institution
should have a clear, multidisciplinary approach to patient evalu-
ation, tissue acquisition, tissue handling/processing, and tissue
analysis (Fig. 19-18). In many cases, tumor morphology differ-
tentiates adenocarcinoma from the other histologic subtypes. If
no clear morphology can be identified, then additional testing for
one immunohistochemistry marker for adenocarcinoma and one
for squamous cell carcinoma will usually enable differentiation.
Immunohistochemistry for neuroendocrine markers is reserved
for lesions exhibiting neuroendocrine morphology. Additional
molecular testing should be performed on all adenocarcinoma
specimens for known predictive and prognostic tumor mark-
ers (e.g., *EGFR*, *KRAS*, and *EML4-ALK* fusion gene). Ideally,
Figure 19-18. Algorithm for adenocarcinoma diagnosis in small biopsies and/or cytology. Step 1: When positive biopsies (fiberoptic bronchoscopy [FOB], transbronchial [TBBx], core, or surgical lung biopsy [SLBx]) or cytology (effusion, aspirate, washings, and brushings) show clear adenocarcinoma (ADC) or squamous cell carcinoma (SQCC) morphology, the diagnosis can be firmly established. If there is neuroendocrine (NE) morphology, the tumor may be classified as small cell carcinoma (SCLC) or non–small cell lung carcinoma (NSCLC), probably large cell neuroendocrine carcinoma (LCNEC) according to standard criteria (+ = positive, – = negative, and ± = positive or negative). If there is no clear ADC or SQCC morphology, the tumor is regarded as NSCLC—not otherwise specified (NOS). Step 2: NSCLC-NOS can be further classified based on (a) immunohistochemical stains, (b) mucin (DPAS or mucicarmine) stains, or (c) molecular data. If the stains all favor ADC-positive ADC marker(s) (i.e., TTF-1 and/or mucin positive) with negative SQCC markers, then the tumor is classified as NSCLC, favor ADC. If SQCC markers (i.e., p63 and/or CK5/6) are positive with negative ADC markers, the tumor is classified as NSCLC, favor SQCC. If the ADC and SQCC markers are both strongly positive in different populations of tumor cells, the tumor is classified as NSCLC-NOS, with a comment it may represent adenosquamous carcinoma. If all markers are negative, the tumor is classified as NSCLC-NOS. †EGFR mutation testing should be performed in (1) classic ADC, (2) NSCLC, favor ADC, (3) NSCLC-NOS, and (4) NSCLC-NOS, possible adenosquamous carcinoma. In NSCLC-NOS, if EGFR mutation is positive, the tumor is more likely to be ADC than SQCC. Step 3: If clinical management requires a more specific diagnosis than NSCLC-NOS, additional biopsies may be indicated. CD = cluster designation; CK = cytokeratin; DPAS = diastase-periodic acid Schiff; DPAS +ve = periodic-acid Schiff with diastase; EGFR = epidermal growth factor receptor; IHC = immunohistochemistry; NB = of note; TTF-1 = thyroid transcription factor-1; –ve = negative; +ve = positive. (Reproduced with permission from Travis WD, Brambilla E, Noguchi M, et al: Diagnosis of lung cancer in small biopsies and cytology: implications of the 2011 International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society classification, Arch Pathol Lab Med. 2013 May;137(5):668-684.)
Table 19-9
Evaluation of patients with lung cancer

<table>
<thead>
<tr>
<th></th>
<th>PRIMARY TUMOR</th>
<th>METASTATIC DISEASE</th>
<th>FUNCTIONAL ASSESSMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>History</td>
<td>Pulmonary</td>
<td>Weight loss</td>
<td>Ability to walk up two flights of stairs</td>
</tr>
<tr>
<td></td>
<td>Nonpulmonary thoracic</td>
<td>Malaise</td>
<td>Ability to walk on a flat surface indefinitely</td>
</tr>
<tr>
<td></td>
<td>Paraneoplastic</td>
<td>New bone pain</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Neurologic signs or symptoms</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Skin lesions</td>
<td></td>
</tr>
<tr>
<td>Physical examination</td>
<td>Voice</td>
<td>Supraclavicular node palpation</td>
<td>Accessory muscle usage</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Skin examination</td>
<td>Air flow by auscultation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Neurologic examination</td>
<td>Force of cough</td>
</tr>
<tr>
<td>Radiographic examination</td>
<td>Chest CT</td>
<td>Chest CT, PET</td>
<td>Chest CT: tumor anatomy, atelectasis</td>
</tr>
<tr>
<td>Tissue analysis</td>
<td>Bronchoscopy</td>
<td>Bone scan, head MRI, abdominal CT</td>
<td>Quantitative perfusion scan</td>
</tr>
<tr>
<td></td>
<td>Transthoracic needle aspiration and biopsy</td>
<td>Bronchoscopic lymph node FNA, Endoscopic ultrasound, Mediastinoscopy, Biopsy of suspected metastasis</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>Thoracoscopy</td>
<td>—</td>
<td>Pulmonary function tests (FEV₁, DLCO, O₂ consumption)</td>
</tr>
</tbody>
</table>

Abbreviations: CT = computed tomography; DLCO = carbon monoxide diffusion capacity; FEV₁ = forced expiratory volume in 1 second; FNA = fine-needle aspiration; MRI = magnetic resonance imaging; O₂ = oxygen; PET = positron emission tomography.

Diagnostic tissue from bronchoscopy can be obtained by one of four methods:
1. Brushings and washings for cytology
2. Direct forceps biopsy of a visualized lesion
3. Endobronchial ultrasound-guided fine-needle aspiration (FNA) of an externally compressing lesion without visualized endobronchial tumor
4. Transbronchial biopsy with fluoroscopy to guide forceps to the lesion or electromagnetic navigational bronchoscopy

Electromagnetic navigation bronchoscopy is a recent addition to the surgeon’s armamentarium for transbronchial biopsy of peripheral lung lesions. Using electromagnetic markers that create a three-dimensional image and align the recorded CT images to the patient’s true anatomy, a transbronchial catheter is advanced, and brushings, FNA, cup biopsy, and washings can be performed. Diagnostic yield using electromagnetic navigation bronchoscopy as an adjunct to standard bronchoscopy is reported as high as 80%. The approach can also be used for placement of fiducial markers for subsequent stereotactic body radiation therapy and for tattooing the perilesional region to guide subsequent video-assisted thoracoscopic resection.

Pneumothorax rates are approximately 1% to 3.5%.

For peripheral lesions (roughly the outer half of the lung), transbronchial biopsy is performed first, followed by brushings and washings. This improves diagnostic yield by disrupting the lesion with the biopsy forceps and mobilizing additional cells. For central lesions, direct forceps biopsy is often possible. For central lesions with external airway compression but no visible endobronchial lesions, endobronchial ultrasound (EBUS) is highly accurate and safe for transbronchial biopsies of both the primary tumor (when it abuts the central airways) as well as the mediastinal lymph nodes.

Image-guided transthoracic FNA (ultrasound or CT FNA) biopsy can accurately diagnose appropriately selected peripheral pulmonary lesions in up to 95% of patients. Three biopsy results are possible after image-guided biopsy procedures: malignant, a specific benign process, or indeterminate. Because false-negative rates range from 3% to 29%, further diagnostic efforts are warranted in the absence of a specific benign diagnosis (such as granulomatous inflammation or hamartoma) because malignancy is not ruled out. The primary complication is pneumothorax in as many as 30% of cases. Intrapulmonary bleeding occurs, but it rarely causes clinically significant hemoptysis or respiratory compromise.

Some groups advocate use of video-assisted thoracoscopic biopsy as the first option for diagnosis, citing superior diagnostic accuracy and low surgical risk. With VATS, the nodule can be excised with a wedge or segmental resection, if less than 3 cm, or a core-needle biopsy can be performed under direct vision for larger lesions. VATS can also provide valuable staging information, including sampling/dissection of mediastinal lymph nodes and assessing whether the primary tumor has invaded a contiguous structure (such as the chest wall or mediastinum).

Lesions most suitable for VATS are those that are located in the outer one-third of the lung. The surgeon should avoid direct manipulation of the nodule or violation of the visceral pleura overlying the nodule. In addition, the excised nodule must be extracted from the chest within a bag to prevent seeding of the chest wall. If the patient’s pulmonary reserve is adequate, the surgeon can proceed to lobectomy (either VATS or open) after frozen section diagnosis.
A thoracotomy is occasionally necessary to diagnose and stage a primary tumor. Although this occurs rarely, two circumstances may require such an approach: (a) a deep-seated lesion that yielded an indeterminate needle biopsy result or that could not be biopsied for technical reasons; or (b) inability to determine invasion of a mediastinal structure by any method short of palpation. In the circumstance of a deep-seated lesion without a diagnosis, tissue can be obtained via thoracotomy using FNA, core-needle biopsy, or excisional biopsy. Intraoperative frozen-section analysis is required; if the open biopsy frozen-section result is indeterminate, a lobectomy may be necessary in extremely rare situations. If a pneumonectomy is required to remove the lesion, a tissue diagnosis of cancer must be made before proceeding.

Assessment for Metastatic Disease

Approximately 40% of patients with newly diagnosed lung cancer present with distant metastasis. The presence of lymph node or systemic metastases may imply inoperability. As with the primary tumor, assessment for the presence of metastatic disease should begin with the history and physical examination, focusing on new bone pain, neurologic symptoms, and new skin lesions. In addition, constitutional symptoms (e.g., anorexia, malaise, and unintentional weight loss of >5% of body weight) suggest either a large tumor burden or the presence of metastases. Physical examination focuses on overall appearance, noting any evidence of weight loss such as redundant skin or muscle wasting, and a complete examination of the head and neck, including...
evaluation of cervical and supraclavicular lymph nodes and the oropharynx. This is particularly true for patients with a significant tobacco history. The skin should be thoroughly examined. Routine laboratory studies include serum levels of hepatic enzymes (e.g., serum glutamic oxaloacetic transaminase and alkaline phosphatase), and serum calcium (to detect bone metastases or the ectopic parathyroid syndrome). Elevation of either hepatic enzymes or serum calcium levels typically occurs with extensive metastases.

**Mediastinal Lymph Nodes.** Chest CT scanning facilitates assessment of mediastinal and hilar nodes for enlargement. However, a positive CT result (i.e., nodal diameter >1.0 cm) predicts actual metastatic involvement in only about 70% of lung cancer patients. Thus, up to 30% of such nodes are enlarged from noncancerous reactive causes (e.g., inflammation due to atelectasis or pneumonia secondary to the tumor). Patients should not be denied an attempt at curative resection just because of a positive CT result for mediastinal lymph node enlargement; any CT finding of metastatic nodal involvement must be confirmed histologically. The negative predictive value of normal-appearing lymph nodes by CT (lymph nodes <1.0 cm) is better than the positive predictive value of a suspicious-appearing lymph node, particularly with small squamous cell tumors. With normal-size lymph nodes and a T1 tumor, the false-negative rate is less than 10%, leading many surgeons to omit mediastinoscopy. However, the false-negative rate increases to nearly 30% with centrally located and T3 tumors. It has also been demonstrated that T1 adenocarcinomas or large cell carcinomas have a higher rate of early micrometastasis. Therefore, all such patients should undergo mediastinoscopy.

Mediastinal lymph node staging by PET scanning appears to have greater accuracy than CT scanning. PET staging of mediastinal lymph nodes has been evaluated in two meta-analyses. The overall sensitivity for mediastinal lymph node metastasis was 79% (95% confidence interval [CI] 76%–82%), with a specificity of 91% (95% CI 89%–93%) and an accuracy of 92% (95% CI 90%–94%).

In comparing PET with CT scans in patients who also underwent lymph node biopsies, PET had a sensitivity of 88% and a specificity of 91%, whereas CT scanning had a sensitivity of 63% and a specificity of 76%. Combining CT and PET scanning may lead to even greater accuracy. In one study of CT, PET, and mediastinoscopy in 68 patients with potentially operable NSCLC, CT correctly identified the nodal stage in 40 patients (59%). It understaged the tumor in 12 patients and overstaged it in 16 patients. PET correctly identified the nodal stage in 59 patients (87%). It understaged the tumor in five patients and overstaged it in four. For detecting N2 and N3 disease, the combination of PET and CT scanning yielded a sensitivity, specificity, and accuracy of 93%, 95%, and 94%, respectively. CT scan alone yielded 75%, 63%, and 68%, respectively. Studies examining combined PET-CT consistently show improved accuracy compared to PET or CT alone; accuracy for PET-CT nodal positivity confirmed by mediastinoscopy is approximately 75%, with a negative predictive value of approximately 90%. Right upper lobe lesions were more likely to have occult N2 disease than other lobes of the lung. PET-positive mediastinal lymph nodes require histologic verification of node positivity, either by EBUS-guided FNA or mediastinoscopy, to minimize the risk of undertreatment, assuming node positivity without histologic confirmation relegates the patient to, at a minimum, induction chemotherapy. If there is a suggestion of N3 disease, the patient would be incorrectly staged as having IIIB disease and would not be considered a candidate for potentially curative surgical resection.

It is important for surgeons who are managing patients with lung cancer to have a clear algorithm for invasive mediastinal staging. In general, invasive staging is underutilized, placing many patients at risk for over- or understaging and, thus, inappropriate treatment. An absolute indication for obtaining a tissue diagnosis is mediastinal lymph node enlargement greater than 1.0 cm by CT scan. There are several options for invasive mediastinal staging (Table 19-10):

1. **Endobronchial Ultrasound (EBUS)-guided transbronchial needle aspiration.** Less invasive than mediastinoscopy, EBUS enables image-guided transtracheal and transbronchial FNA cytologic samples from hilar masses and lymph nodes from level 4R and 4L, level 7, level 10, and level 11. Rapid onsite pathologic evaluation with expert cytopathology evaluation greatly increases the diagnostic accuracy of the procedure; importantly, the intraoperative evaluation will confirm whether the target lesion is being sampled and greatly facilitates acquisition of satisfactory samples for determining the morphologic diagnosis as well as sufficient material for cell block for immunohistochemistry

**TABLE 19-10**

<table>
<thead>
<tr>
<th>Techniques for invasive mediastinal staging</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Endoscopic</strong></td>
</tr>
<tr>
<td>Endobronchial ultrasound with transbronchial needle aspiration</td>
</tr>
<tr>
<td>Endoscopic ultrasound with needle aspiration</td>
</tr>
<tr>
<td>Transbronchial needle aspiration</td>
</tr>
<tr>
<td>Computed tomography–guided transthoracic needle aspiration</td>
</tr>
<tr>
<td><strong>Surgical</strong></td>
</tr>
<tr>
<td>Video-assisted mediastinoscopy</td>
</tr>
<tr>
<td>Transcervical extended mediastinal lymphadenectomy (TEMLA)</td>
</tr>
<tr>
<td>Video-assisted mediastinal lymphadenectomy (VAMLA)</td>
</tr>
<tr>
<td>Thoracoscopic transthoracic lymphadenectomy</td>
</tr>
</tbody>
</table>

**Indications for invasive mediastinal staging in lung cancer**

1. Radiographically enlarged mediastinal lymph nodes
2. Centrally located tumors
3. N1 nodal enlargement
4. Tumor size >3 cm
5. Peripheral clinical stage I tumor with nonenlarged but FDG-avid mediastinal lymph nodes

**Indications for prethoracotomy/thoracoscopy biopsy of stations 5 and 6 lymph nodes**

1. Criteria for invasive staging met and other mediastinal lymph node stations are negative (assuming patient would have induction therapy if any nodal station positive)
2. Enrollment criteria for induction therapy protocol require pathologic confirmation of N2 disease
3. Computed tomography scan shows evidence of bulky nodal metastasis or extracapsular spread that could prevent complete resection
4. Tissue diagnosis of a hilar mass or of lymph nodes causing recurrent laryngeal nerve paralysis is needed
and molecular testing. EBUS does not allow assessment of level 3, 5, or 6 nodal stations.

2. **Endoscopic ultrasound (EUS).** EUS can accurately visualize mediastinal paratracheal lymph nodes (stations 4R, 7, and 4L), paraesophageal (station 8) and inferior pulmonary ligament (station 9) lymph nodes and visualize primary lung lesions contiguous with or near the esophagus (see Fig. 19-8). Using FNA or core-needle biopsies, samples of lymph nodes or primary lesions can be obtained. Diagnostic yield is improved with intraoperative cytologic evaluation, which can be performed with the cytologist in the operating room. Limitations of EUS include the inability to visualize the anterior (pretracheal) mediastinum; thus, EUS does not replace mediastinoscopy/EBUS for complete mediastinal nodal staging. However, it may not be necessary to perform mediastinoscopy if findings on EUS are positive for N2 nodal disease, particularly if more than one station is found to harbor metastases.

3. **Cervical video-assisted mediastinoscopy.** Mediastinoscopy provides tissue sampling of all paratracheal and subcarinal lymph nodes and permits visual determination of the presence of extracapsular extension of nodal metastasis (Fig. 19-20). With complex hilar or right paratracheal primary tumors, it allows direct biopsies and assessment of invasion into the mediastinum. Mediastinoscopy is recommended for centrally located tumors, T2 and T3 primary tumors, and occasionally for T1 adenocarcinomas or large cell carcinomas (due to their higher rate of metastatic spread). Some surgeons perform mediastinoscopy in all lung cancer patients because of the poor survival associated with surgical resection of N2 disease.

4. It is important to note that EBUS or EUS can be used for initial diagnosis in enlarged lymph nodes, but the predictive value of a negative EBUS in a patient with radiographically suspicious mediastinal disease is not sufficient to accurately guide treatment. At the authors’ institutions, it is standard to begin mediastinal lymph node staging with EBUS-guided FNA of clinically suspicious mediastinal lymphadenopathy. If intraoperative rapid onsite cytologic evaluation is negative, mediastinoscopy is performed in the same operative setting to ensure accurate mediastinal staging. However, if the FNA is positive, mediastinoscopy is not performed, and the patient is referred to medical oncology for induction therapy; avoiding a pretreatment mediastinoscopy in this manner facilitates the safe performance of a postinduction mediastinoscopy for restaging of the mediastinum in patients who respond favorably to induction therapy.

5. Left video-assisted thoracoscopic lung node sampling may be needed for patients with left upper lobe tumors who have localized regional spread to stations 5 and 6 lymph nodes, without mediastinal paratracheal involvement (see Fig. 19-8). If there is a low index of suspicion for nodal metastasis, the patient can be scheduled for VATS biopsy and lobectomy under the same anesthesia; the procedure begins by sampling the level 5 and 6 nodes for frozen section, and if the nodes are negative, the anatomic lung resection is performed. If the index of suspicion is high, the VATS biopsy is performed as a separate procedure. Cervical mediastinoscopy should precede VATS biopsy, even if patients have normal paratracheal lymph nodes. Additional diagnostic evaluation of the lymph nodes in stations 5 and 6 may be unnecessary if the mediastinal lymph nodes are proven to be benign with biopsy during cervical mediastinoscopy and the preoperative CT scan suggests complete resectability of the tumor. There are, however, several indications for prethoracotomy biopsy of stations 5 and 6 lymph nodes, which are listed in Table 19-10. It is particularly important to prove that mediastinal lymph nodes are pathologically involved and not just radiographically suspicious for nodal metastasis prior to deciding that the patient is not a candidate for resection.

**Pleural Effusion.** The presence of pleural effusion on radiographic imaging should not be assumed to be malignant. Pleural effusion may be secondary to atelectasis or consolidation (seen with central tumors), cardiac dysfunction, or may be a reactive effusion. When associated with a peripherally based tumor abutting the visceral or parietal pleural surface, probability of being malignant is higher. If this is the only site concerning for metastatic disease, pathologic confirmation is mandatory. It is reasonable to start with thoracentesis, but cytology reveals malignant cells in only 50% of malignant effusions on initial thoracentesis; negative cytology 5 times is needed to have 95% certainty of a benign process. Thoracoscopy may be needed to rule out pleural metastases in select patients and is usually performed as a separate staging procedure, often with subsequent mediastinoscopy if thoracoscopy is negative for metastasis.

**Distant Metastases.** Currently, chest CT and PET are routine in the evaluation of patients with lung cancer. Integrated PET-CT scanners have become standard and have substantially improved accuracy of detection and localization of lymph node and distant metastases, as compared with independently performed PET and CT scans (Fig. 19-21). This technology overcomes the imprecise information on the exact location of focal abnormalities seen on PET and has become the standard imaging modality for lung cancer. Compared to routine chest or abdominal CT and bone scans, PET scanning detects 10% to 15% more distant metastases, but should be confirmed with MRI and/or biopsies if the patient otherwise has early-stage disease. Brain MRI should be performed when the suspicion or risk of brain metastases is increased, such as in patients with
Figure 19-21. Imaging of non–small cell lung cancer by integrated positron emission tomography (PET)-computed tomography (CT) scan. A. CT of the chest showing a tumor in the left upper lobe. B. PET scan of the chest at the identical cross-sectional level. C. Coregistered PET-CT scan clearly showing tumor invasion (confirmed intraoperatively). (Adapted with permission from Lardinois D, Weder W, Hany TF, et al. Staging of non-small-cell lung cancer with integrated positron-emission tomography and computed tomography, N Engl J Med. 2003 Jun 19;348(25):2500-2507.)

clinical stage III disease. In the absence of neurologic symptoms or signs, the probability of a negative head CT scan is 95%. Liver abnormalities that are not clearly simple cysts or hemangiomas and adrenal enlargement, nodules, or masses are further evaluated by MRI scanning and, occasionally, by needle biopsy. Adrenal adenomas have a high lipid content (secondary to steroid production), but metastases and most primary adrenal malignancies contain little if any lipid; thus, MRI is usually able to distinguish the two.

Tumor, Node, and Metastasis: Lung Cancer Staging

The staging of any tumor is an attempt to estimate the extent of disease and determine the patient’s prognosis; in a given patient, tumors are typically classified into a clinical stage and a pathologic stage. Clinical staging includes history and physical examination, radiographic test results, and diagnostic biopsy information. Therapeutic plans are generated based on clinical stage. After surgical resection of tumor and lymph nodes, postoperative pathologic stage (pTNM) is determined, providing further prognostic information.

The staging of solid epithelial tumors is based on the TNM staging system. The primary tumor “T” status provides information about tumor size and relationship to surrounding structures; the “N” status provides information about regional lymph nodes; and the “M” status provides information about the presence or absence of metastatic disease. The designation of lymph nodes as N1, N2, or N3 requires familiarity with the lymph node mapping system (see Fig. 19-8). Based on clearly delineated anatomic boundaries, accurate and reproducible localization of thoracic lymph nodes is possible, facilitating detailed nodal staging for individual patients and standardization of nodal assessment between surgeons.

Pathologic staging criteria are based on the predicted survival relative to each combination of tumor, node, and metastasis status. In 2018, the AJCC eighth edition incorporated multiple changes into the staging system for NSCLC based on analysis of survival predictors from 77,156 lung cancer patients worldwide. Table 19-11a shows the clinical and pathologic criteria changes implemented and each of the TNM descriptors currently used in staging NSCLC (Table 19-11b) and the overall stage classifications (Table 19-11c). T-staging is markedly changed, including T category designation for each centimeter in size up to 5 cm, as well as size of the invasive component in lepidic growth tumors. Visceral pleural invasion increases T-stage to T2 for patients with tumors ≤3 cm in size, and synchronous primary tumors have an added T suffix (m) in tumor staging. Metastatic disease has also been subdivided into intrathoracic, single-site extrathoracic, and multiple extrathoracic metastasis. In addition to the TNM stage, it is recommended that histologic grade, lymphovascular invasion, adequacy of resection margins and mediastinal dissection, tumor mutation status, treatment, and residual tumor after treatment also be recorded into cancer registries to facilitate evaluation of these potential predictors in future analysis of staging criteria.

Staging for small cell lung cancer (SCLC) is typically based on the extent of disease. SCLC presenting with bulky locoregional disease confined to the ipsilateral hemithorax, with no evidence for distant metastatic disease, is termed “limited” SCLC. Limited disease must be treatable within a tolerable field of radiation. Using AJCC descriptors, this includes any T stage, any N stage, without metastatic disease (M0). The only exception is when multiple lung nodules are widely spread throughout the ipsilateral lung in the same hemithorax; in these patients, the size of the involved area would preclude a “safe” radiation field. In contrast, in “disseminated” disease, tumor is beyond the ipsilateral hemithorax or widely spread within the ipsilateral lung and to distant sites. Metastases to the pleura and pericardium, with resultant effusions, are considered disseminated disease.
### Table 19-11a

Changes in Descriptors for non-small cell lung cancer comparing the 7th and 8th editions of the American Joint Committee on Cancer Staging Manual

<table>
<thead>
<tr>
<th>DESCRIPTOR</th>
<th>SEVENTH EDITION</th>
<th>EIGHTH EDITION</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T component</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 cm (pure lepidic adenocarcinoma ≤3 cm total size)</td>
<td>T1a if ≤2 cm; T1b if &gt;2–3 cm</td>
<td>Tis (AIS)</td>
</tr>
<tr>
<td>≤0.5 cm invasive size (lepidic predominant adenocarcinoma ≤3 cm total size)</td>
<td>T1a if ≤2 cm; T1b if &gt;2–3 cm</td>
<td>T1mi</td>
</tr>
<tr>
<td>≤1 cm</td>
<td>T1a</td>
<td>T1a</td>
</tr>
<tr>
<td>&gt;1–2 cm</td>
<td>T1a</td>
<td>T1b</td>
</tr>
<tr>
<td>&gt;2–3 cm</td>
<td>T1b</td>
<td>T1c</td>
</tr>
<tr>
<td>&gt;3–4 cm</td>
<td>T2a</td>
<td>T2a</td>
</tr>
<tr>
<td>&gt;4–5 cm</td>
<td>T2a</td>
<td>T2b</td>
</tr>
<tr>
<td>&gt;5–7 cm</td>
<td>T2b</td>
<td>T3</td>
</tr>
<tr>
<td>&gt;7 cm</td>
<td>T3</td>
<td>T4</td>
</tr>
<tr>
<td>Bronchus &lt;2 cm from carina</td>
<td>T3</td>
<td>T2</td>
</tr>
<tr>
<td>Total atelectasis/pneumonitis</td>
<td>T3</td>
<td>T2</td>
</tr>
<tr>
<td>Invasion of diaphragm</td>
<td>T3</td>
<td>T4</td>
</tr>
<tr>
<td>Invasion of mediastinal pleura</td>
<td>T3</td>
<td>-</td>
</tr>
<tr>
<td><strong>N component</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No assessment, no involvement, or involvement of regional lymph nodes</td>
<td>NX, N0, N1, N2, N3</td>
<td>No change</td>
</tr>
<tr>
<td><strong>M component</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metastases within the thoracic cavity</td>
<td>M1a</td>
<td>M1a</td>
</tr>
<tr>
<td>Single extrathoracic metastasis</td>
<td>M1b</td>
<td>M1b</td>
</tr>
<tr>
<td>Multiple extrathoracic metastases</td>
<td>M1b</td>
<td>M1c</td>
</tr>
</tbody>
</table>

Abbreviations: AIS, adenocarcinoma in situ; mi, minimally invasive adenocarcinoma; Tis, tumor in situ.


### Table 19-11b

American Joint Committee on Cancer Lung Cancer Staging Eighth Edition

<table>
<thead>
<tr>
<th><strong>T</strong></th>
<th><strong>PRIMARY TUMOR</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TX</strong></td>
<td>Primary tumor cannot be assessed, or tumor proven by the presence of malignant cells in sputum or bronchial washings but not visualized by imaging or bronchoscopy</td>
</tr>
<tr>
<td><strong>T0</strong></td>
<td>No evidence of primary tumor</td>
</tr>
<tr>
<td><strong>Tis</strong></td>
<td>Carcinoma <em>in situ</em></td>
</tr>
<tr>
<td></td>
<td>Squamous cell carcinoma <em>in situ</em> (SCIS)</td>
</tr>
<tr>
<td></td>
<td>Adenocarcinoma <em>in situ</em> (AIS): adenocarcinoma with pure lepidic pattern, ≤3 cm in greatest dimension</td>
</tr>
<tr>
<td><strong>T1</strong></td>
<td>Tumor ≤3 cm in greatest dimension, surrounded by lung or visceral pleura, without bronchoscopic evidence of invasion more proximal than the lobar bronchus (i.e., not in the main bronchus)</td>
</tr>
<tr>
<td><strong>T1mi</strong></td>
<td>Minimally invasive adenocarcinoma: adenocarcinoma (≤3 cm in greatest dimension) with a predominantly lepidic pattern and ≤5 mm invasion in greatest dimension</td>
</tr>
<tr>
<td><strong>T1a</strong></td>
<td>Tumor ≤1 cm in greatest dimension. A superficial, spreading tumor of any size whose invasive component is limited to the bronchial wall and may extend proximal to the main bronchus also is classified as T1a, but these tumors are uncommon.</td>
</tr>
<tr>
<td><strong>T1b</strong></td>
<td>Tumor &gt;1 cm but ≤2 cm in greatest dimension</td>
</tr>
<tr>
<td><strong>T1c</strong></td>
<td>Tumor &gt;2 cm but ≤3 cm in greatest dimension</td>
</tr>
<tr>
<td><strong>T2</strong></td>
<td>Tumor &gt;3 cm but ≤5 cm or having any of the following features:</td>
</tr>
<tr>
<td></td>
<td>• Involves the main bronchus regardless of distance to the carina, but without involvement of the carina</td>
</tr>
<tr>
<td></td>
<td>• Invades visceral pleura (PL1 or PL2)</td>
</tr>
<tr>
<td></td>
<td>• Associated with atelectasis or obstructive pneumonitis that extends to the hilar region, involving part or all of the lung</td>
</tr>
<tr>
<td></td>
<td>T2 tumors with these features are classified as T2a if ≤4 cm or if the size cannot be determined and T2b if &gt;4 cm but ≤5 cm.</td>
</tr>
</tbody>
</table>
Metastases to brain, bone, bone marrow, and the pleural and pericardial spaces are common.

**Assessment of Functional Status** Patients with potentially resectable tumors require careful assessment of their functional status and ability to tolerate either lobectomy or pneumonectomy. The surgeon should first estimate the likelihood of pneumonectomy, lobectomy, or possibly sleeve resection, based on the CT images. A sequential process of evaluation then unfolds.

A patient’s history is the most important tool for gauging risk. Specific questions regarding performance status should be routinely asked. If the patient can walk on a flat surface indefinitely, without oxygen and without having to stop and rest secondary to dyspnea, he will be very likely to tolerate lobectomy. If the patient can walk up two flights of stairs (up two standard levels), without having to stop and rest secondary to dyspnea, she will likely tolerate pneumonectomy. Finally, nearly...
all patients, except those with carbon dioxide (CO₂) retention on arterial blood gas analysis, will be able to tolerate periods of single-lung ventilation and wedge resection. Formal assessment of cardiac fitness is mandatory; use of risk scores such as the Thoracic Revised Cardiac Risk Index developed by Dr. Brunelli and colleagues provides useful prognostic information for postresection survival from early-stage lung cancer.

Current smoking status and sputum production are also pertinent. Current smokers and patients with a greater than 60 pack-year history of smoking have a significantly increased risk of postoperative pulmonary complications; heavy smokers are 2.5 times more likely to develop pulmonary complications and three times more likely to develop pneumonia compared to patients with a ≤60 pack-year history (odds ratio [OR] 2.54; 95% CI 1.28–5.04; \( P = 0.0008 \)). Impaired exchange of CO₂ is also predictive of increased risk, independent of the smoking history. For every 10% decline in percent carbon monoxide diffusion capacity (%DLCO), the risk of any pulmonary complication increased by 42% (OR 1.42; 95% CI 1.16–1.75; \( P = 0.008 \)). Risk reduction requires smoking cessation at least 8 weeks preoperatively, a requirement that is often not feasible in a cancer patient. Nevertheless, abstinence for at least 2 weeks before surgery should be encouraged. Smoking cessation on the day of surgery leads to increased sputum production and potential secretion retention postoperatively, and some authors have reported increased rates of pulmonary complications in this group. Patients with chronic daily sputum production will have more problems postoperatively with retention and atelectasis; they are also at higher risk for pneumonia. Sputum culture, antibiotic administration, and bronchodilators may be warranted preoperatively.

Pulmonary function studies are routinely performed when any resection greater than a wedge resection will be performed. Of all the measurements available, the two most valuable are forced expiratory volume in 1 second (FEV₁) and carbon monoxide diffusion capacity (DLCO). General guidelines for the use of FEV₁ in assessing the patient’s ability to tolerate pulmonary resection are as follows: greater than 2.0 L can tolerate pneumonectomy, and greater than 1.5 L can tolerate lobectomy. It must be emphasized that these are guidelines only. It is also important to note that the raw value is often imprecise because normal values are reported as “percent predicted” based on corrections made for age, height, and gender. For example, a raw FEV₁ value of 1.3 L in a 62-year-old, 75-inch (190-cm) male has a percent predicted value of 30% (because the normal expected value is 4.31 L); in a 62-year-old, 62-inch female, the predicted value is 59% (normal expected value 2.21 L). The male patient is at high risk for lobectomy, while the female could potentially tolerate pneumonectomy.

To calculate the predicted postoperative value for FEV₁ or DLCO, the percent predicted value of FEV₁ or DLCO is multiplied by the fraction of remaining lung after the proposed surgery. For example, with a planned right upper lobectomy, a total of three segments will be removed. Therefore, three of a total 20 segments will leave the patient with \((20 – \frac{3}{20}) \times 100 = 85\%\) of their original lung capacity. In the two patients mentioned earlier, the man will have a predicted postoperative FEV₁ of 30% \(\times 0.85 = 25\%\), whereas the woman will have a predicted postoperative FEV₁ of 50%. Percent predicted value of less than 50% for either FEV₁ or DLCO correlates with risk for postoperative complications, particularly pulmonary complications; the risk of complications increases in a stepwise fashion for each 10% decline. Figure 19-22 shows the relationship between predicted postoperative DLCO and estimated operative mortality.

Quantitative perfusion scanning is used in select circumstances to help estimate the functional contribution of a lobe or whole lung. Such perfusion scanning is most useful when the impact of a tumor on pulmonary physiology is difficult to discern. With complete collapse of a lobe or whole lung, the impact is apparent, and perfusion scanning is usually unnecessary. Figure 19-23 shows a tumor with significant right main stem airway obstruction with associated atelectasis and volume loss of the right lung. At presentation, the patient was dyspneic with ambulation, and the FEV₁ was 1.38 L. Six months prior, this patient could walk up two flights of stairs without dyspnea. The surgeon can anticipate that the patient will tolerate pneumonectomy because the lung is already not functioning due to main stem airway obstruction, and may, in fact, be contributing to a shunt. However, with centrally located tumors associated with partial obstruction of a lobar or main bronchus or of the pulmonary artery, perfusion scanning may be valuable in predicting the postoperative result of resection. For example, if the quantitative perfusion to the right lung is measured to be 21% (normal is 55%) and the patient’s percent predicted FEV₁ is 60%, the predicted postoperative FEV₁ after a right pneumonectomy would be 60% \(\times 0.79 = 47\%\), indicating the ability to tolerate pneumonectomy. If the perfusion value is 55%, the predicted postoperative value would be 27%, and pneumonectomy would pose a significantly higher risk.

It is not uncommon to encounter patients with significant reductions in their percent predicted FEV₁ and DLCO whose history shows a functional status that is inconsistent with the pulmonary function tests. In these circumstances, exercise testing that yields maximal oxygen consumption (\(\dot{V}O_{2 \text{max}}\)) has emerged as a valuable decision-making technique to help patients with abnormal FEV₁ and DLCO (Table 19-12). Values <10 mL/kg/min are associated with a 26% mortality after major pulmonary...
resection compared to only 8.3% with \( \dot{V}_{O_2} \text{max} \geq 10 \text{ mL/kg/min.} \)

Values >15 mL/kg/min generally indicate the patient’s ability to tolerate pneumonectomy.

The risk assessment of a patient is an amalgam of clinical judgment and data that must be integrated with the experienced clinician’s sense of the patient and with the patient’s attitude toward the disease and toward life. Figure 19-24 provides a useful algorithm for determining suitability for lung resection.44

### Lung Cancer Treatment

#### Grade IV NEC (Small Cell) Lung Carcinoma.

In rare circumstances where SCLC presents as an isolated lung lesion, lobectomy followed by chemotherapy is warranted after surgical mediastinal staging has confirmed the absence of N2 disease. Often, ultrasound-guided FNA provides a definitive positive diagnosis and more invasive approaches are not needed. However, less than 5% are stage I, and there is no benefit from surgical resection for more advanced-stage disease; treatment is chemotherapy with or without radiation therapy depending on the extent of disease and the patient performance status.

#### Early-Stage Non–Small Cell Lung Cancer.

Early-stage disease includes T1 and T2 tumors (with or without N1 nodal involvement) and T3 tumors (without N1 nodal involvement). This group represents a small but increasing proportion of the total number of patients diagnosed with lung cancer each year (approximately 16% of an estimated 222,500 patients in 2017).18 Surgical resection is the current standard, ideally accomplished by video-assisted lobectomy or pneumonectomy, depending on the tumor location.

Despite the term “early-stage,” the overall 5-year survival rate for all localized lung cancer is 55% and only 26% when regional metastasis was present between 2004 and 2009.45 Median survival for untreated patients with stage IA NSCLC is 14 months, and 5-year survival rate is 22%.46 After surgical resection of postoperative pathologic stage IA disease, 5-year survival is better than with no treatment, but still only 67%.41 Survival declines with higher stages. Advanced age at diagnosis, male sex, low socioeconomic status, nonsurgical treatment, and poor histologic grade are associated with increased mortality risk on multivariate analysis.45

Depending on tumor size and location, lobectomy, sleeve lobectomy, and occasionally pneumonectomy, with mediastinal lymph node dissection or sampling, are appropriate for patients with clinical early-stage disease. Sleeve resection is performed for tumors located at airway bifurcations when an adequate bronchial...
Figure 19-24. Algorithm for preoperative evaluation of pulmonary function and reserve prior to resectional lung surgery. CPET = cardiopulmonary exercise test; CT = computed tomographic scan; CXR = chest radiograph; DLCO = carbon monoxide diffusion capacity; FEV₁ = forced expiratory volume in 1 second; %ppo = percent predicted postoperative lung function; VO₂max = maximum oxygen consumption. (Modified with permission from Colice GL, Shafazand S, Griffin JP: Physiologic evaluation of the patient with lung cancer being considered for resectional surgery: ACCP evidenced-based clinical practice guidelines (2nd edition), Chest. 2007 Sep;132(3 Suppl):161S-177S.)
margin cannot be obtained by standard lobectomy. Pneumonectomy is rarely performed; primary indications for pneumonectomy in early-stage disease include large central tumors involving the distal main stem bronchus and inability to completely resect involved N1 lymph nodes. The latter circumstance occurs with bulky adenopathy or with extracapsular nodal spread.

Management of Early-Stage Lung Cancer in the High-Risk Patient. Lobectomy may not be an option for some patients with early-stage disease, due to poor cardiopulmonary function or other comorbid illnesses. The ultimate decision that a patient is inoperable, both with regard to the patient’s ability to tolerate surgery and the likelihood of successful resection, should be accepted only after evaluation by an expert surgeon. Surgeons with limited expertise, when faced with a complicated patient, should refer the patient to a high-volume center for further evaluation if they are unable to offer the patient surgical resection in their own center.

Rationale for Limited Resection in Early-Stage Lung Cancer. Limited resection, defined as segmentectomy or wedge resection, is a viable option for achieving local control in high-risk patients. Historically, limited resection with wedge or segmentectomy has been considered a compromise operation due to unacceptably high rates of local recurrence and concerns for worse survival. Subsequent meta-analysis of the literature shows that the difference in death rate is likely negligible (Table 19-13). The high rates of local recurrence demonstrated by Ginsberg and others, however, remain a significant concern and continue to restrict the use of limited resection for early-stage lung cancer to the high-risk patient.

With the recent publication of a 20% reduction in lung cancer mortality with screening CT scans in high-risk populations, the topic of limited resection is again the subject of intensive review. Studies investigating anatomic segmentectomy (or extended wedge resection) with hilar and mediastinal lymph node dissection suggest that close attention to the ratio of surgical margin to tumor diameter and a careful assessment of the lymph nodes substantially reduce local recurrence. Recurrence rates were 6.2%, comparable to rates associated with lobectomy, when the margin-to-tumor diameter ratio exceeded 1, compared to 25% if the margin-to-tumor diameter ratio was less than 1.5. In most centers, this requires use of a thoracotomy, although increasing experience with VATS in high-volume centers shows that limited resection is safe and feasible, with perioperative adverse outcomes that are comparable to lobectomy.

Rationale for Tumor Ablation in the Management of Primary Lung Cancer. Limited resection, by definition, requires that the patient has sufficient cardiopulmonary reserve to undergo a general anesthesia and loss of at least one pulmonary segment. For the high-risk or nonoperable patient, as determined by experience pulmonary surgeons, tumor ablation techniques have been developed for treatment of early-stage lung cancers.

Current limitations of this approach include the absence of nodal staging, lack of tissue for molecular profiling, chemoresistance, or sensitivity testing, concerns about definitions of locoregional recurrence, and a lack of uniformity across centers. Surgeons typically define locoregional recurrence as tumor growth within the operative field, including resectable lymph nodes, whereas local recurrence after ablation is most commonly defined as tumor growth within the field of treatment. Despite the fact that in-transit or lymph node metastases are present in up to 27% of clinically stage I NSCLCs at resection, any tumor growth outside the field of ablative treatment is not be considered treatment failure. Despite these limitations, tumor ablative strategies are increasingly proposed as viable alternatives to surgical resection, even in potentially operable patients. While premature, ablative techniques may ultimately be shown to have efficacy equivalent to lobectomy for the primary treatment of very small peripheral early-stage lung cancers. Multidisciplinary collaboration between thoracic surgery, interventional radiology/pulmonology, and radiation oncology is required to ensure that development of these ablative techniques occurs through properly designed and well-controlled prospective studies and will ensure that patients receive the best available therapy, regardless of whether it is surgical resection or ablative therapy.

The two most commonly applied ablation techniques are radiofrequency ablation and stereotactic body radiotherapy.

1. Radiofrequency ablation. Radiofrequency ablation is performed using either monopolar or bipolar delivery of electrical current to electrodes placed within the tumor tissue. In lung tumors, the electrodes are typically inserted into the tumor mass under CT guidance. An electrical current is delivered; the current is converted by means of friction into heat, which quickly leads to immediate and irreparable tissue destruction in the tissue surrounding the electrode. The efficacy of radiofrequency ablation for controlling the primary tumor and improving survival in poor operative candidates (either due to significant comorbid diseases precluding general anesthesia or poor pulmonary function excluding lung resection) is safe and feasible for peripheral lung nodules. In tumors <3.5 cm, the rate of radiographic resolution of tumor is up to 80%, and cancer-specific survival at 2 years was approximately 90%, indicating excellent local control of the primary site. It has become the preferred modality for small peripheral tumors over standard external-beam radiation in centers where the technique is available.

Radiofrequency ablation is an excellent modality for the patient at risk for adverse outcomes with pulmonary resection or for patients who refuse surgery, and surgeons should have an algorithm for determining which patients are optimal for this modality (see Fig. 19-24). Target lesions larger than 5 cm, tumor abutting the hilum, associated malignant pleural or pericardial effusion, greater than three lesions in one lung, and the presence of pulmonary hypertension are all contraindications to radiofrequency ablation. Proximity to a large vessel is a contraindication not only due to the risk of massive bleeding, but also because large blood vessels act as a heat sink and lethal cellular temperatures are less likely to be achieved. For these patients, stereotactic body radiotherapy may provide local tumor control with less risk of major complications. Combination therapy with either external-beam radiotherapy or stereotactic body radiotherapy is also under investigation.

2. Stereotactic body radiotherapy. Stereotactic body radiotherapy applies highly focused, high-intensity, three-dimensional conformal radiation to the target lesion over a few sessions. Tumor motion quantification and image guidance technologies have significantly improved the delivery of radiation with high levels of precision.
Table 19-13
New classification system for lung adenocarcinoma

<table>
<thead>
<tr>
<th>STUDY OR SUBGROUP</th>
<th>log [HAZARD RATIO]</th>
<th>SE</th>
<th>SEGMENTECTOMY TOTAL</th>
<th>LOBECTOMY TOTAL</th>
<th>WEIGHT</th>
<th>HAZARD RATIO IV, RANDOM, 95% CI</th>
<th>HAZARD RATIO IV, RANDOM, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bando 2002</td>
<td>-0.315</td>
<td>0.349</td>
<td>74</td>
<td>132</td>
<td>3.3%</td>
<td>0.73 [0.37, 1.45]</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>-0.21</td>
<td>0.29081</td>
<td>32</td>
<td>32</td>
<td>4.8%</td>
<td>0.81 [0.46, 1.43]</td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>0.368</td>
<td>0.2104</td>
<td>4240</td>
<td>11520</td>
<td>9.2%</td>
<td>1.39 [0.92, 2.11]</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>0.1989</td>
<td>0.2806</td>
<td>32</td>
<td>77</td>
<td>5.1%</td>
<td>1.22 [0.70, 2.11]</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>-0.08338</td>
<td>0.50765</td>
<td>31</td>
<td>55</td>
<td>1.6%</td>
<td>0.92 [0.34, 2.49]</td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>0.131</td>
<td>0.4158</td>
<td>54</td>
<td>147</td>
<td>2.3%</td>
<td>1.14 [0.50, 2.58]</td>
<td></td>
</tr>
<tr>
<td>-0.2357</td>
<td>0.21683</td>
<td>106</td>
<td>78</td>
<td>8.6%</td>
<td>0.79 [0.52, 1.21]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>0.1011</td>
<td>0.5732</td>
<td>46</td>
<td>77</td>
<td>1.2%</td>
<td>1.11 [0.36, 3.40]</td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>0.0773</td>
<td>0.4385</td>
<td>74</td>
<td>159</td>
<td>2.1%</td>
<td>1.08 [0.46, 2.55]</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>0.46</td>
<td>0.2482</td>
<td>102</td>
<td>117</td>
<td>6.6%</td>
<td>1.58 [0.97, 2.58]</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>-1.0435</td>
<td>0.7071</td>
<td>17</td>
<td>17</td>
<td>0.8%</td>
<td>0.35 [0.09, 1.41]</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>0.239</td>
<td>1</td>
<td>38</td>
<td>289</td>
<td>0.4%</td>
<td>1.27 [0.18, 9.02]</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>-0.1165</td>
<td>0.7168</td>
<td>68</td>
<td>104</td>
<td>0.8%</td>
<td>0.89 [0.22, 3.63]</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>-0.1632</td>
<td>0.3595</td>
<td>67</td>
<td>273</td>
<td>3.1%</td>
<td>0.85 [0.42, 1.72]</td>
<td></td>
</tr>
<tr>
<td>-0.02</td>
<td>0.2569</td>
<td>119</td>
<td>131</td>
<td>2.3%</td>
<td>1.37 [0.60, 3.11]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3148</td>
<td>0.4184</td>
<td>113</td>
<td>131</td>
<td>2.3%</td>
<td>1.37 [0.60, 3.11]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>1.0402</td>
<td>0.9158</td>
<td>49</td>
<td>150</td>
<td>0.5%</td>
<td>2.83 [0.47, 17.03]</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>0.2852</td>
<td>1.6735</td>
<td>56</td>
<td>178</td>
<td>0.1%</td>
<td>1.33 [0.05, 35.35]</td>
<td></td>
</tr>
<tr>
<td>-0.0202</td>
<td>0.25</td>
<td>43</td>
<td>95</td>
<td>6.5%</td>
<td>0.98 [0.60, 1.60]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7133</td>
<td>0.3087</td>
<td>98</td>
<td>383</td>
<td>4.3%</td>
<td>0.49 [0.27, 0.90]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>0.5481</td>
<td>0.5</td>
<td>66</td>
<td>103</td>
<td>1.6%</td>
<td>1.73 [0.65, 4.61]</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>-0.6162</td>
<td>0.89796</td>
<td>20</td>
<td>57</td>
<td>0.5%</td>
<td>0.54 [0.09, 3.14]</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>0.19886</td>
<td>0.73469</td>
<td>90</td>
<td>124</td>
<td>0.8%</td>
<td>1.22 [0.29, 5.15]</td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>0.3001</td>
<td>0.4056</td>
<td>153</td>
<td>277</td>
<td>2.5%</td>
<td>1.35 [0.61, 2.99]</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>0.239</td>
<td>0.27562</td>
<td>162</td>
<td>2599</td>
<td>5.3%</td>
<td>1.27 [0.74, 2.18]</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>-0.1625</td>
<td>0.19388</td>
<td>39</td>
<td>81</td>
<td>10.8%</td>
<td>0.85 [0.58, 1.24]</td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.00; Chi² = 25.04, df = 26 (P = 0.52); I² = 0%

Overall effect: Z = 0.68 (P = 0.50)

*HR for overall survival impact of operative approach (segmentectomy versus lobectomy) of stage I NSCLC patients. The pooled HR displayed in this figure when compared with segmentectomy, but there was not a significant benefit of lobectomy on HR of stage I patients (7-21) (HR 1.04; 95% CI, 0.92–1.18, P = 0.50) (22–33).

Abbreviations: HR = hazard ratio; NSCLC = non–small cell lung cancer; CI = confidence interval; df = degree of freedom; SE = standard error.


Favours Segmentectomy Favours Lobectomy
to the target lesion. This accuracy is important because the lung is extremely sensitive to radiation injury and the majority of patients with early-stage lung cancer who are currently considered candidates for ablative therapy have marginal lung function; excessive injury to normal surrounding lung tissue is not desirable. Importantly, these techniques allow the safe delivery of up to 66 Gy of radiation to the target tumors without exceeding the maximum-tolerated dose. A phase II North American multicenter study recently demonstrated the safety and efficacy of this approach in 59 nonoperable patients.62

Patients with biopsy-proven, node-negative peripheral NSCLCs less than 5 cm in diameter (T1 or T2) were treated with stereotactic body radiotherapy after they were deemed inoperable, based on coexisting medical conditions, by a thoracic surgeon and/or pulmonologist. Primary tumor control was excellent; at 3 years, 97.6% were deemed to have primary tumor control by the authors, and 90.6% had local control. However, it is important to note that primary tumor failure was defined specifically as at least a 20% increase in the longest diameter of the gross tumor volume by CT scan and evidence of tumor viability either by biopsy confirming carcinoma or by demonstration of FDG avidity on PET scan. For viability to be confirmed with PET scan, the uptake was required to be of similar intensity to the pretreatment staging PET scan. Failure beyond a 1.5- to 2-cm margin around the primary tumor volume was considered local failure. Failure in regional node basins was seen in two patients. When compared to locoregional control rates of approximately 6.5% with limited resection, the 3-year locoregional recurrence was higher at 12.8%.

3. Patient selection for stereotactic body radiotherapy, as with limited resection and radiofrequency ablation, is important. Because the radiation field is so precise, patients with severe emphysema and chronic obstructive pulmonary disease can be safely treated without significant concern for worsening lung function. However, patients with central tumors near the mediastinum and hilum have increased incidence of significant hypoxia, hemoptysis, atelectasis, pneumonitis, and reduced pulmonary function. In the multicenter trial detailed earlier, treated tumors were required to be greater than 2 cm from the proximal bronchial tree in all directions (which they defined as the distal 2 cm of the trachea, carina, and named major lobar bronchi up to their first bifurcation).62

**Rationale for Chemotherapy in the Management of Early-Stage NSCLC.** The role of chemotherapy in early-stage (stages I and II) NSCLC is evolving, with several prospective phase 2 studies having shown a potential benefit. Initial concerns that induction chemotherapy may result in increased perioperative morbidity or mortality appear to be unwarranted, as the incidence of perioperative morbidity and mortality is not different between the two groups, except in patients undergoing right-sided pneumonectomy after induction chemotherapy. As shown in Table 19-14, an absolute survival benefit of 4% to 7% can be realized using induction for all stages of lung cancer, and in situations where use of adjuvant chemotherapy is anticipated, induction chemotherapy is an acceptable alternative.

### Table 19-14

<table>
<thead>
<tr>
<th>STAGE</th>
<th>5-YEAR SURVIVAL (%)</th>
<th>ABSOLUTE BENEFIT (%)</th>
<th>NEW 5-YEAR SURVIVAL (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IA</td>
<td>75</td>
<td>4</td>
<td>79</td>
</tr>
<tr>
<td>IB</td>
<td>55</td>
<td>6</td>
<td>61</td>
</tr>
<tr>
<td>IIA</td>
<td>50</td>
<td>7</td>
<td>57</td>
</tr>
<tr>
<td>IIB</td>
<td>40</td>
<td>7</td>
<td>47</td>
</tr>
<tr>
<td>IIIA</td>
<td>15–35</td>
<td>6–7</td>
<td>21–42</td>
</tr>
<tr>
<td>IIIB</td>
<td>5–10</td>
<td>3–5</td>
<td>8–15</td>
</tr>
</tbody>
</table>

National Comprehensive Cancer Network guidelines currently recommend observation for T1a (≤1 cm), T1b (>1–2 cm), and T1c (>2–3 cm), node-negative, completely resected NSCLCs (T1abcNOM0). For patients with larger tumors (T2a tumor >3–4 cm; T2b tumor >4–5 cm) that are node-negative, it is recommended that chemotherapy be considered in high-risk patients, ideally in the setting of a clinical trial. High-risk tumor characteristics include poorly differentiated tumors, moderately to poorly differentiated lung neuroendocrine tumors, vascular invasion, resection limited to wedge resection only, tumors >4 cm in size, visceral pleural involvement, and when lymph node sampling at the time of resection was incomplete (Nx).

**Evaluation and Management of Locally Advanced NSCLC.** Five-year relative survival in patients with locoregional disease is 28%, but there is significant heterogeneity within the group. Stage III disease includes patients with small tumors that have metastasized to the mediastinal lymph nodes as well as large tumors (>7 cm), and tumor invading unresectable structures or the major carina with no nodal metastasis at all. Patients with clinically evident N2 disease (i.e., bulky adenopathy present on CT scan or mediastinoscopy, with lymph nodes often replaced by tumor) have a 5-year survival rate of 5% to 10% with surgery alone. In contrast, patients with microscopic N2 disease discovered incidentally in one lymph node station after surgical resection have a 5-year survival rate that may be as high as 30%. As a result, many surgeons and oncologists differentiate between microscopic and bulky N2 lymphadenopathy and the number of involved N2 nodal stations in determining whether to proceed with resection following induction therapy. It is generally accepted that surgical resection is appropriate for patients with a single-station metastasis with a single lymph node smaller than 3 cm, although randomized trials specifically investigating resection following induction therapy for patients with single-station microscopic disease have not yet been performed.

Histologic confirmation of N2 nodal metastases is imperative; false-positive findings on PET scan are unacceptably high, and reliance on this modality will lead to significant undertreatment of patients with earlier stage cancers. This is particularly true in regions with high incidence of granulomatous diseases.
When N2 nodes are found, incidentally, to harbor metastasis at the time of planned anatomic lung resection, the decision to proceed with resection varies depending on surgeon preference; it is acceptable to either proceed with anatomic resection and mediastinal lymph node sampling/dissection or to stop the procedure, refer the patient for induction therapy, and reevaluate for resection after induction therapy is completed.

When histologically confirmed metastases are found during preoperative staging evaluation, patients should be referred for induction chemotherapy; patients in whom the mediastinal nodes are sterilized by induction therapy have a better prognosis, and surgical resection is generally warranted as part of a multimodal approach. As with preinduction evaluation, histologic confirmation of persistent N2 disease after induction therapy is imperative; patients should not be denied surgical resection following induction chemotherapy based on radiographic evidence for N2 disease because the survival for resected NSCLC is significantly better than with definitive chemotherapy.

**Surgery in T4 and Stage IV Disease** Surgery is occasionally appropriate for highly selected patients with tumors invading the SVC, carinal or vertebral body involvement, or satellite nodules in the same lobe (T3, N0-1, M0) or in T4, N0-1 tumors with limited invasion into the mediastinum, heart, great vessels, trachea, recurrent laryngeal nerve, esophagus, vertebral body, diaphragm or carina through direct extension. Surgery generally does not have a role in the care of patients with any tumor with N3 disease or T4 tumors with N2 disease. Survival rates remain extremely low for these patients. Similarly, the treatment of patients with stage IV disease is chemotherapy. However, on occasion, patients with a single site of metastasis are encountered, particularly with adenocarcinomas presenting with a solitary brain metastasis. In this highly select group, 5-year survival rates of 10% to 15% can be achieved with surgical excision of the brain metastasis and the primary tumor, assuming it is early stage.

**Surgery for Management of Pancoast’s Tumor** Carcinoma arising in the extreme apex of the chest with associated arm and shoulder pain, atrophy of the muscles of the hand, and Horner’s syndrome presents a unique challenge to the surgeon. Any tumor of the superior sulcus, including tumors without evidence for involvement of the neurovascular bundle, is now commonly known as Pancoast’s tumors, after Henry Pancoast who described the syndrome in 1932. The designation is reserved for tumors involving the parietal pleura or deeper structures overlying the first rib. Chest wall involvement at or below the second rib is not a Pancoast’s tumor. Treatment is multidisciplinary; due to the location of the tumor and involvement of the neurovascular bundle that supplies the ipsilateral extremity, preserving postoperative function of the extremity is critical. For this reason, resection should only be performed in patients who are proven negative for mediastinal lymph node involvement. Survival with N2 positive nodes is poor, and the morbidity and mortality associated with surgical resection are high. If bulky lymphadenopathy is present, EBUS- or EUS-guided FNA may prove nodal involvement. However, a negative FNA is not sufficient for proving the absence of mediastinal involvement and should be followed by mediastinoscopy to ensure accurate and complete evaluation of the mediastinum.

Because Pancoast’s tumors have high rates of local recurrence and incomplete resection, induction chemoradiotherapy followed by surgery is recommended. This treatment regimen was well tolerated in a study performed by the Southwest Oncology Group, with 95% of patients completing induction treatment. Complete resection was achieved in 76%. Five-year survival was 44% overall and 54% when complete resection was achieved. Disease progression with this regimen was predominantly at distant sites, with the brain being the most common. The current treatment algorithm for Pancoast’s tumors is presented in Fig. 19-25.

Surgical excision is performed via thoracotomy with en bloc resection of the chest wall and vascular structures and anatomic lobectomy. A portion of the lower trunk of the brachial plexus and the stellate ganglion are also typically resected. With chest wall involvement, en bloc chest wall resection, along with lobectomy, is performed, with or without chest wall reconstruction.

For small rib resections or those posterior to the scapula, chest wall reconstruction is usually unnecessary. Larger defects (two rib segments or more) are usually reconstructed with Gore-Tex to provide chest wall contour and stability. En bloc resection is also used for other locally advanced tumors (T3) with direct invasion of the adjacent chest wall, diaphragm, or pericardium. If a large portion of the pericardium is removed, reconstruction with thin Gore-Tex membrane will be required to prevent cardiac herniation and venous obstruction.

**Preoperative (Induction) Chemotherapy for NSCLC** The use of chemotherapy before anatomic surgical resection has a number of potential advantages:

1. The tumor’s blood supply is still intact, allowing better chemotherapy delivery and avoiding tumor cell hypoxia (in any residual microscopic tumor remaining postoperatively), which would increase radioresistance.
2. The primary tumor may be downstaged, enhancing resectability.
3. Patients are better able to tolerate chemotherapy before surgery and are more likely to complete the prescribed regimen than after surgery.
4. It functions as an in vivo test of the primary tumor’s sensitivity to chemotherapy.
5. Response to chemotherapy can be monitored and used to guide decisions about additional therapy.
6. Systemic micrometastases are treated.
7. It identifies patients with progressive disease/nonresponders and spares them a pulmonary resection.

Potential disadvantages include:

1. There is a possible increase in the perioperative complication rate in patients requiring right pneumonectomy after induction chemotherapy.
2. While the patient is receiving chemotherapy, potentially curative resection is delayed; if the patient does not respond, this delay could result in tumor spread.

In stage IIIA N2 disease, the response rates to induction chemotherapy are high, in the range of 70%. The treatment is generally safe, as it does not cause a significant increase in perioperative morbidity. Two randomized trials have now compared surgery alone for patients with N2 disease to preoperative chemotherapy followed by surgery. Both trials were stopped before complete accrual because of a significant increase in survival for the chemotherapy arm.

The initially observed survival differences have been maintained for up to 3 years and beyond (5-year data not shown). Given these results, induction chemotherapy with
Confirm T3–4, N0-1 M0 NSCLC
No evidence for metastatic or N2 nodal disease

Tumor progression or poor performance status
Definitive chemoradiotherapy

Concurrent induction chemotherapy (Cisplatin/Etoposide)
And radiotherapy: 45 Gy over 5 weeks

CT chest/upper abdomen
MRI/MRA of vessels/brachial plexus
Mediastinoscopy
Brain CT or MRI and PET scan

Reassessment performance score, physiologic reserve, tumor response
Radiographic evaluation: CT scans of the chest, upper abdomen, and brain. PET scan for metastases

Tumor stable/regression; good to excellent performance status
Additional chemotherapy as tolerated

Poor performance status

Thoracotomy, en bloc chest wall resection, lobectomy, chest wall with reconstruction

**Figure 19-25.** Treatment algorithm for Pancoast’s tumors. CT = computed tomography; MRA = magnetic resonance angiography; MRI = magnetic resonance imaging; NSCLC = non–small cell lung cancer; PET = positron emission tomography.

cisplatin-based regimen (two to three cycles) has become standard for patients with N2 disease. Table 19-15 summarizes the findings of a systematic review and meta-analysis reporting response rates, progression-free survival, and overall survival after induction chemotherapy followed by surgical resection.

**Postoperative (Adjuvant) Chemotherapy for NSCLC**

Postoperative adjuvant chemotherapy was previously thought to confer no benefit based on multiple prospective randomized trials, in part because patients who had undergone thoracotomy and lung resection had difficulty tolerating the adjuvant regimens. More recently, however, newer, more effective agents have shown promise, and adjuvant therapy is better tolerated after minimally invasive lung resection (i.e., VATS or robotic anatomic resection). Targeted therapies, which have been shown to be beneficial in advanced-stage lung cancer, are of particular interest.

Any patient with nodal metastasis (N1 or N2) or with T3 tumors (defined as tumors >5 to ≤7 cm or separate tumor in same lobe or direct invasion of chest wall [includes parietal pleura and superior sulcus/parietal pericardium/phrenic nerve]) should receive adjuvant chemotherapy if they are able to tolerate the regimen. In the situation where the margins of resection are positive, re-resection is recommended. If not possible, concurrent chemoradiation is recommended for macroscopic residual tumor and sequential chemoradiation for microscopic residual tumor.

**Definitive Nonsurgical Treatment for NSCLC.** Recent advances in targeted therapies have changed the management of advanced NSCLC from a generalized, platinum-based approach to one in which molecular analysis and targeted, personalized therapies are now standard of care. It is now mandatory that the pathologist clearly differentiate between squamous cell carcinoma and adenocarcinoma because the therapeutic options are different and use of bevacizumab, while beneficial in patients with adenocarcinoma, has been found to cause excessive pulmonary hemorrhage in patients with squamous histology. For the surgeon, this requirement translates into a much more aggressive approach to tissue diagnosis. At our institution, the cytopathologist provides onsite rapid assessment of the fine-needle aspirate to determine whether tumor cells are present and confirm that sufficient tumor cells are present to enable molecular testing. This has increased the number of passes performed during an EBUS-guided FNA or during CT-guided aspiration of a pulmonary or intrathoracic lesion; typically, an additional two to three passes are made for cell block material after confirming the presence of tumor cells in the target area. When insufficient cells are obtained for molecular testing, despite having a diagnosis, additional sampling is warranted; this is mandatory in patients with adenocarcinoma and likely to become necessary for other non–small cell histologic types as advances in targeted therapies become available for clinical use. Acquiring adequate tissue for diagnosis may require mediastinoscopy or VATS; close communication between the oncologist, surgeon, pathologist, and patient is needed to ensure that the benefits to the patient clearly outweigh the risks and that results obtained through more aggressive diagnostic measures are needed to direct subsequent care.
### Table 19-15

<table>
<thead>
<tr>
<th>Trial</th>
<th>No. of Patients (Stage III)</th>
<th>Chemotherapy</th>
<th>Response Rate (%)</th>
<th>PCR (%)</th>
<th>Complete Resection</th>
<th>PFS</th>
<th>OS</th>
<th>5-Year Survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rosell et al</td>
<td>60 (60)</td>
<td>Mitomycin, Ifosfamide, Cisplatin</td>
<td>60</td>
<td>4</td>
<td>85%</td>
<td>12 vs. 5 mo (DFS; ( P = .006 ))</td>
<td>22 vs. 10 mo (( P = .005 ))</td>
<td>16% vs.</td>
</tr>
<tr>
<td>Roth et al</td>
<td>90 (60)</td>
<td>Cyclophosphamide, Etoposide, Cisplatin</td>
<td>35</td>
<td>NR</td>
<td>39% vs. 31%</td>
<td>Not reached vs. 9 mo (( P = .006 ))</td>
<td>64 vs. 11 mo (( P = .008 ))</td>
<td>56% vs.</td>
</tr>
<tr>
<td>Pass et al</td>
<td>91</td>
<td>Etoposide, Cisplatin</td>
<td>62</td>
<td>8</td>
<td>85% vs. 86%</td>
<td>12.7 vs. 5.8 mo (( P = .083 ))</td>
<td>28.7 vs. 15.6 mo (( P = .095 ))</td>
<td>NR</td>
</tr>
<tr>
<td>Nagai et al</td>
<td>92</td>
<td>Cisplatin, Vindesine</td>
<td>28</td>
<td>0</td>
<td>65% vs. 77%</td>
<td>NR</td>
<td>17 vs. 16 mo (( P = .5274 ))</td>
<td>10% vs.</td>
</tr>
<tr>
<td>Gilligan et al</td>
<td>519 (80)</td>
<td>Platinum based</td>
<td>49</td>
<td>4</td>
<td>82% vs. 80%</td>
<td>NR</td>
<td>54 vs. 55 mo (( P = .86 ))</td>
<td>44% vs.</td>
</tr>
<tr>
<td>Depierre et al</td>
<td>355 (167)</td>
<td>Mitomycin, Ifosfamide, Cisplatin</td>
<td>64</td>
<td>11</td>
<td>92% vs. 86%</td>
<td>26.7 vs. 12.9 mo (( P = .033 ))</td>
<td>37 vs. 26 mo (( P = .15 ))</td>
<td>43.9% vs.</td>
</tr>
<tr>
<td>Pisters et al</td>
<td>354 (113)</td>
<td>Carboplatin, Paclitaxel</td>
<td>41</td>
<td>NR</td>
<td>94% vs. 89%</td>
<td>33 vs. 21 mo (( P = .07 ))</td>
<td>75 vs. 46 mo (( P = .19 ))</td>
<td>50% vs.</td>
</tr>
<tr>
<td>Mattson et al</td>
<td>94</td>
<td>Docetaxel, Carboplatin</td>
<td>28</td>
<td>NR</td>
<td>77% vs. 76%</td>
<td>9 vs. 7.6 mo (NS)</td>
<td>14.8 vs. 12.6 mo (NS)</td>
<td>NR</td>
</tr>
</tbody>
</table>

Abbreviations: DFS = disease-free survival; NR = not recorded; NS = not significant; OS = overall survival; pCR = pathologic complete response; PFS = progression-free survival.

Once a treatment plan has been devised, two strategies for delivery are available. “Sequential” chemoradiation involves full-dose systemic chemotherapy (i.e., cisplatin combined with a second agent) followed by standard radiotherapy (approximately 60 Gy). The combination of chemotherapy followed by radiation has improved 5-year survival from 6% with radiotherapy alone to 17%. An alternative approach, referred to as “concurrent chemoradiation,” administers chemotherapy and radiation at the same time. Certain chemotherapeutic agents sensitize tumor cells to radiation and, thus, enhance the radiation effect. The advantages of this approach are improved primary tumor and locoregional lymph node control and elimination of the delay in administering radiotherapy that occurs with sequential treatment. A disadvantage, however, is the necessary reduction in chemotherapy dosage in order to diminish overlapping toxicities; this can potentially lead to undertreatment of systemic micrometastases. Randomized trials have shown a modest 5-year survival benefit as compared with chemotherapy. In a systematic review of 47 trials and six meta-analyses, an absolute survival benefit of 4% at 2 years was seen when concurrent platinum-based chemoradiation was given compared to sequential radiation.

Definitive radiotherapy is predominantly used for palliation of symptoms in patients with poor performance status; cure rates with radiation as a single modality in patients with N2 or N3 disease is less than 7%. Recent improvement has been seen with three-dimensional conformal radiotherapy and altered fractionation. Such poor results for patients with stage III lung cancer reflect the limitations of locoregional treatment in a disease where death results from systemic metastatic spread.

**Options for Thoracic Surgical Approaches**

Thoracic surgical approaches have changed over recent years with advancements in minimally invasive surgery. A surgeon trained in advanced minimally invasive techniques can now perform pleural-based, pulmonary and mediastinal procedures through multiple thoracoscopic ports without the need for a substantial, rib-spreading incision. Subjective measures of quality of life after VATS, such as pain (Fig. 19-26) and perceived functional recovery, consistently and reproducibly favor VATS over thoracotomy. Objective measures such as functional status as measured by 6-minute walk, return to work, and ability to tolerate chemotherapy also favor VATS over thoracotomy. Finally, recovery of respiratory function occurs earlier in VATS patients. These findings are pronounced in patients with chronic obstructive pulmonary disease and in the elderly—populations whose quality of life can be dramatically impacted by changes in their respiratory symptoms and function, thoracic pain, and physical performance. Table 19-16 provides a summary of populations that may benefit from VATS approaches.

**Video-Assisted (VATS)/Robotic-Assisted Thoracoscopic Surgery (RATS).** VATS/RATS has become the recommended approach to diagnosis and treatment of pleural effusions, recurrent pneumothoraces, lung biopsies, lobectomy or segmental resection, resection of bronchogenic and mediastinal cysts, and intrathoracic esophageal mobilization for esophagectomy. These approaches are also utilized for pneumonectomy in some centers of excellence with very high

![Pie chart comparison of pain control at 3 weeks after lobectomy by standard thoracotomy or video-assisted thoracic surgery (VATS). The pie charts show that patients undergoing VATS have significantly less pain (P < .01) as measured by the most potent analgesic still required: severe—schedule II narcotic; moderate—schedule III or lower narcotic; mild—nonsteroidal anti-inflammatory drug (NSAID) or acetaminophen. (Reproduced with permission from Demmy TL, Nwoque C. Is video-assisted thoracic surgery lobectomy better? Quality of life considerations, Ann Thorac Surg. 2008 Feb;85(2):S719-S728.)

**Table 19-16**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulmonary compromise</td>
<td>Poor FEV1/DLco, heavy smoking, sleep apnea, recent pneumonia</td>
</tr>
<tr>
<td>Cardiac dysfunction</td>
<td>Congestive heart failure, severe coronary artery disease, recent myocardial infarction, valvular disease</td>
</tr>
<tr>
<td>Extrathoracic malignancy</td>
<td>Solitary brain metastasis from lung cancer, deep pulmonary metastases requiring lobectomy</td>
</tr>
<tr>
<td>Poor physical performance</td>
<td>Performance status equivalent to a Zubrod score of 2 or 3, morbid obesity</td>
</tr>
<tr>
<td>Rheumatologic/orthopedic condition</td>
<td>Spinal disease, severe rheumatoid arthritis, severe kyphosis, lupus erythematosus, osteomyelitis</td>
</tr>
<tr>
<td>Advanced age</td>
<td>Age ≥ 70 years</td>
</tr>
<tr>
<td>Vascular problems</td>
<td>Aneurysm, severe peripheral vascular disease</td>
</tr>
<tr>
<td>Recent or impending major operation</td>
<td>Urgent abdominal operation, joint replacement requiring use of crutches, need for contralateral thoracotomy</td>
</tr>
<tr>
<td>Psychological/neurologic conditions</td>
<td>Substance abuse, poor command following, pain syndromes</td>
</tr>
<tr>
<td>Imunosuppression/impaired wound healing</td>
<td>Recent transplantation, diabetes</td>
</tr>
</tbody>
</table>

Abbreviations: DLco = carbon monoxide diffusion capacity; FEV1 = forced expiratory volume in 1 second.
volumes of VATS lung resection. VATS is performed via two to four incisions measuring 0.5 to 1.2 cm in length to allow insertion of the thoracoscope and instruments. An access incision, typically in the fourth or fifth intercostal space in the anterior axillary line, is used for dissection of the hilum during lung resection. The incision location varies according to the procedure. With respect to VATS lobectomy, port placement varies according to the lobe being resected and is highly variable among surgeons. The basic principle is to position the ports high enough on the thoracic cage to have access to the hilar structures. Endoscopic staplers are used to divide the major vascular structures and bronchus (Fig. 19-27). Robotic approaches are similarly tailored to the side and lobe undergoing resection, with the entire operation performed using

Figure 19-27. Selected video-assisted thoracic surgery lobectomy maneuvers. All the maneuvers are shown with the patient positioned in the left lateral decubitus position. The same maneuvers can be performed in mirror image for left-sided work. A. Medial viewing and inferior holding of lung to allow dissection through the access incision. Example shows dissection of the apical hilum. B. Medial viewing and access holding of lung to allow stapling of hilar structures from below. Example shows division of the apical pulmonary artery trunk to the right upper lobe (upper lobe branch of vein divided and reflected away). C. Standard viewing and use of working port to dissect and divide structures while lung is retracted through access incision. Example shows use of stapler to divide pulmonary artery to right lower lobe. D. Standard viewing and use of working port to retract lung and access incision to dissect structures. This method is commonly used to dissect the pulmonary artery in the major fissure. Example shows inferior pulmonary vein after the pulmonary ligament was divided using this maneuver. E. Standard viewing and use of access incision to deliver stapler to divide fissures. Example shows division of the posterior fissure between the right lower lobe and the upper lobe.
robotic arms, except for one assistant port through which specimens can be removed and suctioning and stapling performed (in early models; the newest robotic models have robotic staplers, allowing the surgeon to have complete control of the entire operation).

**Open Approaches to Thoracic Surgery.** When a thorascoscopic approach is not possible, an open approach, most frequently the posterolateral thoracotomy, is used to gain access to the intrathoracic space. The posterolateral thoracotomy incision can be used for most pulmonary resections, esophageal operations, and operations in the posterior mediastinum and vertebral column (Fig. 19-28). The anterolateral thoracotomy has traditionally been used in trauma victims. This approach allows quick entry into the chest with the patient supine. In the face of hemodynamic instability, the lateral decubitus position significantly compromises control over the patient’s cardiopulmonary system and resuscitation efforts, whereas the supine position allows the anesthesiologist full access to the patient. A bilateral anterior thoracotomy incision with a transverse sternotomy (“clamshell” thoracotomy) is a standard operative approach to the heart and mediastinum in certain elective circumstances. It is the preferred incision for double-lung transplantation in many centers. A partial median sternotomy can also be added to an anterior thoracotomy (“trap-door” or “hemiclamshell” thoracotomy) for access to mediastinal structures. A hypesthetic nipple is a frequent complication of this approach. The median sternotomy incision allows exposure of anterior mediastinal structures and is principally used for cardiac operations. Although the surgeon has access to both pleural cavities, incision into the pleural cavity can be avoided if entry is unnecessary (Fig. 19-29).

**Postoperative Care**

**Chest Tube Management.** At the conclusion of most thoracic operations, the pleural cavity is drained with a chest tube(s). If the visceral pleura has not been violated and there is no concern
The median sternotomy incision.

**A.** Skin incision from the suprasternal notch to the xiphoid process. **B.** Exposure of the pleural space. a. = artery; v. = vein.

for pneumo- or hemothorax (e.g., after VATS sympathectomy), a chest tube is unnecessary. After chest tube placement, the lung is reexpanded with positive-pressure ventilation. There are two reasons for the use of plural tubes in this setting: first, the tube allows evacuation of air if an air leak is present; second, blood and pleural fluid can be drained, thereby preventing accumulation within the pleural space that would compromise the patient’s respiratory status. The tube is removed when the air leak is resolved and when the volume of drainage decreases below an acceptable level over 24 hours.

Historically, many surgeons have somewhat arbitrarily required less than 150 mL of drainage volume over 24 hours prior to removing a chest tube to minimize risk of reaccumulation. The pleural lymphatics, however, can absorb up to 0.40 mL/kg per hour in a healthy individual, which may be as much as 500 mL over a 24-hour period. In fact, studies have shown that pleural tubes can be removed after VATS lobectomy or thoracotomy with 24-hour drainage volumes as high as 400 mL, without subsequent development of plural effusions.\(^82\)

It is our current practice to remove chest tubes with 24-hour outputs of 400 mL or less after lobectomy or lesser pulmonary resections. In settings where normal pleural fluid dynamics have been altered, such as malignant pleural effusion, pleural space infections or inflammation, and pleurodesis, strict adherence to a volume requirement before tube removal is appropriate (typically 100 to 150 mL over 24 hours).

For operations involving lung resection or parenchymal injury, suction levels of –20 cm H\(_2\)O are routinely used to eradicate residual air spaces and to control postoperative parenchymal air leaks for the first 12 to 24 hours. The following day, however, the decision to continue suction or place the patient to water seal (off suction) must be made. Applying suction to an air leak has been shown to prolong the duration of the air leak and extend the time frame during which tube thoracostomy is needed.\(^83\) The main guidelines for the continued use of suction if an air leak is present depend on the expansion of the remaining lung as determined by CXR. If the lung is well expanded, the chest tube can remain to water seal drainage. If an undrained pneumothorax is present on CXR, the chest tube and its attached tubing should be examined to ensure that the chest tube is patent and the attached tubing is not kinked or mechanically obstructed, such as occurs when the patient is lying on the tube. If the tube is a small caliber tube (pigtail catheter), it should be flushed with sterile saline through a three-way stopcock that has been cleaned with alcohol because these tubes tend to become clogged with fibrin. These catheters are also prone to kinking at the insertion site into the skin. Once the surgeon has confirmed that the chest tube is patent, the patient is asked to voluntarily cough or perform the Valsalva maneuver. This maneuver increases the intrathoracic pressure and will push air that is contained within the hemithorax out of the chest tube. During the voluntary cough, the fluid level in the water seal chamber should move up and down with the cough and with deep respiration, reflecting the pleural pressure changes occurring with these maneuvers. A stationary fluid level implies either a mechanical blockage (e.g., due to external tube compression or to a clot/debris within the tube) or pleurodesis. If bubbles pass through the water seal chamber, an air leak is presumed. If the leak is significant enough to induce atelectasis or collapse of the lung during use of water seal, suction should be used to achieve lung reexpansion.

**Pain Control.** Good pain control after intrathoracic procedures is critical; it permits the patient to actively clear and manage secretions and promotes ambulation and a feeling of well-being. The most common techniques of pain management are epidural, paravertebral, and intravenous. Epidural catheters are commonly used, although we prefer to use paravertebral catheters in our center. Epidural catheters should be inserted at about the T6 level, roughly at the level of the scapular tip. Lower placement risks inadequate pain control, and higher placement may provoke hand and arm numbness. Typically, combinations of fentanyl at 0.3 µg/mL with either bupivacaine (0.125%) or ropivacaine (0.1%) are used. Ropivacaine has less cardiotoxicity than bupivacaine; thus, in the case of inadvertent intravenous injection, the potential for refractory complete heart block is significantly less with ropivacaine. Paravertebral blocks can be placed using the same epidural catheter kit 2.5 cm lateral to the spinous process at T4 to T6. Combinations of narcotic and topical analgesia are then infused as with the epidural catheter.

When properly placed, a well-managed epidural can provide outstanding pain control without significant systemic sedation.\(^84\) Thoracic epidurals do not commonly cause urinary retention, although a low thoracic epidural may block the sensory fibers to the bladder. Motor function, however, remains intact. In some patients who are having difficulty voiding, it...
may be possible to avoid Foley catheterization by simply reminding the patient to void on a regular basis. In male patients with voiding difficulty prior to surgery, urinary catheterization may be required. In addition, the use of local anesthetics may cause sympathetic outflow blockade, leading to vasodilation and hypotension often requiring intravenous vasconstrictors (an α-agonist such as phenylephrine) and/or fluid administration. In such circumstances, fluid administration for hypotension may be undesirable in pulmonary surgery patients, particularly after pneumonectomy. Paravertebral catheters provide equivalent pain control with less effect on hemodynamics.86 Recently, liposomal bupivacaine was introduced and has become the standard approach to pain management after thoracic surgery in several centers. The formulation allows for slow-release of bupivacaine for up to 72 hours after injection. Injected directly into the intercostal spaces immediately prior to chest closure, the formulation has shown great promise in several retrospective reports, with randomized trials not yet completed. One limiting factor in use of the formulation is the cost; future trials are needed, with cost-analysis, to determine whether the benefits in pain, reduction in narcotic use, shorter hospital stay and possibly decreased pulmonary complications justify the cost.

Alternatively, intravenous narcotics via patient-controlled analgesia can be used, often in conjunction with ketorolac, gabapentin, and intravenous Tylenol. Dosing must be titrated to balance the degree of pain relief with the degree of sedation. Oversedated patients are as ominous as patients without adequate pain control because of the significant risk of secretion retention, atelectasis/pneumonia, and pulmonary aspiration, especially in elderly patients who should be carefully assessed for aspiration risk when ordered for dietary advancement. Proper pain control with intravenous narcotics requires a carefully regulated balance of pain relief and sedation to maximize the benefits of pain control while minimizing these very real and potentially life-threatening complications.

Whether on epidural, paravertebral, or intravenous pain control, the patient is typically transitioned to oral pain medication on the third or fourth postoperative day. During both the parenteral and oral phase of pain management, a standardized regimen of stool softeners and laxatives is advisable in order to prevent severe constipation.

Respiratory Care. The best respiratory care is achieved when the patient is able to deliver an effective cough to clear secretions and results from the commitment and proper training of all involved healthcare providers. The process begins preoperatively, with clear instructions on using pillows (or other support techniques) over the wound and then applying pressure.

Postoperatively, proper pain control without oversedation (as outlined earlier) is essential. Daily morning rounds should include a careful assessment of the patient’s pulmonary status, reminders to the patient and family about the importance of coughing and deep breathing, including use of adjunctive respiratory equipment if ordered, and mobilization of the patient. Early transition to a chair and to ambulation is the best respiratory therapy and should be strongly encouraged. When available, physical and/or cardiopulmonary rehabilitation services are vital additional members of the care team.

Postoperative Complications

Postoperative complications after pulmonary resection range from minor to life-threatening. Strict attention to volume status, early and aggressive pulmonary toilet, and good pain control can reduce the risk of most complications, but does not completely eliminate them, even in centers of excellence. The most devastating complication after pulmonary resection is postpneumonectomy pulmonary edema, which occurs in 1% to 5% of patients undergoing pneumonectomy and more often after right compared to left pneumonectomy. Clinically, symptoms of respiratory distress manifest hours to days after surgery. Radiographically, diffuse interstitial infiltration or frank alveolar edema is seen. The pathophysiologic causes are related increased permeability and filtration pressure and decreased lymphatic drainage from the affected lung. Judicious use of intravenous fluids perioperatively, including use of vasopressors rather than fluid boluses for hypotension intra- and postoperatively, is critical to minimizing the risk of this syndrome. Treatment consists of ventilatory support, fluid restriction, and diuretics. Extracorporeal membrane oxygenation may be lifesaving in centers where this option is available. The syndrome reportedly has a nearly 100% mortality rate despite aggressive therapy.

Other postoperative complications include air leak and bronchopleural fistula. Although these are two very different problems, distinguishing between them may be difficult. Postoperative air leaks are common after pulmonary resection, particularly in patients with emphysematous lung, because the fibrosis and destroyed blood supply impairs healing of surface injuries. Prolonged air leaks (i.e., those lasting >5 days) may be treated by diminishing or discontinuing suction (if used), by continuing chest drainage, or by instilling a pleurodesis agent, usually doxycycline or talcum powder, which will cause pleurodesis of the lung within the chest cavity and minimize the possible collapse of the lung due to persistent air leak. This is useful only in patients in whom full lung expansion is achieved, either with suction or on water seal, as patients with a persistent pneumothorax on CXR will not have adequate lung-to-parietal pleural apposition to achieve adequate pleurodesis.

If the leak is moderate to large, a high index of suspicion for bronchopleural fistula from the resected bronchial stump should be maintained, particularly if the patient is immunocompromised or had induction chemo- and/or radiation therapy. If suspected, flexible bronchoscopy is performed to evaluate the bronchial stump. Management options include continued prolonged chest tube drainage, reoperation, and reclosure (with stump reinforcement with intercostal or pedicled serratus muscle flap). If the fistula is very small (<4 mm), bronchoscopic fibrin glue application has been used successfully to seal the hole in some patients. Patients often have concomitant empyema, and open drainage may be necessary.

Spontaneous Pneumothorax

Spontaneous pneumothorax is secondary to intrinsic abnormalities of the lung and can be classified as primary and secondary. Primary spontaneous pneumothorax is defined as a spontaneous pneumothorax without underlying lung disease. The most common cause is rupture of an apical subpleural bleb. The cause of these blebs is unknown, but they occur more frequently in smokers and males, and they tend to predominate in young postadolescent males with a tall thin body habitus. Treatment is generally chest tube insertion with water seal. If a leak is present and persists for greater than 3 days, thoracoscopic management (i.e., bleb resection with pleurodesis by talc or pleural abrasion) is performed. Recurrences or complete lung collapse with the first episode are generally indications for thoracoscopic
Pulmonary Infections

Lung Abscess. A lung abscess is a localized area of pulmonary parenchymal necrosis caused by an infectious organism; tissue destruction results in a solitary or dominant cavity measuring at least 2 cm in diameter. Less often, there may be multiple, smaller cavities (<2 cm). In that case, the infection is typically referred to as a necrotizing pneumonia. An abscess that is present for more than 6 weeks is considered chronic.

Based on the etiology (Table 19-17), lung abscesses are further classified as primary or secondary. A primary lung abscess occurs, for example, in immunocompromised patients, as a result of highly virulent organisms inciting a necrotizing pulmonary infection, or in patients who have a predisposition to aspirate oropharyngeal or gastrointestinal secretions. A secondary lung abscess occurs in patients with an underlying condition such as a partial bronchial obstruction, a lung infarct, or adjacent suppurative infections (subphrenic or hepatic abscesses).

Pathogenesis Lung abscesses result when necrotizing microorganisms infect the lower respiratory tract via inhalation of aerosolized particles, aspiration of oropharyngeal secretions, or hematogenous spread from distant sites. Direct extension from a contiguous site is less frequent. Most primary lung abscesses are suppurative bacterial infections secondary to aspiration. Risk factors for pulmonary aspiration include advanced age, conditions of impaired consciousness, suppressed cough reflex, dysfunctional esophageal motility, laryngopharyngeal reflex dysfunction, and centrally acting neurologic diseases (e.g., stroke). At the time of aspiration, the composition of the oropharyngeal flora determines the etiologic organisms. With increasing use of proton pump inhibitors to suppress acid secretion in the stomach, the oropharyngeal flora has shifted, and the risk of developing bacterial lung infections after an aspiration event has increased.

Secondary lung abscesses occur most often distal to an obstructing bronchial carcinoma. Infected cysts or bullae are not considered true abscesses.

Microbiology Normal oropharyngeal secretions contain many more Streptococcus species and more anaerobes (approximately $1 \times 10^8$ organisms/mL) than aerobes (approximately $1 \times 10^7$ organisms/mL). Pneumonia that follows from aspiration, with or without abscess development, is typically polymicrobial. An average of two to four isolates present in large numbers have been cultured from lung abscesses sampled percutaneously. Overall, at least 50% of these infections are caused by purely anaerobic bacteria, 25% are caused by mixed aerobes and anaerobes, and 25% or fewer are caused by aerobes only. In nosocomial pneumonia, 60% to 70% of the organisms are gram-negative bacteria. Immunocompromised patients may develop abscesses because of the usual pathogens as well as less virulent and opportunistic organisms such as Salmonella species, Legionella species, Pneumocystis carinii, atypical mycobacteria, and fungi.

Clinical Features and Diagnosis The typical presentation may include productive cough, fever (>38.9°C), chills, leukocytosis (>15,000 cells/mm$^3$), weight loss, fatigue, malaise, pleuritic chest pain, and dyspnea. Lung abscesses may also present in a more indolent fashion, with weeks to months of cough, malaise, weight loss, low-grade fever, night sweats, leukocytosis, and anemia. After aspiration pneumonia, 1 to 2 weeks typically elapse before cavitation occurs; 40% to 75% of such patients produce putrid, foul-smelling sputum. Severe complications

### Table 19-17

<table>
<thead>
<tr>
<th>Causes of lung abscess</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Primary</td>
</tr>
<tr>
<td>A. Necrotizing pneumonia</td>
</tr>
<tr>
<td>1. <em>Staphylococcus aureus</em>, <em>Klebsiella</em>, <em>Pseudomonas</em>, <em>Mycobacterium</em></td>
</tr>
<tr>
<td>2. <em>Bacteroides</em>, <em>Fusobacterium</em>, <em>Actinomyces</em></td>
</tr>
<tr>
<td>3. <em>Entamoeba</em>, <em>Echinococcus</em></td>
</tr>
<tr>
<td>B. Aspiration pneumonia</td>
</tr>
<tr>
<td>1. Anesthesia</td>
</tr>
<tr>
<td>2. Stroke</td>
</tr>
<tr>
<td>3. Drugs or alcohol</td>
</tr>
<tr>
<td>C. Esophageal disease</td>
</tr>
<tr>
<td>1. Achalasia, Zenker’s diverticulum, gastroesophageal reflux</td>
</tr>
<tr>
<td>D. Immunodeficiency</td>
</tr>
<tr>
<td>1. Cancer (and chemotherapy)</td>
</tr>
<tr>
<td>2. Diabetes</td>
</tr>
<tr>
<td>3. Organ transplantation</td>
</tr>
<tr>
<td>4. Steroid therapy</td>
</tr>
<tr>
<td>5. Malnutrition</td>
</tr>
<tr>
<td>II. Secondary</td>
</tr>
<tr>
<td>A. Bronchial obstruction</td>
</tr>
<tr>
<td>1. Neoplasm</td>
</tr>
<tr>
<td>2. Foreign body</td>
</tr>
<tr>
<td>B. Systemic sepsis</td>
</tr>
<tr>
<td>1. Septic pulmonary emboli</td>
</tr>
<tr>
<td>2. Seeding of pulmonary infarct</td>
</tr>
<tr>
<td>C. Complication of pulmonary trauma</td>
</tr>
<tr>
<td>1. Infection of hematoma or contusion</td>
</tr>
<tr>
<td>2. Contaminated foreign body or penetrating injury</td>
</tr>
<tr>
<td>D. Direct extension from extraparenchymal infection</td>
</tr>
<tr>
<td>1. Pleural empyema</td>
</tr>
<tr>
<td>2. Mediastinal, hepatic, subphrenic abscess</td>
</tr>
</tbody>
</table>

such as massive hemoptysis, endobronchial spread to other portions of the lungs, rupture into the pleural space and development of pyopneumothorax, or septic shock and respiratory failure are rare in the modern antibiotic era. The mortality rate is about 5% to 10%, except in the presence of immunosuppression, where rates range from 9% to 28%.

The CXR is the primary tool for diagnosing a lung abscess (Fig. 19-30). Its distinguishing characteristic is a density or mass with a relatively thin-walled cavity. An air-fluid level observed within the abscess indicates communication with the tracheobronchial tree. CT scan of the chest clarifies the diagnosis when CXR is equivocal and identifies endobronchial obstruction and/or an associated mass and other pathologic anomalies. A cavitating lung carcinoma is frequently mistaken for a lung abscess. Differential diagnosis also includes loculated or interlobar empyema, infected lung cysts or bullae, tuberculosis, bronchiectasis, fungal infections, and noninfectious inflammatory conditions (e.g., Wegener’s granulomatosis).

Ideally, the specific etiologic organism is identified before antibiotic administration. Bronchoscopy, which is essential to

Figure 19-30. Lung abscess resulting from emesis and aspiration after an alcoholic binge. A. Chest X-ray showing an abscess cavity in the left upper lobe. B. A coronal tomogram highlights the thin wall of the abscess. C. Healing of the abscess cavity after 4 weeks of antibiotic therapy and postural drainage.
rule out endobronchial obstruction due to tumor or foreign body, is ideal for obtaining uncontaminated cultures using bronchoalveolar lavage. Culture samples can also be obtained by percutaneous, transthoracic FNA under ultrasound or CT guidance. Routine sputum cultures are often of limited usefulness because of contamination with upper respiratory tract flora.

Actinomycosis and nocardiosis, although rare, are particularly virulent infections associated with lung abscess. Diagnosis can be difficult. Both frequently masquerade as other clinical syndromes; thus, it is important for the surgeon to keep these bacteria in mind when considering the differential diagnosis for cavitary lung lesions. Actinomyces, a normal oropharyngeal bacterium, causes extensive pulmonary damage. Actinomycosis lung infection typically begins as acute pneumonitis after pulmonary aspiration. The symptoms mimic pulmonary tuberculosis, including chronic cough, night sweats, weight loss, and hemoptysis.

Ongoing infection leads to chronic inflammation and fibrosis; cavitation occurs due to destruction of the pulmonary tissues. Without treatment, the infection continues to destroy surrounding structures, resulting in fistula formation to adjacent structures, including the adjacent lung, interlobar fissures, pleural space, chest wall, and mediastinum. Actinomyces israelii is the most common of the species to cause disease. Nocardiosis is also a rare opportunistic infection that usually occurs in an immunocompromised host (HIV or cancer patients) and causes both local and systemic suppurative infections. The most common site is pulmonary, caused by Nocardia asteroides in 90% of cases. Infection is slowly progressive, with weight loss, fatigue, cough, and hemoptysis. An acute pulmonary infection is common, with necrotizing pneumonia and cavitation or slowly enlarging pulmonary nodule. In some cases, empyema also develops.

Management of Lung Abscess Systemic antibiotics directed against the causative organism represent the mainstay of therapy. The duration of antimicrobial therapy varies from 3 to 12 weeks for necrotizing pneumonia and lung abscess. It is likely best to treat until the cavity is resolved or until serial radiographs show significant improvement. Parenteral therapy is generally used until the patient is afebrile and able to demonstrate consistent enteral intake. Oral therapy can then be used to complete the course of therapy. For community-acquired infections secondary to aspiration, likely pathogens are oropharyngeal streptococci and anaerobes. Penicillin G, ampicillin, and amoxicillin are the main therapeutic agents, but a β-lactamase inhibitor or metronidazole should be added to cover the increasing prevalence of gram-negative anaerobes that produce β-lactamase. Clindamycin is also a primary therapeutic agent. For hospital-acquired infections, Staphylococcus aureus and aerobic gram-negative bacilli are common organisms of the oropharyngeal flora. Piperacillin with a β-lactamase inhibitor (or equivalent alternatives) provide better coverage of likely pathogens.

Surgical drainage of lung abscesses is uncommon since drainage usually occurs spontaneously via the tracheobronchial tree. Indications for intervention are listed in Table 19-18. Drainage and resection may be required for actinomycosis and nocardiosis; diagnosis is often delayed because the bacteria are difficult to culture; invasion of the infection into surrounding structures is, therefore, common. Once identified, long-term antibiotics (months to years) are typically required along with drainage, debridement, and resection as needed. While penicillin derivatives are effective against most Actinomyces species, the infections are typically polymicrobial, and broad-spectrum parenteral antibiotics may be required. Nocardia species, in contrast, are highly variable; specific identification of the infecting species with antibiotic sensitivities is needed to direct appropriate therapy. Evaluation for malignant spread, particularly to the brain, is also required in the management of nocardiosis, as systemic dissemination occurs early and frequently.

External drainage may be accomplished with tube thoracostomy, percutaneous drainage, or surgical cavernostomy. The choice between tube thoracostomy versus radiographically guided catheter placement depends on the treating physician’s preference and institutional technical expertise in placing image-guided thoracostomy tubes. Surgical resection is required in fewer than 10% of lung abscess patients. Lobectomy is the preferred intervention for bleeding from a lung abscess or pyopneumothorax. An important intraoperative consideration is to protect the contralateral lung with a double-lumen tube, bronchial blocker, or contralateral main stem intubation. Surgical treatment has a 90% success rate, with associated mortality of 1% to 13%.

Bronchiectasis. Bronchiectasis is defined as a pathologic and permanent dilation of bronchi with bronchial wall thickening. This condition may be localized to certain bronchial segments, or it may be diffuse throughout the bronchial tree, typically affecting the medium-sized airways. Overall, this is a rare clinical entity in the United States with a prevalence of less than 1 in 10,000, although the incidence has increased in recent years and noncystic fibrosis–related bronchiectasis is now thought to affect 27.5 out of every 10,000 persons over age 75.

Pathogenesis Development of bronchiectasis can be attributed to either congenital or acquired causes. The principal congenital diseases that lead to bronchiectasis include cystic fibrosis, primary ciliary dyskinesia, and immunoglobulin deficiencies (e.g., selective IgA deficiency). Congenital causes commonly produce a diffuse pattern of bronchial involvement. Acquired causes are categorized broadly as infectious and inflammatory. Bronchial obstruction from cancer, inhaled objects, extrinsic airway compression, or inspissated sputum promotes localized infection and subsequent medium airway destruction. Diffuse pneumatic processes from pathogens including necrotizing bacterial pneumonia, pertussis and measles pneumonia, severe influenza, or varicella pneumonia can lead to widespread bronchiectasis. Chronic granulomatous disease, immunodeficiency disorders, and hypersensitivity disorders can also lead to diffuse bronchiectasis.

Noninfectious causes of bronchiectasis include inhalation of toxic gases such as ammonia, which results in severe

---

**Table 19-18**

<table>
<thead>
<tr>
<th>Indications for surgical drainage procedures for lung abscesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Failure of medical therapy</td>
</tr>
<tr>
<td>2. Abscess under tension</td>
</tr>
<tr>
<td>3. Abscess increasing in size during appropriate treatment</td>
</tr>
<tr>
<td>4. Contralateral lung contamination</td>
</tr>
<tr>
<td>5. Abscess &gt;4–6 cm in diameter</td>
</tr>
<tr>
<td>6. Necrotizing infection with multiple abscesses, hemoptysis, abscess rupture, or pyopneumothorax</td>
</tr>
<tr>
<td>7. Inability to exclude a cavitating carcinoma</td>
</tr>
</tbody>
</table>

---
and destructive airway inflammatory responses. Allergic bronchopulmonary aspergillosis, Sjögren’s syndrome, and α1-antitrypsin deficiency are some additional examples of presumed immunologic disorders that may be accompanied by bronchiectasis.

In addition, recent studies have suggested an association between chronic gastroesophageal reflux disease, acid suppression, and nontuberculous mycobacterial infection with bronchiectasis. This interaction is thought to be related to chronic aspiration of colonized gastric secretions in the setting of acid suppression; while not proven to be causative, these findings suggest a role for gastroesophageal reflux disease in the pathogenesis of bronchiectasis.

The process shared by all causes of bronchiectasis is impairment of airway defenses or deficits in immunologic mechanisms, which permit bacterial colonization and chronic infection. Common organisms include Haemophilus species (55%), Pseudomonas species (26%), and Streptococcus pneumoniae (12%). Both the bacterial organisms and the inflammatory cells recruited to thwart the bacteria elaborate proteolytic and oxidative molecules, which progressively destroy the muscular and elastic components of the airway walls; those components are then replaced by fibrous tissue. Thus, chronic airway inflammation is the essential pathologic feature of bronchiectasis. The dilated airways are usually filled with thick purulent material; more distal airways are often occluded by secretions or obliterated by fibrous tissue. Bronchial wall vascularity increases, bronchial arteries become hypertrophied, and abnormal anastomoses form between the bronchial and pulmonary arterial circulation.

There are three principal types of bronchiectasis, based on pathologic morphology: cylindrical—uniformly dilated bronchi; varicose—an irregular or beaded pattern of dilated bronchi; and saccular (cystic)—peripheral balloon-type bronchial dilation. The saccular type is the most common after bronchial obstruction or infection (Fig. 19-31).

**Clinical Manifestations and Diagnosis** Typical symptoms are a daily persistent cough and purulent sputum production; the quantity of daily sputum production (10 mL to >150 mL) correlates with disease extent and severity. Other patients may appear asymptomatic or have a dry nonproductive cough (“dry bronchiectasis”). These patients are prone to have involvement of the upper lobes. The clinical course is characterized by progressive

![Figure 19-31. Multiple cystic-type bronchiectatic cavities can be seen on a cut section of right lower lobe lung.](image-url)
shown either no change or a worsening of pulmonary status and require further study in the non–cystic fibrosis population.

Surgical resection of a localized bronchiectatic segment or lobe, preserving as much functional lung as possible, may benefit patients with refractory symptoms while on maximal medical therapy. Multifocal disease must be excluded before any attempt at surgery; any uncorrectable predisposing factor (e.g., ciliary dyskinesia) must also be excluded. Patients with end-stage lung disease from bronchiectasis may be potential candidates for a bilateral lung transplant. Surgical resection is also indicated in patients with significant hemothysis, although bronchial artery embolization is the preferred first option. Antireflux surgery may also prove beneficial in patients with chronic aspiration, but further studies are required. It is particularly important to recognize that antireflux surgery in patients with severe underlying pulmonary dysfunction has higher risk for perioperative adverse outcomes than in the general population. It should be undertaken only by very experienced surgeons with direct involvement of the pulmonary medicine physicians to minimize postoperative pulmonary compromise.

**Mycobacterial Infections**

**Epidemiology**

Tuberculosis is a widespread problem that affects nearly one-third of the world’s population. It is the ninth leading cause of death worldwide and the leading infectious cause of death. The rate of death from tuberculosis has declined, from 1.7 million in 2000 to 1.3 million among HIV-negative people in 2016. There were 6.3 million new cases of tuberculosis worldwide in 2016 according to the World Health Organization (WHO); 56% are in the countries of India, Indonesia, China, the Philippines, and Pakistan. Treatment success rate was 83%. Only 9257 new cases were reported to the WHO in the United States in 2016. HIV infection is the strongest risk factor for developing active tuberculosis. The elderly, minorities, and recent immigrants are the most common populations to have clinical manifestations of infection, yet no age group, sex, or race is exempt from infection. In most large urban centers, reported cases of tuberculosis are more numerous among the homeless, prisoners, and drug-addicted populations. Immunocompromised patients additionally contribute to an increased incidence of tuberculosis infection, often developing unusual systemic as well as pulmonary manifestations. As compared with past decades, presently surgical intervention is required more frequently in patients with multidrug-resistant or rifampin-resistant (but isoniazid-susceptible) tuberculosis organisms (MDR/RR-TB) who do not respond to medical treatment and in selected patients with nontuberculous mycobacterial infections (NTM).

**Microbiology**

Mycobacterial species are obligate aerobes. They are primarily intracellular parasites with slow growth rates. Their defining characteristic is the property of acid-fastness, which is the ability to withstand decolorization by an acid-alcohol mixture after being stained. *Mycobacterium tuberculosis* is the highly virulent bacillus of this species that produces invasive infection among humans, principally pulmonary tuberculosis.

Infections with *M tuberculosis* are primary when they are the first infection in a previously unsensitized host and secondary or postprimary when reactivation of a previous infection occurs.

Because of improper application of antitubercular drugs and multifactorial interactions, MDR-TB organisms, defined by their resistance to at least two of the first-line antitubercular drugs (isoniazid and rifampin), and rifampicin-resistant (but isoniazid-susceptible) (RR-TB), have emerged. According to the WHO Global Tuberculosis Report 2017, in 2014, there were 108 reported cases of TB from MDR/RR-TB organisms, with 78% of cases successfully treated. In addition, there is another rare variant termed extensively drug-resistant tuberculosis (XDR-TB). These organisms are resistant to isoniazid and rifampin and have also developed resistance to either fluoroquinolones and injectable second-line drugs (e.g., capreomycin, amikacin, kanamycin), the two other classes of medications in the MDR-TB treatment regimen. In 2014, there were 109,680 MDRTB cases globally, with 6777 (6.2%) extensively drug resistant. Successful treatment was achieved in 54% of MDRTB and only 30% of XDR-TB. In 2016, it was estimated that MDR/RR was the responsible organism for more than 4% of new cases and nearly 20% of previously treated cases.

The more important NTM organisms include *Mycobacterium kansasii*, *M avium* and *M intracellulare* complex (MAC), and *M fortuitum*. The highest incidence of *M kansasii* infection is in Midwestern U.S. cities among middle-aged men from good socioeconomic surroundings. MAC organisms are important infections in elderly and immunocompromised patient groups. *M fortuitum* infections are common complications of underlying severe debilitating disease. None of these organisms are as contagious as *M tuberculosis*.

**Pathogenesis and Pathology**

The main route of transmission is via airborne inhalation of viable mycobacteria. Three stages of primary infection have been described. In the first stage, alveolar macrophages become infected through ingesting the bacilli. In the second stage, from days 7 to 21, the patient typically remains asymptomatic while the bacteria multiply within the infected macrophages. The third stage is characterized by the onset of cell-mediated immunity (CD4+ helper T cells) and delayed-type hypersensitivity. Activated macrophages acquire an increased capacity for bacterial killing. Macrophage death increases, resulting in the formation of a granuloma, the characteristic lesion found on pathologic examination.

Tuberculous granulomas are composed of blood-derived macrophages, degenerating macrophages or epithelioid cells, and multinucleated giant cells (fused macrophages with nuclei around the periphery; also known as Langerhans cells). The low oxygen content of this environment inhibits macrophage function and bacillary growth, with subsequent central caseation as macrophage death occurs. A Ghon complex is a single, small lung lesion that is often the only remaining trace of a primary infection. The primary infection is usually located in the peripheral portion of the middle zone of the lungs.

Reactivation tuberculosis may occur after hydrolytic enzymes liquefy the caseum. Typically, the apical and posterior segments of the upper lobes and the superior segments of the lower lobes are involved. Edema, hemorrhage, and mononuclear cell infiltration are also present. The tuberculous cavity may become secondarily infected with other bacteria, fungi, or yeasts, all of which may contribute to enhanced tissue destruction.

The pathologic changes caused by NTM organisms are similar to those produced by *M tuberculosis*. *M intracellulare* complex infections commonly occur, not only in immunocompromised patients but also in patients with previously damaged lungs. Caseous necrosis is uncommon and is characterized by clusters of tissue macrophages filled with mycobacteria. It has a poor granulomatous response and confinement of immune cell infiltration to the interstitium and alveolar walls. Cavitative disease is infrequent, although nodules may be noted.
Clinical Presentation and Diagnosis. The clinical course of infection and the presentation of symptoms are influenced by many factors, including the site of primary infection, the stage of disease, and the degree of cell-mediated immunity. About 80% to 90% of tuberculosis patients present with clinical disease in the lungs. In 85% to 90% of these patients, involution and healing occur, leading to a dormant phase that may last a lifetime. The only evidence of tuberculosis infection may be a positive skin reaction to tuberculin challenge or a Ghon complex observed on CXR. Within the first 2 years of primary infection, reactivation may occur in up to 10% to 15% of infected patients. In 80%, reactivation occurs in the lungs; other reactivation sites include the lymph nodes, pleura, and the musculoskeletal system.

After primary infection, pulmonary tuberculosis is frequently asymptomatic. Systemic symptoms of low-grade fever, malaise, and weight loss are subtle and may go unnoticed. A productive cough may develop, usually after tubercle cavitation. Many radiographic patterns can be identified at this stage, including local exudative lesions, local fibrotic lesions, cavitation, bronchial wall involvement, acute tuberculous pneumonia, bronchiectasis, bronchostenosis, and tuberculous granulomas. Hemoptysis often develops from complications of disease such as bronchiectasis or erosion into vascular malformations associated with cavitation.

Extrapulmonary involvement is due to hematogenous or lymphatic spread from pulmonary lesions. Virtually any organ can become infected, giving rise to the protean manifestations of tuberculosis. The pleura, chest wall, and mediastinal organs may all be involved. More than one-third of immunocompromised patients have disseminated disease, with hepatomegaly, diarrhea, splenomegaly, and abdominal pain.

The definitive diagnosis of tuberculosis requires identification of the mycobacterium in a patient’s bodily fluids or involved tissues. Skin testing using purified protein derivative is important for epidemiologic purposes and can help exclude infection in uncomplicated cases. For pulmonary tuberculosis, sputum examination is inexpensive and has a high diagnostic yield.

Bronchoscopy with alveolar lavage may also be a useful diagnostic adjunct and has high diagnostic accuracy. Chest CT scan can delineate the extent of parenchymal disease.

Management. Medical therapy is the primary treatment of pulmonary tuberculosis and is often initiated before a mycobacterial pathogen is definitively identified. Combinations of two or more drugs are routinely used in order to minimize resistance, which inevitably develops with only single-agent therapy. A current treatment algorithm is outlined in Fig. 19-32. Generally, therapy lasts about 26 weeks (2 months intensive therapy followed by 4 months continuation therapy). A 7-month continuation phase should be considered for patients with cavitary pulmonary TB who remain positive on sputum culture after the 2-month intensive therapy, those patients who did not receive pyrazinamide during the intensive phase, HIV-positive patients who did not receive concomitant antiretroviral therapy, and patients treated with INH and rifapentine once weekly (not recommended) who have persistent positive sputum after 2 months of intensive therapy.

In the United States, surgical intervention is most often required in order to treat patients with MDR/RR-TB organisms whose lungs have been destroyed and who have persistent thick-walled cavitation. The indications for surgery related to mycobacterial pulmonary infections are presented in Table 19-19. The governing principle of mycobacterial surgery is to remove all gross disease while preserving any uninvolved lung tissue. Scattered nodular disease may be left intact, given its low mycobacterial burden. Antimycobacterial medications should be given preoperatively (for about 3 months) and continued postoperatively for 12 to 24 months. Overall, more than 90% of patients who were deemed good surgical candidates are cured when appropriate medical and surgical therapy is used.

Pulmonary Fungal Infections. The incidence of fungal infections has increased significantly, with many new opportunistic fungi emerging. This increase is attributed to the growing population of immunocompromised patients (e.g., organ transplant recipients, cancer patients undergoing chemotherapy, HIV patients, and young and elderly patients) who are more likely to become infected with fungi. Clinically significant examples include species of Aspergillus, Cryptococcus, Candida, and Blastomyces. Other at-risk patient populations include those who are malnourished, severely debilitated, or diabetic or who have hematologic disorders.

Patients receiving high-dose, intensive antibiotic therapies are also susceptible. There are, however, some fungi that are primary or true pathogens, able to cause infections in otherwise healthy patients. Some endemic examples in the United States include species of Histoplasma, Coccidioides, and Blastomyces.

Direct identification of the organism in body exudates or tissues, preferably as growth in culture, provides definitive diagnosis. Serologic testing to identify mycotic-specific antibodies may also be useful. Several new classes of antifungal agents have proven effective against many life-threatening fungi and are less toxic than older agents. In addition, thoracic surgery may be a useful therapeutic adjunct for patients with pulmonary mycoses.

Aspergillus. The genus Aspergillus comprises over 150 species and is the most common cause of mortality due to invasive mycoses in the United States. It is typically acute in onset and life-threatening and occurs in the setting of neutropenia, chronic steroid therapy, or cytotoxic chemotherapy. It can also occur in the general intensive care unit population of critically ill patients, including patients with underlying chronic obstructive pulmonary disease (COPD), postoperative patients, patients with cirrhosis or alcoholism, and postinfluenza patients, without any of these factors present. The species most commonly responsible for clinical disease include A fumigatus, A flavus, A niger, and A terreus. Aspergillus is a saprophytic, filamentous fungus with septate hyphae. Spores (2.5 to 3 µm in diameter) are released and easily inhaled by susceptible patients; because the spores are microns in size, they are able to reach the distal bronchi and alveoli.

Diagnosis of aspergillosis requires one or more cavities on lung imaging, with or without fungal ball or nodules, microscopy or culture positive for Aspergillus, or antibodies (precipitins) to Aspergillus on serum testing. The characteristics must have been present for 3 months. Aspergillus can manifest as one of three clinical syndromes: Aspergillus hypersensitivity lung disease, aspergillosis, or invasive pulmonary aspergillosis. Overlap occurs between these syndromes, depending on the patient’s immune status. Aspergillus hypersensitivity manifests as a productive cough, fever, wheezing, pulmonary infiltrates, eosinophilia, and elevation of IgE antibodies.
Figure 19-32. Treatment algorithm for tuberculosis. Patients in whom tuberculosis is proven or strongly suspected should have treatment initiated with isoniazid (INH), rifampin (RIF), pyrazinamide (PZA), and ethambutol (EMB) for the initial 2 months. A repeat smear and culture should be performed when 2 months of treatment has been completed. If cavities were seen on the initial chest radiograph (CXR) or the acid-fast bacillus (AFB) smear results are positive at completion of 2 months of treatment, the continuation phase of treatment should consist of INH and RIF daily or twice daily for 4 months to complete a total of 6 months of treatment. If cavitation was present on the initial CXR and the culture results at the time of completion of 2 months of therapy are positive, the continuation phase should be lengthened to 7 months (total of 9 months of treatment). If the patient has HIV infection and the CD4+ cell count is <100/µL, the continuation phase should consist of daily or three times weekly INH and RIF. In HIV-uninfected patients with no cavitation on CXR and negative results on AFB smears at completion of 2 months of treatment, the continuation phase may consist of either once weekly INH and rifapentine (RPT) or daily or twice weekly INH and RIF to complete a total of 6 months of treatment (bottom). For patients receiving INH and RPT whose 2-month culture results are positive, treatment should be extended by an additional 3 months (total of 9 months). EMB may be discontinued when results of drug susceptibility testing indicate no drug resistance. PZA may be discontinued after it has been taken for 2 months (56 doses). RPT should not be used in HIV-infected patients with tuberculosis or in patients with extrapulmonary tuberculosis. Therapy should be extended to 9 months if results of 2-month culture are positive. (Reproduced with permission from Blumberg HM, Burman WJ, Chaisson RE, et al. American Thoracic Society/Centers for Disease Control and Prevention/Infectious Diseases Society of America: treatment of tuberculosis, Am J Respir Crit Care Med. 2003 Feb 15;167(4):603-662.)

Table 19-19

Indications for surgery to treat mycobacterial pulmonary infections

1. Complications resulting from previous thoracic surgery to treat tuberculosis
2. Failure of optimized medical therapy (e.g., progressive disease, lung gangrene, or intracavitary aspergillosis superinfection)
3. Need for tissue acquisition for definitive diagnosis
4. Complications of pulmonary scarring (e.g., massive hemoptysis, cavernomas, bronchiectasis, or bronchostenosis)
5. Extrapulmonary thoracic involvement
6. Pleural tuberculosis
7. Nontuberculous mycobacterial infection

to Aspergillus, whereas aspergilloma (fungal ball) is a matted sphere of hyphae, fibrin, and inflammatory cells that tends to colonize preexisting intrapulmonary cavities. Grossly, aspergilloma appears as a round or oval, friable, gray (or red, brown, or even yellow), necrotic-looking mass (Fig. 19-33). This form is the most common presentation of noninvasive pulmonary aspergillosis. The most common symptoms are hemoptysis, chronic and productive cough, clubbing, malaise, or weight loss. CXR can suggest the diagnosis by the finding of a crescentic radiolucency above a rounded radiopaque lesion (Monad sign).

The natural history varies greatly between patients and, therefore, treatment is individualized. Factors associated with poor prognosis include severe underlying pulmonary disease, growth in the number or size of the aspergilloma(s) during observation, immunosuppression or HIV infection, history of lung transplantation, chronic pulmonary sarcoidosis, and increasing Aspergillus-specific IgG titers.
Asymptomatic patients can be observed without any additional therapy. Antifungals have limited utility due to the poor blood supply to the aspergilloma. Oral triazole therapy is now considered the standard of care for chronic, cavitary pulmonary aspergillosis. Hemoptysis is a harbinger of erosion of the disease into adjacent bronchial arteries and typically requires intervention. In the setting of very mild hemoptysis (e.g., blood-streaked sputum), cough suppression is warranted while further therapeutic evaluation is performed.

Bronchial artery embolization is the first-line therapy for massive hemoptysis and may be definitive therapy.\textsuperscript{105} This is particularly important to consider for patients with severely impaired pulmonary function who may not have sufficient reserve to tolerate even a very small pulmonary resection. Operative intervention may be required for recurrent hemoptysis, particularly after bronchial artery embolization, chronic cough with systemic symptoms, progressive infiltrate around the mycetoma, and a pulmonary mass of unknown cause.\textsuperscript{106}

When operative intervention is indicated, the surgeon must remain cognizant of the goals of the procedure. In the setting of simple aspergilloma, VATs wedge resection is preferred. As this disease typically occurs in patients with significantly impaired pulmonary function, attempts should be made to excise all diseased tissue with as limited a resection as possible. Once resection is completed, the postresection space in the hemithorax should be obliterated with a pleural tent, pneumoperitoneum, decortication of the remaining lung, intrathoracic rotation of a muscle or omental flap, or thoracoplasty. If completely resected for single aspergilloma, antifungal therapy is not needed. If multiple nodules are present or the disease is incompletely resected, however, antifungal therapy should be considered. Long-term follow-up is necessary, given that the recurrence rate after surgery is about 7%. 

\textbf{Figure 19-33.} Pulmonary aspergilloma. \textbf{A.} The chest X-ray shows a solid mass within a cavity surrounded by a rim of air between the mass and cavity wall (Monad sign, arrows). \textbf{B.} A cut section shows the “fungus ball” occupying an old, fibrotic cavity. \textbf{C.} Histologic stain reveals characteristic \textit{Aspergillus} hyphae invading the wall of the cavity.
Invasive pulmonary aspergillosis typically affects immunocompromised patients who have dysfunctional cellular immunity, namely defective polymorphonuclear leukocytes. Invasion of pulmonary parenchyma and blood vessels by a necrotizing bronchopneumonia may be complicated by thrombosis, hemorrhage, and then dissemination. Patients present with fever that is nonresponsive to antibiotic therapy in the setting of neutropenia. They may also have pleuritic chest pain, cough, dyspnea, or hemoptysis. Characteristic signs on CT scan include the halo sign and cavitary lesions. Treatment with voriconazole must be prompt and aggressive, including reversal of neutropenia, if there is to be any chance for recovery. Mortality ranges from 93% to 100% in bone marrow transplant recipients, to approximately 38% in kidney transplant recipients, although this improves to approximately 60% at 12 weeks with antifungal therapy. Several other advances in diagnosis and treatment, including CT scans in high-risk populations and development of additional triazoles and echinocandins, have improved the early identification and response to therapy in this patient population. Additional treatment considerations include the use of hematopoietic growth factors to minimize the neutropenic period, which contributes to uncontrolled disease. Surgical removal of the infectious nidus is advocated by some groups because medical treatment has such poor outcomes. Treatment continues until microbiologic clearance is achieved and clinical signs and radiographic imaging indicate resolution of disease. In addition, the patient should no longer be immunosuppressed. If continuation of immunosuppressive medications is required, antifungal therapy should also continue to prevent recurrence of invasive disease.

Cryptococcosis

Cryptococcosis is a subacute or chronic infection caused by Cryptococcus neoformans, a round, budding yeast (5 to 20 μm in diameter) that is sometimes surrounded by a characteristic wide gelatinous capsule. Cryptococci are typically present in soil and dust contaminated by pigeon droppings. When inhaled, such droppings can cause a nonfatal disease primarily affecting the pulmonary and central nervous systems. At present, cryptococcosis is one of the most common opportunistic infections in patients with HIV infection, affecting ~3% of that population. Four basic pathologic patterns are seen in the lungs of infected patients: granulomas; granulomatous pneumonia; diffuse alveolar or interstitial involvement; and proliferation of fungi in alveoli and lung vasculature. Symptoms are nonspecific, as are the radiographic findings. Cryptococcus neoformans may be isolated from sputum, bronchial washings, percutaneous needle aspiration of the lung, or cerebrospinal fluid. If disease is suspected, serum cryptococcal antigen tilters should be obtained; if positive or if the patient has persistent fever, evidence of progression, physiologic compromise, or dissemination, treatment should be promptly initiated. According to the CDC, multiple antifungal agents are effective against C neoformans; asymptomatic infections, such as those identified through targeted screening, should be treated with fluconazole while severe lung infections require amphotericin B combined with flucytosine followed by fluconazole for an extended length of time. Duration of therapy is longer in patients who are immunocompromised.

Candidiasis

Candida organisms are oval, budding cells (with or without mycelial elements) that colonize the oropharynx of many healthy individuals. The fungi of this genus are common hospital and laboratory contaminants. Usually, Candida albicans causes disease in the oral or bronchial mucosa, among other anatomic sites. Approximately 95% of all invasive Candida infections are caused by five species: C albicans, C tropicalis, C parapsilosis, C glabrata, and C krusei. The specific pathogen varies between patient populations and geographic regions. Non-C albicans infections now constitute nearly 70% of all cases in the United States, with C glabrata leading the list. Resistance to fluconazole is common in the non-C albicans species, either natural or developed in response to antifungal therapy, and the shift is likely related to the widespread use of this antifungal agent.

The incidence of Candida infections has increased and is no longer confined to immunocompromised patients. Increasing incidence of infection has been identified in patients with any of the following risk factors: critical illness of long duration; use of long-term antibiotics, particularly multiple; indwelling urinary or vascular catheter; gastrointestinal perforation; or burn wounds. With respect to the thorax, such patients commonly have candidial pneumonia, pulmonary abscess, esophagitis, and mediastinitis. Pulmonary candidal infections typically result in an acute or chronic granulomatous reaction. Because Candida can invade blood vessel walls and a variety of tissues, systemic or disseminated infections can occur, but are less common.

Treatment for candidal infection includes both fungicidal and fungistatic agents. The fungicidal medications include polyenes (amphotericin B deoxycholate [AmB-D] and various lipid-associated amphotericin B preparations) and the echinocandins (caspofungin, micafungin, and anidulafungin). Fungistatic drugs include the triazoles (fluconazole, itraconazole, voriconazole, and posaconazole). The availability of multiple effective therapies allows for specific tailoring of treatment, including combination regimens, based on the patient’s ability to tolerate associated toxicities, the microbiologic information for the specific candidal species, and the route of administration. While demonstrated efficacy is similar, the triazoles and echinocandins appear to have fewer side effects and are better tolerated than the other classes of antifungal drugs. The current initial recommended regimen for adults with invasive candidiasis is an echinocandin.

In addition to prompt institution of antifungal therapy, it is advisable to remove all central venous catheters as soon as can be safely achieved. For fungemia, an eye examination should be performed. Treatment should continue for at least 2 weeks after the last positive blood culture. For patients with Candida mediastinitis (which has a mortality rate of >50%), surgical intervention to debride all infected tissues is required, in addition to prolonged administration of antifungal drugs.

Mucormycosis

The Mucor species, rare members of the class Zygomycetes, are responsible for rapidly fatal disease in immunocompromised patients. Other disease-causing species of the class Zygomycetes include Absidia, Rhizopus, and Mortierella. Characteristic of these fungi are nonseptate, branching hyphae that are difficult to culture. Infection occurs via inhalation of spores. Immunocompromised patients, including patients with neutropenia, acidosis, diabetes, and hematologic malignancy all predispose to clinical susceptibility. In the lungs, disease consists of blood vessel invasion, thrombosis, and infarction of infected organs.

Tissue destruction is significant, along with cavitation and abscess formation. Initial treatment is to correct underlying risk factors and administer antifungal therapies, although
the optimal duration and optimal total dose are unknown. Lipid formulations of amphotericin B are recommended at this time. Surgical resection of any localized disease should be performed after initial medical treatment attempts fail.

**Primary Fungal Pathogens**

*Histoplasma capsulatum* is a dimorphic fungus existing in mycelial form in soil contaminated by fowl or bat excreta and in yeast form in human hosts. The most common of all fungal pulmonary infections, histoplasmosis, primarily affects the respiratory system after spores are inhaled. It is endemic in the Midwest and Mississippi River Valley of the United States, where about 500,000 new cases arise each year. In immunocompromised patients, the infection becomes systemic and more virulent; because cell-mediated immunity is impaired, uninhibited fungal proliferation occurs within pulmonary macrophages and then spreads. Acute forms of the disease present as primary or disseminated pulmonary histoplasmosis; chronic forms present as pulmonary granulomas (*histoplasmosis*), chronic cavitary histoplasmosis, mediastinal granulomas, fibrosing mediastinitis, or bronchiolithiasis. Histoplasmosis is definitively diagnosed by fungal smear, culture, direct biopsy of infected tissues, or serologic testing.

The clinical presentation depends on the inoculum size and on host factors.

Symptoms of acute pulmonary histoplasmosis are fever, chills, headache, chest pain, musculoskeletal pain, and nonproductive cough. CXRs may be normal or may show mediastinal lymphadenopathy and patchy parenchymal infiltrates. Most patients improve in a few weeks; mild to moderate disease can be treated with itraconazole. Amphotericin B is the treatment of choice if moderate symptoms persist for 2 to 4 weeks or if the illness is extensive, including dyspnea and hypoxia, and if patients are immunosuppressed.110

As the pulmonary infiltrates from acute histoplasmosis heal, consolidation into an asymptomatic solitary nodule or histoplasmosma may occur and is usually seen incidentally on radiographs as a coin-shaped lesion. Central and concentric calcification may occur; if so, no further treatment is required. Noncalcification of the lesion requires further diagnostic workup including chest CT scan, needle biopsy, or surgical excision to rule out a malignancy. Figure 19-34 demonstrates the differences in pathologic findings between infections in normal and immunocompromised hosts.111

When lymph nodes and pulmonary granulomas calcify over time, pressure atrophy on the bronchial wall may result in erosion and migration of the granulomatous mass into the bronchus, causing bronchiolithiasis. Typical symptoms include cough, hemoptysis, and dyspnea. Life-threatening complications include massive hemoptysis or bronchoesophageal fistula. In addition to radiography, bronchoscopy should be performed to aid in diagnosis. Definitive treatment requires surgical excision of the bronchial mass and repair of the airway and contiguous structures. Endobronchial debridement is not advised as this can result in massive, fatal bleeding.

Fibrosing mediastinitis is an uncommon manifestation of histoplasmosis but can be fatal due to progressive distortion and compression of the major vessels and central airways.

Diagnosis can be difficult, and symptoms may be present for extended periods, even years, before the diagnosis is made. The differential diagnosis for the disease process includes granulomatous mediastinitis related to recent infection, malignancy, and chronic pulmonary thromboembolism. A trial of itraconazole is worthwhile, although it is not proven to be effective. In cases where radiographic or physiologic improvement is achieved after a trial of 12 week of therapy, continuation of therapy is considered for a full 12 months. In the majority of patients, however, antifungal therapy has not been proven. There is no role for corticosteroids at this time or for antifibrotics. Occasionally, intravascular stents have been helpful for severe vascular compromise. Balloon dilatation and endobronchial silicone stents may be needed for airway compromise, although this should be directed by a surgeon with expertise in mediastinal and airway disease management.

Chronic pulmonary histoplasmosis occurs in about 10% of patients who become symptomatic after infection. Most such patients have preexisting lung pathology, particularly emphysema, which becomes colonized, and subsequent pneumonitis and necrosis, cavity enlargement, new cavity formation, and pulmonary dissemination occur. Nonspecific symptoms, such as cough, sputum production, fever, weight loss, weakness, and hemoptysis are common. Chest radiography may reveal intrapulmonary cavitation and scarring. Occasionally, partial resolution of the inflammatory changes may be observed. Itraconazole provides effective therapy, but must be given for 12 to 24 months. It is superior to ketoconazole and fluconazole; these should only be used if itraconazole is not tolerated. Voriconazole and posaconazole have been found to be useful for salvage therapy. Serum itraconazole levels should be monitored to ensure the drug is being absorbed. Occasionally, lipid-associated amphotericin B is necessary for more severe infections. Surgical excision should be considered in patients patients with adequate pulmonary reserve and localized, thick-walled cavities that have been unresponsive to antifungal therapy.

Disseminated histoplasmosis occurs most frequently in patients who are severely immunocompromised, such as posttransplantation patients, patients with HIV, and patients using immunosuppressive medications. Presentation ranges from nonspecific signs of fever, weight loss, and malaise, to shock, respiratory distress, and multiorgan failure. Diagnosis can be made with a combination of *Histoplasma* urine antigen, serologic assay, and fungal culture and should be suspected in patients with the above symptoms in any endemic area, particularly if the patient is immunosuppressed.112 Any of the antifungal therapies can be used in treatment of disseminated histoplasmosis. Use of amphotericin B has decreased the mortality rate to less than 25% in this type of serious infection.

*Coccidioides immitis* is an endemic fungus found in soil and dust of the southwestern United States. Agricultural workers, military personnel, and other occupations with extensive exposure to soil, especially in areas of endemic growth, are at highest risk, as are immunocompromised individuals.113 Spores (arthroconidia) are inhaled, swell into spherules, and subdivide into endospores, and subsequent infection develops. Diagnosis can be achieved through serum analysis for antecoccidioidal antibody, spherule identification in tissue, or by isolating the fungus in cultures from sputum, other body fluid, or tissue.

Inhalation of the fungus causes pulmonary involvement in 95% of patients with symptomatic disease. Three main categories of pulmonary involvement, based on the associated signs and symptoms, are possible: primary, complicated, and residual pulmonary coccidioidomycosis. Primary pulmonary
Figure 19-34. Pathologic findings of infection in normal and immunocompromised hosts. Histopathologic preparations are shown contrast- ing acute diffuse pulmonary involvement in a lung segment of a normal host with a probable primary infection (A through D) with pulmonary granulomas from an immunocompromised patient who had an opportunistic reinfection with *Histoplasma capsulatum* (E, F). A. Diffuse interstitial pneumonitis in an adult (normal host) with recent heavy environmental exposure and subsequent development of progressive pulmonary disease. There is an inflammatory cell infiltrate primarily involving the interalveolar interstitial spaces but present within many alveolar spaces as well. The exudate consists mostly of mononuclear phagocytes, lymphocytes, and occasional plasma cells. Many of the alveolar walls are markedly thickened (hematoxylin and eosin stain [H&E], ×50). B. Another area from the same lung as in A showing focal vasculitis with an infiltrate of lymphocytes and macrophages (H&E, ×25). C. Relatively large alveolar macrophages packed with single and budding yeasts 2 to 4 µm in diameter (same lung as in A and B). The basophilic cytoplasm of these yeasts is retracted from their thin outer cell walls, leaving halo-like clear areas that can be confused with capsules (H&E, ×500). D. Intracellular and extracellular yeasts, 2 to 4 µm in diameter, some of which are single, budding, or in short chains (Gomori methenamine silver stain, ×500). E. Nonnecrotizing (sometimes called epithelioid cell or noncaseating) granuloma from a patient who had recently received chemotherapy for a germ cell tumor (different patient than in A through D). This lesion consists of a focal collection of macrophages (sometimes referred to as histiocytes or epithelioid cells) plus lymphocytes and occasional plasma cells. A few multinucleated macrophages are present. A thin layer of fibroblasts circumscribes the lesion. Yeasts of *H capsulatum*, probably present within macrophages of this lesion at an earlier stage, were not identified in this granuloma or in any of several other nonnecrotizing granulomas within the specimen. Lesions of this type often undergo necrosis to become necrotizing granulomas (H&E, ×50). F. Necrotizing (sometimes referred to as caseating) granuloma from the same lung as in E. This lesion has a necrotic center surrounded by macrophages, encapsulating fibroblasts, fibrous connective tissue in the periphery, and scattered lymphocytes. A prominent giant cell is present in the lower left of the granuloma (at approximately 8 o’clock). Microorganisms are usually present only in relatively small numbers in these types of lesions. They are most frequently detected within the most central necrotic material in these granulomas (H&E, ×25). (Reproduced with permission from Hage CA, Wheat LJ, Loyd J, et al: Pulmonary histoplasmosis, Semin Respir Crit Care Med. 2008 Apr;29(2):151-165.)
coccidioidomycosis occurs in about 40% of people who inhale spores. The other 60% will remain asymptomatic and develop life-long immunity. The constellation of symptoms of “valley fever,” including fever, chills, headache, erythema multiforme, erythema nodosum, polyarthralgias, nonproductive cough, and chest pain, and a CXR showing hiliar and paratracheal adenopathy are highly suggestive of pulmonary coccidioidomycosis. In many patients, initial diagnosis is community-acquired pneumonia, and it is only when the patient fails to respond to appropriate antibiotic therapy that pulmonary coccidioidomycosis is considered. The disease is self-limited in the majority of patients, and treatment is not required in these cases.

Therapy should be considered for (a) patients with impaired cellular immunity; (b) comorbid illnesses that are adversely impacted by the infection, including chronic pulmonary dysfunction, renal failure, and congestive heart failure; and (c) when symptoms and radiographic findings persist for more than 6 to 8 weeks, at which time the disease is considered to be persistent coccidioidal pneumonia and occurs in approximately 1% of patients. Progression to caseous nodules, cavities, and calcified, fibrotic, or ossified lesions indicates complicated or residual stages of coccidioidomycosis.

There are several relative indications for surgery in pulmonary coccidioidomycosis. A rapidly expanding (>4 cm) cavity that is close to the visceral pleura is a high risk for rupture into the pleural space and subsequent empyema. Other indications for operative intervention include life-threatening hemoptysis; hemoptysis that is persistent despite medical therapy; symptomatic fungus ball; bronchopleural fistula; cavitory lesions with persistent positive sputum; and pulmonary nodules that degenerate over time. Finally, any nodule with signs that are concerning for malignancy should undergo further evaluation, including biopsy or resection, to determine the underlying etiology.

Diagnosis of coccidioidomycosis is confirmed by histopathologic, mycologic, and serologic evaluation. Extrapulmonary disease may develop in approximately 0.5% of infected patients, with involvement of meninges, bones, joints, skin, or soft tissues. Immunocompromised patients are especially susceptible to disseminated coccidioidomycosis, which carries a mortality rate over 40%. Treatment options for this disease vary depending on the severity of the disease as well as the stage. Amphotericin B deoxycholate or the triazoles continue to be the primary antifungal medications. If meningeal involvement is identified, flucytosine or itraconazole therapy is required for the remainder of the patient’s life. Intrathecal amphotericin B can also be administered in some cases.

*Blastomyces dermatitidis* Blastomyces dermatitidis is a round, single-budding yeast with a characteristic thick, refractile cell wall. It resides in the soil as a nonmotive spore called conidia. Exposure occurs when contaminated soil is disturbed and the conidia are aerosolized. The spore is inhaled and transforms into a yeast phase at body temperature. Infection is typically self-limited. A small minority of patients will develop chronic pulmonary infection or disseminated disease, including cutaneous, osteoarticular, and genitourinary involvement. *B dermatitidis* has a worldwide distribution; in the United States, it is endemic in the central states. With chronic infection, the organism induces a granulomatous and pyogenic reaction with microabscesses and giant cells; caseation, cavitation, and fibrosis may also occur. Symptoms are nonspecific and consistent with chronic pneumonia in 60% to 90% of patients. They include cough, mucoid sputum production, chest pain, fever, malaise, weight loss, and, uncommonly, hemoptysis. In acute disease, radiographs are either completely negative or have nonspecific findings; in chronic disease, fibronodular lesions (with or without cavitation) similar to tuberculosis are noted. Pulmonary parenchymal abnormalities in the upper lobe(s) may be noted. Mass lesions similar to carcinoma are frequent, and lung biopsy is frequently used. Over 50% of patients with chronic blastomycosis also have extrapulmonary manifestations, but less than 10% of patients present with severe clinical manifestation.

Once a patient manifests symptoms of chronic blastomycosis, antifungal treatment is required to achieve resolution. Mortality approaches 60% if untreated. While controversial, a short course of triazole therapy (oral itraconazole 200 mg daily) for 6 months is the treatment of choice for most patients with mild to moderate forms of the disease. Because itraconazole has poor CNS penetration, the most common site of recurrence after apparently successful therapy is in the CNS. In the absence of therapy, close follow-up is warranted for evidence of progression to chronic or extrapulmonary disease. Amphotericin B is warranted for patients with severe or life-threatening disease, CNS involvement, disseminated disease, or extensive lung involvement and in immunocompromised patients. After adequate drug therapy, surgical resection of known cavitary lesions should be considered because viable organisms are known to persist in such lesions.

**Massive Hemoptysis**

Massive hemoptysis is generally defined as expectoration of over 600 mL of blood within a 24-hour period. It is a medical emergency associated with a mortality rate of 30% to 50%. Most clinicians would agree that losing over a liter of blood via the airway within 1 day is significant, yet use of an absolute volume criterion presents difficulties. First, it is difficult for the patient or caregivers to quantify the volume of blood being lost. Second, and most relevant, the rate of bleeding necessary to incite respiratory compromise is highly dependent on the individual’s prior respiratory status. For example, the loss of 100 mL of blood over 24 hours is significant, yet use of an absolute volume criterion presents difficulties. First, it is difficult for the patient or caregivers to quantify the volume of blood being lost. Second, and most relevant, the rate of bleeding necessary to incite respiratory compromise is highly dependent on the individual’s prior respiratory status.

Anatomy. The lungs have two sources of blood supply: the pulmonary and bronchial arterial systems. The pulmonary system is a high-compliance, low-pressure system, and the walls of the pulmonary arteries are very thin and delicate. The bronchial arteries, part of the systemic circulation, have systemic pressures and thick walls; most branches originate from the proximal thoracic aorta. Most cases of massive hemoptysis involve bleeding from the bronchial artery circulation or from the pulmonary circulation pathologically exposed to the high pressures of the bronchial circulation. In many cases of hemoptysis, particularly those due to inflammatory disorders, the bronchial arterial tree becomes hyperplastic and tortuous. The systemic pressures within these arteries, combined with a disease process within the airway and erosion, lead to bleeding.

**Causes.** Significant hemoptysis has many causes, most commonly including pulmonary, extrapulmonary, and iatrogenic. Table 19-20 summarizes the most common causes of hemoptysis. Most are secondary to inflammatory processes. Aneurysms
Table 19-20

Pulmonary and extrapulmonary causes of massive hemoptysis

<table>
<thead>
<tr>
<th>PULMONARY</th>
<th>EXTRAPULMONARY</th>
<th>IATROGENIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulmonary parenchymal disease</td>
<td>Congestive heart failure</td>
<td>Intrapulmonary catheter</td>
</tr>
<tr>
<td>Bronchitis</td>
<td>Coagulopathy</td>
<td></td>
</tr>
<tr>
<td>Bronchiectasis</td>
<td>Mitral stenosis</td>
<td></td>
</tr>
<tr>
<td>Tuberculosis</td>
<td>Medications</td>
<td></td>
</tr>
<tr>
<td>Lung abscess</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pneumonia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cavitary fungal infection (e.g., aspergilloma)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lung parasitic infection (ascariasis, schistosomiasis, paragonimiasis)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pulmonary neoplasm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pulmonary infarction or embolism</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trauma</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arteriovenous malformation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pulmonary vasculitis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pulmonary endometriosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wegener’s granulomatosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cystic fibrosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pulmonary hemosiderosis</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

of the pulmonary artery (referred to as Rasmussen’s aneurysm) can develop within pulmonary cavities and can result in massive bleeding. Hemoptysis due to lung cancer is usually mild, resulting in blood-streaked sputum. Massive hemoptysis in patients with lung cancer is typically caused by malignant invasion of pulmonary artery vessels by large central tumors. Although rare, it is often a terminal event.

**Management.** Life-threatening hemoptysis is best managed by a multidisciplinary team of intensive care physicians, interventional radiologists, and thoracic surgeons. Treatment priorities begin with respiratory stabilization; intubation with isolation of the bleeding lung may be required to prevent asphyxiation. This can be done with main-stem intubation into the nonbleeding lung, endobronchial blockers into the bleeding lung, or double-lumen endotracheal intubation, depending on the urgency of the situation and the expertise of the providers. Once adequate ventilation has been achieved, the bleeding site should be localized; bronchoscopy can often provide direct visualization of blood coming from a specific area of the tracheobronchial anatomy. Control of the hemorrhage is then achieved endobronchially with laser or bronchial occlusion, endovascularly with bronchial and/or pulmonary artery embolization, or surgically with resection of the involved area. The order of priorities in management is detailed in Table 19-21.

The clinically pragmatic definition of massive hemoptysis is a degree of bleeding that threatens respiratory stability. Therefore, clinical judgment of respiratory compromise is the first step in evaluating a patient. Two scenarios are possible: (a) bleeding is significant and persistent, but its rate allows a rapid but sequential diagnostic and therapeutic approach, and (b) bleeding is so rapid that emergency airway control and therapy are necessary.

**Scenario 1: Significant, Persistent, but Nonmassive Bleeding**

Although bleeding is brisk in scenario 1, the patient may be able to maintain clearance of the blood and secretions with his or her own respiratory reflexes. Immediate measures are admission to an intensive care unit; strict bed rest; Trendelenburg positioning with the affected side down (if known); administration of humidified oxygen; cough suppression; monitoring of oxygen saturation and arterial blood gases; and insertion of large-bore intravenous catheters. Strict bed rest with sedation may lead to slowing or cessation of bleeding, and the judicious use of intravenous narcotics or other relaxants to mildly sedate the patient and diminish some of the reflexive airway activity is often necessary. Also recommended are administration of aerosolized adrenaline, intravenous antibiotic therapy if needed, and correction of abnormal blood coagulation study results. Finally, unless contraindicated, intravenous vasopressin (20 U over 15 minutes, followed by an infusion of 0.2 U/min) can be given.

A CXR is the first test and often proves to be the most revealing. Localized lesions may be seen, but the effects of blood soiling of other areas of the lungs may predominate, obscuring the area of pathology. Chest CT scan provides more detail and is nearly always performed if the patient is stable. Pathologic areas may be obscured by blood soiling.

Flexible bronchoscopy is the next step in evaluating the patient’s condition. Some clinicians argue that rigid bronchoscopy should always be performed. However, if the patient is clinically stable and the ongoing bleeding is not imminently threatening, flexible bronchoscopy is appropriate. It allows diagnosis of airway abnormalities and will usually permit
localization of the bleeding site to either a lobe or even a segment. The person performing the bronchoscopy must be prepared with excellent suction and must be able to perform saline lavage with a dilute solution of epinephrine.

Most cases of massive hemoptysis arise from the bronchial arterial tree; therefore, the next therapeutic option frequently is selective bronchial arteriography and embolization.

Prearteriogram bronchoscopy is extremely useful to direct the angiographer. However, if bronchoscopy fails to localize the bleeding site, then bilateral bronchial arteriograms can be performed. More recently, use of multidetector CT angiography in patients with hemoptysis that is not immediately life-threatening has been shown to facilitate endovascular intervention; reformatting of the images in multiple projections allows clear delineation of the pulmonary vascular anatomy. With this approach, abnormal bronchial and nonbronchial arteries can be visualized and subsequently targeted for therapeutic arterial embolization. Once the targeted arterial system has been embolized, immediate control and cessation of the hemoptysis is achieved in more than 80% of patients. If bleeding persists after bronchial artery embolization, a pulmonary artery source should be suspected and a pulmonary angiogram performed at the same setting.

Recurrence is seen in 30% to 60% of cases and is very common in the setting of invasive fungal infections such as aspergilloma. Recurrence after bronchial artery embolization is less common in the setting of malignancy and active tuberculosis but does occur and can ultimately result in patient death. Repeat embolization can be effective and is warranted for initial management of recurrent hemoptysis, but early surgical intervention should be considered, particularly in the setting of aspergilloma or other cavitary lesions.

If respiratory compromise is impending, orotracheal intubation should be performed. After intubation, flexible bronchoscopy should be performed to clear blood and secretions and to attempt localization of the bleeding site. Depending on the possible causes of the bleeding, bronchial artery embolization or (if appropriate) surgery can be considered.

**Scenario 2: Significant, Persistent, and Massive Bleeding**

Life-threatening bleeding requires emergency airway control and preparation for potential surgery. Such patients are best cared for in an operating room equipped with rigid bronchoscopy. Immediate orotracheal intubation may be necessary to gain control of ventilation and suctioning. However, rapid transport to the operating room with rigid bronchoscopy should be facilitated. Rigid bronchoscopy allows adequate suctioning of bleeding with visualization of the bleeding site; the nonbleeding side can be cannulated with the rigid scope and the patient ventilated. After stabilization, ice-saline lavage of the bleeding site can then be performed (up to 1 L in 50-mL aliquots); bleeding stops in up to 90% of patients.

Alternatively, blockade of the main stem bronchus of the affected side can be accomplished with a double-lumen endotracheal tube, with a bronchial blocker, or by intubation of the nonaffected side by an uncut standard endotracheal tube. Placement of a double-lumen endotracheal tube is challenging in these circumstances, given the bleeding and secretions.

Proper placement and suctioning may be difficult, and attempts could compromise the patient’s ventilation. The best option is to place a bronchial blocker in the affected bronchus with inflation. Endovascular embolization can be performed to stop the bleeding after control has been achieved with the bronchial blocker. The blocker is left in place for 24 hours; after 24 hours, the area is reexamined bronchoscopically.

**Surgical Intervention.** In most patients, bleeding can be stopped, recovery can occur, and plans to definitively treat the underlying cause can be made. In scenario 1 (significant, persistent, but nonmassive bleeding), the patient may undergo further evaluation as an inpatient or outpatient. A chest CT scan and pulmonary function studies should be obtained preoperatively. In scenario 2 (patients with significant, persistent, and massive bleeding), surgery, if appropriate, will usually be performed during the same hospitalization as the rigid bronchoscopy or main stem bronchus blockade. In a small number of patients (<10%), immediate surgery will be necessary due to the extent of bleeding. The bleeding site in these patients is localized using rigid bronchoscopy with immediate thoracotomy or sternotomy to follow.

Surgical treatment is individualized according to the source of bleeding and the patient’s medical condition, prognosis, and pulmonary reserve. General indications for urgent surgery are presented in Table 19-22. In patients with significant cavitary disease or with fungus balls, the walls of the cavities are eroded and necrotic; rebleeding will likely ensue. In addition, bleeding from cavitary lesions may be due to pulmonary artery erosion, which requires surgery for control.

**End-Stage Lung Disease**

**Lung Volume Reduction Surgery.** The ideal patient for lung volume reduction surgery (LVRS) has heterogeneous emphysema with apical predominance, meaning the worst emphysematous changes are in the apex (seen on chest CT scan) of the lungs. The physiologic lack of function of these areas is demonstrated by quantitative perfusion scan, which shows minimal or no perfusion. By surgically excising these nonfunctional areas, the volume of the lung is reduced, theoretically restoring respiratory mechanics. Diaphragm position and function are improved, and there may be an improvement in the dynamic small airway collapse in the remaining lung.

Operative mortality in the initial experience was 16.9%, with a 1-year mortality of 23%. In response, the National Emphysema Treatment Trial (NETT) performed a randomized trial of 1218 patients in a noncrossover design to medical versus surgical management after a 10-week pretreatment pulmonary rehabilitation program. Subgroup analysis demonstrated that in patients with the anatomic changes delineated by Cooper and colleagues, LVRS significantly improved exercise capacity, lung function, quality of life, and dyspnea compared to medical therapy. After 2 years, functional improvements began to decline toward baseline. Similar parameters in medically treated patients steadily decline below baseline. LVRS was associated with increased short-term morbidity and mortality and did not confer a survival benefit over medical therapy.

**Table 19-22**

<table>
<thead>
<tr>
<th>General indications for urgent operative intervention for massive hemoptysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Presence of a fungus ball</td>
</tr>
<tr>
<td>2. Presence of a lung abscess</td>
</tr>
<tr>
<td>3. Presence of significant cavitary disease</td>
</tr>
<tr>
<td>4. Failure to control the bleeding</td>
</tr>
</tbody>
</table>
Lung Transplantation. The most common indications for lung transplant are COPD and idiopathic pulmonary fibrosis (IPF). Most patients with IPF and older patients with COPD are offered a single-lung transplant. Younger COPD patients and patients with α1-antitrypsin deficiency and severe hyperinflation of the native lungs are offered a bilateral-lung transplant. Most patients with primary pulmonary hypertension and almost all patients with cystic fibrosis are treated with a bilateral-lung transplant. A heart-lung transplant is reserved for patients with irreversible ventricular failure or uncorrectable congenital cardiac disease.

Patients with COPD are considered for placement on the transplant waiting list when their FEV₁ has fallen to below 25% of its predicted value. Patients with significant pulmonary hypertension should be listed earlier. IPF patients should be referred when their forced vital capacity has fallen to less than 60% or their DLCO has fallen to less than 50% of their predicted values.

In the past, patients with primary pulmonary hypertension and New York Heart Association (NYHA) class III or IV symptoms were listed for a lung transplant. However, treatment of such patients with intravenous prostacyclin and other pulmonary vasodilators has now markedly altered that strategy. Virtually all patients with primary pulmonary hypertension are now treated with intravenous epoprostenol. Several of these patients have experienced a marked improvement in their symptoms associated with a decrease in their pulmonary arterial pressures and an increase in exercise capacity. Listing of these patients is deferred until they develop NYHA class III or IV symptoms or until their mean pulmonary artery pressure rises above 75 mmHg.

Medium-term and bronchiolitis obliterans syndrome (BOS)—free survival rates of patients who underwent a lung transplant at the University of Minnesota are shown in Figs. 19-35 and 19-36. The mortality of patients while waiting for transplants is about 10%. In an effort to expand the number of lung donors, many transplant groups have liberalized their criteria for donor selection. Still, the partial pressure of arterial oxygen (Pao₂) should be greater than 300 mmHg on a fraction of inspired oxygen (Fio₂) of 100%. In special circumstances, lungs may be used from donors with a smoking history; from donors older than 50 years of age; and from donors with positive Gram stains or infiltrates on CXR. The use of two living donors, each donating a single lower lobe, is another strategy for increasing the donor pool. Recipient outcomes are similar to those with cadaver donors in carefully selected patients.

Most of the early mortality after lung transplant is related to primary graft failure resulting from a severe ischemia-reperfusion injury to the lung(s) (Fig. 19-37). Reperfusion injury is characterized radiographically by interstitial and alveolar edema and clinically by hypoxia and ventilation-perfusion mismatch. Donor neutrophils and recipient lymphocytes probably play an important role in the pathogenesis of reperfusion injury. The most important impediment to longer-term survival after a lung transplant is the development of BOS, a manifestation of chronic rejection. Episodes of acute rejection are the major risk factors for developing BOS. Other injuries to the lung (including early reperfusion injury and chronic gastroesophageal reflux disease) may also adversely affect long-term outcomes of patients.

CHEST WALL

Chest Wall Mass

Clinical Approach. Surgeons confronted with a patient with a chest wall mass must be cognizant that their approach to diagnosis and treatment has significant impact on the patient’s chances.
for long-term survival. All chest wall tumors should be considered malignant until proven otherwise. It is critically important that the surgeon(s) be mindful of this tenet and well versed in the diagnostic and treatment principles for chest wall malignancies. These tenets must be applied from the initial biopsy, as the placement of the incision can impact significantly on the successful complete resection and reconstruction of the chest wall. Complete resection is imperative if there is any hope for cure and/or long-term survival. A general approach is outlined in Figs. 19-38 and 19-39.

Patients with chest wall tumors, regardless of etiology, typically complain of a slowly enlarging palpable mass (50% to 70%), chest wall pain (25% to 50%), or both. Interestingly, growing masses are often not noticed by the patient until they suffer a trauma to the area. Pain from a chest wall mass is typically localized to the area of the tumor; it occurs more often and more intensely with malignant tumors, but it can also be present in up to one-third of patients with benign tumors. With Ewing’s sarcoma, fever and malaise may also be present. Benign chest wall tumors tend to occur in younger patients (average age 26 years), whereas malignant tumors tend to be found in older patients (average age 40 years). Overall, between 50% and 80% of chest wall tumors are malignant.

**Evaluation and Management.** Laboratory evaluations are useful in assessing chest wall masses for the following:

1. **Plasmacytoma.** Serum protein electrophoresis demonstrates a single monoclonal spike, which is measuring the overproduction of one immunoglobulin from the malignant plasma cell clone.

2. **Osteosarcoma.** Alkaline phosphatase levels may be elevated.

3. **Ewing’s sarcoma.** Erythrocyte sedimentation rates may be elevated.

**Radiography** CXR may reveal rib destruction, calcification within the lesion, and if old films are available, a clue to growth rate. CT scanning, however, is necessary to determine the relationship of the chest wall mass to contiguous structures (e.g., mediastinum, lung, soft tissues, and other skeletal elements), evaluate for pulmonary metastases, and assess for extraosseous bone formation and bone destruction, both typically seen with osteosarcoma.

Because MRI provides multiple planes of imaging (coronal, sagittal, and oblique), better definition of the relationship between tumor and muscle, and tumor and contiguous or nearby neurovascular structures or the spine, it is an important radiographic adjunct for preoperative planning. Compared to CT scan alone, MRI may further delineate tissue abnormalities, potentially enhancing the ability to distinguish benign from malignant sarcoma.

**Biopsy** The first step in the management of all chest wall tumors is to obtain a tissue diagnosis. Inappropriate or misguided attempts at tissue diagnosis through casual open biopsy techniques have the potential (if the lesion is a sarcoma) to seed surrounding tissues and contiguous body cavities (e.g., the pleural space) with tumor cells, potentially compromising local

---

**Figure 19-38.** Systematic approach for evaluating a chest wall mass when the clinical scenario is uncomplicated and initial imaging studies suggest a clear diagnosis. CT = computed tomography; MRI = magnetic resonance imaging.

**Figure 19-39.** Systematic approach for evaluating a chest wall mass for which the diagnosis is not unequivocal. A tissue diagnosis is critical for effective management of chest wall masses. CT = computed tomography; MRI = magnetic resonance imaging; PNET = primitive neuroectodermal tumor.
tumor control and patient survival. Tissue diagnosis is accomplished using one of three methods: needle biopsy (typically CT-guided, FNA, or a core biopsy), incisional biopsy, or excisional biopsy in limited and specific situations.

1. Needle biopsy. Pathologists experienced with sarcomas can accurately diagnose approximately 90% of patients using FNA cytology. A needle biopsy (FNA or core) has the advantage of avoiding wound and body cavity contamination (a potential complication with an incisional biopsy).

2. Incisional biopsy. If a needle biopsy is nondiagnostic, an incisional biopsy may be performed, with caveats. First, the skin incision must be placed directly over the mass and oriented to allow subsequent scar excision and skin flaps, and drains should be avoided. If the surgeon believes a hematoma is likely to develop, a drain is useful for limiting soft tissue contamination by tumor cells. At the time of definitive surgical resection, the en bloc resection includes the biopsy scar and the drain tract along with the tumor.

3. Excisional biopsy. Any lesion less than 2.0 cm can be excised as long as the resulting wound is small enough to close primarily. Otherwise, excisional biopsy is performed only when the initial diagnosis (based on radiographic evaluation) indicates that the lesion is benign or when the lesion has the classic appearance of a chondrosarcoma (in which case, definitive surgical resection can be undertaken).

**Benign Chest Wall Neoplasms**

1. Chondroma. Chondromas, seen primarily in children and young adults, are one of the more common benign tumors of the chest wall. They usually occur at the costochondral junction anteriorly and may be confused with costochondritis, except that a painless mass is present. Radiographically, the lesion is lobulated and radiodense; it may have diffuse or focal calcifications; and it may displace the bony cortex without penetration. Chondromas may grow to huge sizes if left untreated. Treatment is surgical resection with a 2-cm margin. Large chondromas may harbor well-differentiated chondrosarcoma and should be managed with a 4-cm margin to prevent local recurrence.

2. Fibrous dysplasia. As with chondromas, fibrous dysplasia most frequently occurs in young adults and may be associated with trauma. Pain is an infrequent complaint, and the lesion is typically located in the posterolateral aspect of the rib cage. Radiographically, an expansile mass is present, with cortical thinning and no calcification. Local excision with a 2-cm margin is curative.

3. Osteochondroma. Osteochondromas, often found incidentally as a solitary lesion on radiograph, are the most common benign bone tumor. Osteochondromas occur in the first two decades of life, and they arise at or near the growth plate of bones. Osteochondromas in the thorax arise from the rib cortex. They are one of several components to the autosomal dominant syndrome, hereditary multiple exostoses. When part of this syndrome, osteochondromas have a high rate of degeneration into chondrosarcomas. Any patient with hereditary multiple exostoses syndrome who develops new pain at the site of an osteochondroma or who notes gradual growth in the mass over time should be carefully evaluated for osteosarcoma. Local excision of a benign osteochondroma is sufficient. If malignancy is determined, wide excision is performed with a 4-cm margin.

4. Eosinophilic granuloma. Eosinophilic granulomas are benign osteolytic lesions. Eosinophilic granulomas of the ribs can occur as solitary lesions or as part of a more generalized disease process of the lymphoreticular system termed Langerhans cell histiocytosis (LCH). In LCH, the involved tissue is infiltrated with large numbers of histiocytes (similar to Langerhans cells seen in skin and other epithelia), which are often organized as granulomas. The cause is unknown. Of all LCH bone lesions, 79% are solitary eosinophilic granulomas, 7% involve multiple eosinophilic granulomas, and 14% belong to other forms of more systemic LCH. Isolated single eosinophilic granulomas can occur in the ribs or skull, pelvis, mandible, humerus, and other sites. They are diagnosed primarily in children between the ages of 5 and 15 years. Because of the associated pain and tenderness, they may be confused with Ewing’s sarcoma or with an inflammatory process such as osteomyelitis. Healing may occur spontaneously, but the typical treatment is limited surgical resection with a 2-cm margin.

5. Desmoid tumors. Soft tissue neoplasms arising from fascial or musculopaponeurotic structures, desmoid tumors consist of proliferations of benign-appearing fibroblastic cells, abundant collagen, and few mitoses. Desmoid tumors possess alterations in the adenomatous polyposis coli (APC)/β-catenin pathway. Cyclin D1 dysregulation is thought to play a significant role in their pathogenesis. Associations with other diseases and conditions are well documented, especially those with similar alterations in the APC pathway, such as familial adenomatous polyposis (Gardner’s syndrome). Other conditions with increased risk of desmoid tumor formation include increased estrogen states (pregnancy) and trauma. Surgical incisions (abdominal and thorax) have been the site of desmoid development, either in or near the scar.

Clinically, patients are usually in the third to fourth decade of life and have pain, a chest wall mass, or both. The tumor is usually fixed to the chest wall, but not to the overlying skin. There are no typical radiographic findings, but MRI may delineate muscle or soft tissue infiltration. Desmoid tumors do not metastasize, but they have a significant propensity to recur locally, with rates as high as 5% to 50%, sometimes despite complete initial resection with histologically negative margins. Such locally aggressive behavior is secondary to microscopic tumor infiltration of muscle and surrounding soft tissues and prompts some to consider them a low-grade form of fibrosarcoma.

Because the lesions have low cellularity and poor yield with FNA, an open incisional biopsy for lesions over 3 to 4 cm is often necessary, following the caveats listed earlier (see biopsy section). Surgery consists of wide local excision with a 2- to 4-cm margin and intraoperative frozen section assessment of resection margins. Typically, chest wall resection, including the involved rib(s) and one rib above and below the tumor with a 4- to 5-cm margin of rib, is required. A margin of less than 1 cm results in much higher local recurrence rates. If a major neurovascular structure would have to be sacrificed, leading to high morbidity, then a margin of less than 1 cm would have to suffice. Survival after wide local excision with negative margins is 90% at 10 years.
Primary Malignant Chest Wall Tumors

Malignant tumors of the chest wall are either metastatic lesions from another primary tumor or sarcoma. Soft tissue sarcomas of the chest wall include fibrosarcomas, liposarcomas, malignant fibrous histiocytomas (MFHs), rhabdomyosarcomas, angiosarcomas, and other extremely rare lesions (Fig. 19-40). Despite the prevalence of localized disease, soft tissue sarcomas of the chest wall have significantly worse survival than similar tumors located on the extremities or the head and neck region. The factors impacting on risk of death from soft tissue sarcomas of the chest wall are presented in Table 19-23. All sarcomas have a propensity to spread to the lungs.

While many varieties of sarcoma exist, the primary features affecting prognosis are histologic grade and responsiveness to chemotherapy (Table 19-24). Preoperative (neoadjuvant) chemotherapy offers the ability to (a) assess tumor chemosensitivity by the degree of tumor size reduction and microscopic necrosis; (b) determine tumor sensitivity to specific chemotherapeutic agents; and (c) improve resectability by reducing tumor size. Patients whose tumors are responsive to preoperative chemotherapy have a much better prognosis than those with a poor response. Information about tumor response to chemotherapy, the patient’s physiologic state and capacity to receive treatment, and metastatic disease status is used to determine optimal therapy. The initial treatment is either (a) preoperative chemotherapy (for patients with osteosarcoma, rhabdomyosarcoma, primitive neuroectodermal tumor, or Ewing’s sarcoma) followed by surgery and postoperative chemotherapy; (b) primary surgical resection and reconstruction (for patients with nonmetastatic MFH, fibrosarcoma, liposarcoma, or synovial sarcoma); or (c) preoperative chemotherapy followed by surgical resection if indicated in patients presenting with metastatic soft tissue sarcomas. Contiguous involvement of underlying lung or other soft tissues or the presence of pulmonary metastases does not preclude successful surgery. In fact, patients receiving surgical intervention have significantly better overall survival. Median survival with surgical resection is 25 months compared to 8 months without resection. Additional prognostic variables that are important for long-term survival include tumor size, grade, stage, and negative re-resection margin. With the exception of rhabdomyosarcomas, the primary treatment of these lesions is wide surgical resection with 4-cm margins and reconstruction.

The following is an overview of several chest wall sarcomas.

1. **Chondrosarcoma.** Chondrosarcomas are the most common primary chest wall malignancy. As with chondromas, they usually arise anteriorly from the costochondral arches. CT scan shows a radiolucent lesion often with stippled calcifications pathognomonic for chondrosarcomas (Fig. 19-41). The involved bony structures are also destroyed. Most chondrosarcomas are slow-growing, low-grade tumors; these often painful masses can reach massive proportions. For this reason, any lesion in the anterior chest wall likely to be a low-grade chondrosarcoma should be treated with wide (4-cm) resection after metastatic disease to the lungs or bones is ruled out.

   Chondrosarcomas are not sensitive to radiation or chemotherapy. Prognosis is determined by tumor grade and extent of resection. With a low-grade tumor and wide resection, patient survival at 5 to 10 years can be as high as 60% to 80%.

2. **Osteosarcoma.** While osteosarcomas are the most common bone malignancy, they represent only 10% to 15% of all malignant chest wall tumors. They primarily occur in young adults as rapidly enlarging, painful masses; however, osteosarcomas can occur in older patients as well.

![Figure 19-40](image-url)  
Figure 19-40. Chest computed tomography scan showing a right chest wall tumor (arrow). Tissue diagnosis revealed that this mass was a leiomyosarcoma.
Table 19-23
Cox proportional hazards model for risk of death from soft tissue sarcoma

<table>
<thead>
<tr>
<th>Gender</th>
<th>N</th>
<th>HAZARD RATIO</th>
<th>95% CI</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>3937</td>
<td>Reference group</td>
<td>0.897</td>
<td>Reference group</td>
</tr>
<tr>
<td>Female</td>
<td>4113</td>
<td>Reference group</td>
<td>0.843–0.955</td>
<td>Reference group</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 years</td>
<td></td>
<td>Reference group</td>
<td>1.131</td>
<td>Reference group</td>
</tr>
<tr>
<td>51–70 years</td>
<td></td>
<td>Reference group</td>
<td>1.026–1.247</td>
<td>Reference group</td>
</tr>
<tr>
<td>&gt;70 years</td>
<td></td>
<td>Reference group</td>
<td>1.395–1.697</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>7152</td>
<td>Reference group</td>
<td>1.212</td>
<td>Reference group</td>
</tr>
<tr>
<td>Non-Caucasian</td>
<td>898</td>
<td>Reference group</td>
<td>1.093–1.344</td>
<td>Reference group</td>
</tr>
<tr>
<td>Histologic type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fibrosarcoma</td>
<td></td>
<td>Reference group</td>
<td>1.281</td>
<td>Reference group</td>
</tr>
<tr>
<td>MFH</td>
<td>489</td>
<td>Reference group</td>
<td>1.097–1.495</td>
<td>Reference group</td>
</tr>
<tr>
<td>Liposarcoma</td>
<td>2529</td>
<td>Reference group</td>
<td>0.759–1.054</td>
<td>.182</td>
</tr>
<tr>
<td>LMS/GIST</td>
<td>1534</td>
<td>Reference group</td>
<td>1.033–1.403</td>
<td>.018</td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Head and neck</td>
<td>576</td>
<td>Reference group</td>
<td>1.255</td>
<td>Reference group</td>
</tr>
<tr>
<td>Trunk</td>
<td>4054</td>
<td>Reference group</td>
<td>1.096–1.438</td>
<td>.960</td>
</tr>
<tr>
<td>Extremity</td>
<td>2474</td>
<td>Reference group</td>
<td>0.875–1.151</td>
<td>.002</td>
</tr>
<tr>
<td>Retroperitoneum</td>
<td>946</td>
<td>Reference group</td>
<td>1.093–1.489</td>
<td></td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Localized</td>
<td>5006</td>
<td>Reference group</td>
<td>1.575</td>
<td>Reference group</td>
</tr>
<tr>
<td>Regional</td>
<td>1724</td>
<td>Reference group</td>
<td>1.458–1.702</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Distant</td>
<td>1320</td>
<td>Reference group</td>
<td>2.660–3.155</td>
<td></td>
</tr>
<tr>
<td>Surgical treatment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>6754</td>
<td>Reference group</td>
<td>1.562</td>
<td>Reference group</td>
</tr>
<tr>
<td>No</td>
<td>1296</td>
<td>Reference group</td>
<td>1.443–1.691</td>
<td></td>
</tr>
<tr>
<td>Radiation therapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>2175</td>
<td>Reference group</td>
<td>1.151</td>
<td>Reference group</td>
</tr>
<tr>
<td>No</td>
<td>5875</td>
<td>Reference group</td>
<td>1.070–1.239</td>
<td></td>
</tr>
<tr>
<td>Chemotherapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1062</td>
<td>Reference group</td>
<td>0.909</td>
<td>Reference group</td>
</tr>
<tr>
<td>No</td>
<td>6988</td>
<td>Reference group</td>
<td>0.829–0.996</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CI = confidence interval; GIST = gastrointestinal stromal tumor; LMS = leiomyosarcoma; MFH = malignant fibrous histiocytoma.

Table 19-24
Classification of sarcomas by therapeutic response

<table>
<thead>
<tr>
<th>TUMOR TYPE</th>
<th>CHEMOTHERAPY SENSITIVITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Osteosarcoma</td>
<td>+</td>
</tr>
<tr>
<td>Rhabdomyosarcoma</td>
<td>+</td>
</tr>
<tr>
<td>Primitive neuroectodermal tumor</td>
<td>+</td>
</tr>
<tr>
<td>Ewing’s sarcoma</td>
<td>+</td>
</tr>
<tr>
<td>Malignant fibrous histiocytoma</td>
<td>±</td>
</tr>
<tr>
<td>Fibrosarcoma</td>
<td>±</td>
</tr>
<tr>
<td>Liposarcoma</td>
<td>±</td>
</tr>
<tr>
<td>Synovial sarcoma</td>
<td>±</td>
</tr>
</tbody>
</table>

sometimes in association with previous radiation, Paget’s disease, or chemotherapy. Radiographically, the typical appearance consists of spicules of new periosteal bone formation producing a sunburst appearance. Osteosarcomas have a propensity to spread to the lungs, and up to one-third of patients present with metastatic disease. Osteosarcomas are potentially sensitive to chemotherapy. Currently, pre-operative chemotherapy is common. After chemotherapy, complete resection is performed with wide (4-cm) margins, followed by reconstruction. In patients presenting with lung metastases that are potentially amenable to surgical resection, induction chemotherapy may be given, followed by surgical resection of the primary tumor and of the pulmonary metastases. Following surgical treatment of known disease, additional maintenance chemotherapy is usually recommended.
3. **Malignant fibrous histiocytoma.** Originally thought to derive from histiocytes because of the microscopic appearance of cultured tumor cells, these tumors likely originate from the fibroblast. MFHs are generally the most common soft tissue sarcoma of late adult life, although they are rare on the chest wall. The typical age at presentation is between age 50 and 70 years. Presentation is pain, with or without a palpable mass. Radiographically, a mass is usually evident, with destruction of surrounding tissue and bone. Treatment is wide resection with a margin of 4 cm or more and reconstruction. Over two-thirds of patients suffer from distant metastasis or local recurrence.

4. **Liposarcoma.** Liposarcomas make up 15% of chest wall sarcomas. Most liposarcomas are low-grade tumors that have a propensity to recur locally, given their infiltrative nature. They typically present as a painless mass. Treatment is wide resection and reconstruction. Intraoperative margins should be evaluated (as with all sarcomas) and resection continued, if feasible, until margins are negative. Local recurrence can be treated with reexcision, with occasional use of radiotherapy.

5. **Fibrosarcoma.** Often presenting as a large, painful mass, these lesions are visible on plain radiograph or CT, with surrounding tissue destruction. Treatment is wide local excision with intraoperative frozen-section analysis of margins, followed by reconstruction. Local and systemic recurrence is frequent. Patient survival at 5 years is about 50% to 60%.

6. **Rhabdomyosarcoma.** Rhabdomyosarcomas are rare tumors of the chest wall. Microscopically, they are a spindle cell tumor. The diagnosis often depends on immunohistochemical staining for muscle markers. Rhabdomyosarcomas are sensitive to chemotherapy. Treatment consists of preoperative chemotherapy with subsequent surgical resection.

**Other Tumors of the Chest Wall**

1. **Primitive neuroectodermal tumors (PNETs) and Ewing’s sarcoma.** PNETs (neuroblastomas, ganglioneuroblastomas, and ganglioneuromas) derive from primordial neural crest cells that migrate from the mantle layer of the developing spinal cord. Histologically, PNETs and Ewing’s sarcomas are small, round cell tumors; both possess a translocation between the long arms of chromosomes 11 and 22 within their genetic makeup. They also share a consistent pattern of proto-oncogene expression and have been found to express the product of the *MIC2* gene.

   Ewing’s sarcoma occurs in adolescents and young adults who present with progressive chest wall pain, but without the presence of a mass. Systemic symptoms of malaise and fever are often present. Laboratory studies reveal an elevated erythrocyte sedimentation rate and mild white blood cell elevation. Radiographically, the characteristic onion peel appearance is produced by multiple layers of periosteum in the bone formation. Evidence of bony destruction is also common. The diagnosis can be made by a percutaneous needle biopsy or an incisional biopsy.

   These tumors have a strong propensity to metastasize to the lungs and skeleton; patient survival rates are thus only 50% or less at 3 years. Increasing tumor size is associated with decreasing survival. Treatment has improved significantly and now consists of multiagent chemotherapy, radiation therapy, and surgery. Patients are typically treated preoperatively with chemotherapy and reevaluated with radiologic imaging. When residual disease is identified, surgical resection and reconstruction are performed followed by maintenance chemotherapy.

2. **Plasmacytoma.** Solitary plasmacytomas of the chest wall are very rare, with approximately 25 to 30 cases per year in the United States. The typical presentation is pain without a palpable mass. Plain radiographs show an osteolytic lesion in the region of the pain. As with other chest wall tumors, a needle biopsy under CT guidance is performed for diagnosis. Histologically, the lesion is identical to multiple myeloma, with sheets of plasma cells. It occurs at an average age of 55 years. Evaluation for systemic myeloma is performed with bone marrow aspiration, testing of calcium levels, and measurement of urinary Bence Jones proteins. If the results of these studies are negative, then a solitary
plasmacytoma is diagnosed. Surgery is usually limited to a biopsy only, which may be excisional.\textsuperscript{134} Treatment consists of radiation with doses of 4000 to 5000 cGy. Up to 75\% of patients develop systemic multiple myeloma with 10-year survival of approximately 20\%.

**Chest Wall Reconstruction**

The primary determinant of long-term freedom from recurrence and overall survival is margin status; therefore, adequate margins of normal tissue must be included in the en bloc resection. En bloc resection should include involved ribs, sternum, superior sulcus, or spine if necessary; invasion of these structures should not be considered a contraindication to surgery in an otherwise fit patient. The resection should include at least one normal adjacent rib above and below the tumor, with all intervening intercostal muscles and pleura. In addition, an en bloc resection of overlying chest wall muscles is often necessary, such as of the pectoralis minor or major, serratus anterior, or latissimus dorsi. When the periphery of the lung is involved with the neoplasm, it is appropriate to resect the adjacent part of the pulmonary lobe in continuity (Fig. 19-42). Involvement of the sternum by a malignant tumor requires total resection of the sternum with the adjacent cartilage. Techniques for postoperative respiratory support are now good enough that resection should not be compromised because of any concern about the patient’s ability to be adequately ventilated in the early postoperative period.

The extent of resection depends on the tumor’s location and on any involvement of contiguous structures. Laterally based lesions often require simple wide excision, with resection of any contiguously involved lung, pleura, muscle, or skin. Anteriorly based lesions contiguous with the sternum require partial sternectomy. Primary malignant tumors of the sternum may require complete sternectomy. Posterior lesions involving the rib heads over their articulations with the vertebral bodies may, depending on the extent of rib involvement, require partial en bloc vertebrectomy.

Optimal management of larger tumors includes careful preoperative planning and execution of the surgery by the thoracic surgeon and an experienced plastic surgeon in order to ensure optimal physiologic and cosmetic results. With these measures, reconstruction at the same operation can be accomplished.\textsuperscript{135} Reconstruction of a large defect in the chest wall requires the use of some type of material to prevent lung herniation and to provide stability for the chest wall (see Fig. 19-42). Mild degrees of paradoxical motion are often well tolerated if the area of instability is relatively small. Historically, a wide variety of materials have been used to reestablish chest wall stability, including rib autografts, steel struts, acrylic plates, and numerous synthetic meshes. The current preference is either a 2-mm polytetrafluoroethylene (Gore-Tex) patch or a double-layer polypropylene (Marlex) mesh sandwiched with methylmethacrylate. There are several properties that make Gore-Tex an excellent material for use in chest wall reconstruction: (a) it is impervious to fluid, which prevents pleural fluid from entering the chest wall and minimizes the formation of seromas, which can compromise the myocutaneous flap viability and provide a nidus for infection; and (b) it provides excellent rigidity and stability when secured taut to the surrounding bony structure and, as a result, provides a firm platform for myocutaneous flap reconstruction. Except for smaller lesions, tissue coverage requires the use of myocutaneous flaps (latissimus dorsi, serratus anterior, rectus abdominis, or pectoralis major muscles).\textsuperscript{136,137}

**MEDIASTINUM**

**Anatomy and Pathologic Entities**

The mediastinum can be divided into compartments for classification of anatomic components and disease processes, which, despite substantial overlap, facilitates understanding of general concepts of surgical interest. Several classification schemes exist, but for the purposes of this chapter, the three-compartment model is used (Fig. 19-43). The anterior compartment lies between the sternum and the anterior surface of the heart and great vessels. The visceral or middle compartment is located between the great vessels and the trachea. As the name implies, the posterior compartment lies posterior and includes the paravertebral sulci, bilaterally, and the paraesophageal area.

**Figure 19-42.** Principles of reconstruction after resection of a chest wall tumor (osteogenic sarcoma) are shown. A. En bloc resection of the involved chest wall, including normal ribs above and below the tumor as well as pulmonary parenchyma, must be performed. The resected specimen is shown. B. A prosthesis has been sewn in place. In the lower third of the prosthesis, the line of diaphragm reattachment is seen. The skin defect was closed with a myocutaneous flap from the ipsilateral rectus muscle.
CHAPTER 19
CHEST WALL, LUNG, MEDIASTINUM, AND PLEURA

Posterior mediastinum

Anterosuperior mediastinum

Middle mediastinum

Figure 19-43. Anatomic division of the mediastinum.

The normal content of the anterior compartment includes the thymus gland or its remnant, the internal mammary artery and vein, lymph nodes, and fat. The thymus gland is large during childhood, occupying the entire anterior mediastinum (Fig. 19-44) but decreases in both thickness and length after adolescence and takes on a more fatty content, with only residual islands of thymic cellular components (Fig. 19-45). The middle mediastinal compartment contains the pericardium and its contents, the ascending and transverse aorta, the superior and inferior venae cavae, the brachiocephalic artery and vein, the phrenic and upper vagus nerves, the trachea and main bronchi and corresponding lymph nodes, and the central portions of the pulmonary arteries and veins. The posterior compartment contains the descending aorta, esophagus, thoracic duct, azygos and hemiazygos veins, and lymph nodes. Numerous pathologic variants may be present in the various compartments, with much overlap. Table 19-25 includes the most common pathologic entities listed by compartment.138,139

History and Physical Examination
Mediastinal pathology varies significantly by patient age. In children, neurogenic tumors of the posterior mediastinum are most common, followed by lymphoma, which is usually located in the anterior or middle compartment. Thymoma in childhood is rare (Table 19-26). In adults, the most common tumors include neurogenic tumors of the posterior compartment, benign cysts occurring in any compartment, and thymomas of the anterior mediastinum (Table 19-27). In both age groups, about 25% of mediastinal tumors are malignant. Pediatric tumors will be discussed in Chapter 39.

Up to two-thirds of mediastinal tumors in adults are discovered as asymptomatic abnormalities on radiologic studies ordered for other problems, particularly now that screening CT examinations are more prevalent. When symptomatic, these tumors are significantly more likely to be malignant. Characteristics such as size, location, rate of growth, and associated inflammation are important factors that correlate with symptoms. Large, bulky tumors, expanding cysts, and teratomas can cause compression of mediastinal structures, in particular the trachea, and lead to cough, dyspnea on exertion, or stridor. Chest pain or dyspnea may be reported secondary to associated pleural effusions, cardiac tamponade, or phrenic nerve involvement. Occasionally, a mediastinal mass near the aortopulmonary window may be identified in a workup for hoarseness because of left recurrent laryngeal nerve involvement. The patient in Fig. 19-46 presented with hoarseness due to nodal compression of the left recurrent laryngeal nerve from a primary lung cancer with metastases to the level 5 and 6 lymph nodes in the region of the aortopulmonary window.

The history and physical examination in conjunction with the imaging findings may suggest a specific diagnosis (Table 19-28). In one series, systemic symptoms were present in 50% of patients with a mediastinal mass and a lymphoproliferative disorder, as compared with only 29% of patients with other masses (such as thymic or neurogenic). Laboratory signs of inflammation were also noted; the erythrocyte sedimentation rate and C-reactive protein levels were elevated and leukocytosis was present in 86% of patients with a lymphoproliferative disorder, as compared with only 58% of patients with other types of mediastinal masses.

Imaging and Serum Markers
Chest CT or MRI is required to fully delineate the anatomy.140 A contrast-enhanced CT scan enables clear delineation of the

Figure 19-44. Normal appearance of the thymus gland in childhood. Ao = aorta; PA = pulmonary artery; VC = vena cava.
soft tissue structures from the vasculature and is preferred over noncontrast studies. If there is concern for invasion of vascular structures or spinal involvement, MRI is more accurate than CT scan and provides important information regarding respectability.

If an endocrine origin is suspected, several other imaging modalities are available (Table 19-29). Single-photon emission CT (SPECT) technology may be used to improve image contrast and give information on three-dimensional localization, largely replacing conventional two-dimensional nuclear imaging studies. If a thyroid origin is suspected, a thyroid scan using $^{131}$I or $^{123}$I can identify most intrathoracic goiters and identify the extent of functioning thyroid tissue. If indicated, the thyroid scan should precede other scans requiring iodine-containing contrast agents because they would subsequently interfere with iodine tracer uptake by thyroid tissue. If a pheochromocytoma or neuroblastoma is suspected, the octreotide scan or $^{123}$I-metaiodobenzylguanidine (MIBG) scans are helpful in diagnosis and localization. The sestamibi scan may be useful for diagnosing and localizing a mediastinal parathyroid gland. PET is useful for distinguishing malignant from benign tumors and may help detect distant metastases in some patients. However, the role of routine PET imaging for staging surgically resectable lesions of the mediastinum has not been established.

The use of serum markers to evaluate a mediastinal mass can be invaluable in some patients. For example, nonseminomatous and seminomatous germ cell tumors can frequently be diagnosed and often distinguished from one another by the levels of α-fetoprotein (AFP) and human chorionic gonadotropin (hCG). In over 90% of nonseminomatous germ cell tumors, either the AFP or the hCG level will be elevated. Results are close to 100% specific if the level of either AFP or hCG is greater than 500 ng/mL. Some centers institute chemotherapy based on this result alone, without biopsy confirmation of the diagnosis. In contrast, the AFP level in patients with mediastinal seminoma is always normal; only 10% will have elevated hCG, which is usually less than 100 ng/mL. Other serum markers, such as intact parathyroid hormone level for ectopic parathyroid

### Table 19-25

**Usual location of the common primary tumors and cysts of the mediastinum**

<table>
<thead>
<tr>
<th>ANTERIOR COMPARTMENT</th>
<th>VISCERAL COMPARTMENT</th>
<th>PARAVERTEBRAL SULCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thymoma</td>
<td>Enterogenous cyst</td>
<td>Neurilemoma-schwannoma</td>
</tr>
<tr>
<td>Germ cell tumor</td>
<td>Lymphoma</td>
<td>Neurofibroma</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>Pleuropericardial cyst</td>
<td>Malignant schwannoma</td>
</tr>
<tr>
<td>Lymphangioma</td>
<td>Mediastinal granuloma</td>
<td>Ganglioneuroma</td>
</tr>
<tr>
<td>Hemangioma</td>
<td>Lymphoid hamartoma</td>
<td>Ganglioneuroblastoma</td>
</tr>
<tr>
<td>Lipoma</td>
<td>Mesothelial cyst</td>
<td>Neuroblastoma</td>
</tr>
<tr>
<td>Fibroma</td>
<td>Neuroenteric cyst</td>
<td>Paraganglioma</td>
</tr>
<tr>
<td>Fibrosarcoma</td>
<td>Paraganglioma</td>
<td>Pheochromocytoma</td>
</tr>
<tr>
<td>Thymic cyst</td>
<td>Pheochromocytoma</td>
<td>Fibrosarcoma</td>
</tr>
<tr>
<td>Parathyroid adenoma</td>
<td>Thoracic duct cyst</td>
<td>Lymphoma</td>
</tr>
</tbody>
</table>


### Table 19-26

**Mediastinal tumors in children**

<table>
<thead>
<tr>
<th>TUMOR TYPE</th>
<th>PERCENTAGE OF TOTAL</th>
<th>LOCATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neurogenic tumors</td>
<td>40</td>
<td>Posterior</td>
</tr>
<tr>
<td>Lymphomas</td>
<td>18</td>
<td>Anterior/middle</td>
</tr>
<tr>
<td>Cysts</td>
<td>18</td>
<td>All</td>
</tr>
<tr>
<td>Germ cell tumors</td>
<td>11</td>
<td>Anterior</td>
</tr>
<tr>
<td>Mesenchymal tumors</td>
<td>9</td>
<td>All</td>
</tr>
<tr>
<td>Thymomas</td>
<td>Rare</td>
<td>Anterior</td>
</tr>
</tbody>
</table>

adenomas, may be useful for diagnosing and also for intraoperatively confirming complete resection. After successful resection of a parathyroid adenoma, this hormone level should rapidly normalize.

**Diagnostic Nonsurgical Biopsies of the Mediastinum**

The treatment of up to 60% of patients with anterior mediastinal masses is ultimately nonsurgical, so it is essential to understand all options for obtaining adequate tissue for a definitive diagnosis using the least invasive approach. CT-guided needle biopsy, EBUS- and EUS-guided FNA, and even core-needle biopsy (either CT-guided and EUS-guided) have proven most useful for cytologic and tissue diagnosis of mediastinal masses and lymphadenopathy.

When FNA and core-needle biopsy were combined, the accuracy was 98%, compared to 79% for each modality independently. In addition, core-needle biopsy changed the diagnosis in nine cases that had been missed by FNA due to inadequate specimens. Finally, core-needle biopsy was better at diagnosis for benign diseases compared to FNA. Accessible nodal stations include subcarinal (level 7), aortopulmonary (level 5), paraeosophageal (level 8), and inferior pulmonary ligament (level 9) as well as paratracheal (level 4). Technical expertise in these modalities should be pursued by thoracic and general surgeons.

Historically, needle biopsies of anterior mediastinal masses were reportedly sensitive and specific for most carcinomatous tumors, but there were questions regarding accuracy for diagnosing lymphomas. However, advances in cytopathology as well as needle biopsy technology have substantially improved diagnostic accuracy such that most centers are reporting yields ranging from 75% to 80% for the diagnosis of lymphoma as well.

To achieve maximal diagnostic yield for mediastinal masses suggestive of a lymphoma, it is necessary to obtain multiple fine-needle aspirates, preferably with immediate onsite rapid cytologic analysis to confirm sampling of the target tissue and adequate cellularity. This also facilitates processing of the sample to ensure that proper studies for lymphoma, including flow cytometry, are obtained. If the needle biopsy is inconclusive, surgical biopsy can be performed. If the lesion is accessible by CT-guided or EUS-guided core-needle biopsy, intraoperative frozen section or immediate cytologic smear of a core biopsy can also be performed. Currently, core-needle biopsy with EBUS is not possible. The authors perform their own endobronchial, endoscopic, and CT-guided transbronchial and transthoracic biopsies, and in our experience, lack of cellularity in the aspirate is readily apparent.

In general, plans to proceed with surgical biopsy are made in combination with the image-guided aspiration and, as such, are performed in the same setting. This enables the authors to avoid a more invasive surgical procedure when FNA or core-needle biopsy is sufficient without contributing to delays in diagnosis by having multiple attempts from multiple providers (such as interventional radiology and pulmonology) before involvement of the surgeon in the diagnostic workup.
**Surgical Biopsies and Resection of Mediastinal Masses**

For tumors of the mediastinum that are not amenable to an endoscopic or CT-guided needle biopsy or that do not yield sufficient tissue for diagnosis, a surgical biopsy is indicated. The definitive approach to a surgical biopsy of the anterior mediastinum is through a median sternotomy. At the time of sternotomy, if the lesion is easily resectable, it should be completely removed. Given the invasiveness of the procedure and the inability in some patients to obtain a definitive diagnosis by frozen section, less invasive procedures are preferable if the lesion is large or if the CT scan or history suggests that surgery is not the best definitive treatment.

Masses in the paratracheal region are easily biopsied by mediastinoscopy. For tumors of the anterior or posterior mediastinum, a left or right VATS approach often allows safe and adequate surgical biopsies. In some patients, an anterior mediastinotomy (i.e., Chamberlain procedure) may be ideal for an anterior tumor or a tumor with significant parasternal extension. Before a surgical biopsy is pursued, a discussion should be held with the pathologist regarding routine histologic assessment, special stains and markers, and requirements for lymphoma workup.

Surgical resection using minimally invasive approaches, including video-assisted and robotic thoracoscopic surgery and transcervical, are now routine for the vast majority of middle and posterior tumors and for moderate sized (<5 to 6 cm) anterior mediastinal tumors. Outcomes comparing VATS to open thymectomy in patients with myasthenia gravis without thymoma were prospectively evaluated by Chang and colleagues in 2005, and no differences were seen in terms of response to therapy and recurrence of symptoms. Pain scores were significantly better in the VATS approach. These reports and others support application of VATS for the majority of anterior mediastinal masses.

Other minimally invasive approaches are under study. For example, good results have been reported using a cervical incision with a sternal retractor for thymus removal. The upward lift allows the surgeon reasonable access to the anterior mediastinum and has proven adequate in some centers for definitive resection of the thymus gland for myasthenia gravis. For larger anterior mediastinal masses or in centers where expertise in thoracoscopy is not available, median sternotomy and thoracotomy remain excellent options for resection of anterior mediastinal masses. Occasionally, a lateral thoracotomy with sternal extension (hemi-clamshell) provides excellent exposure for extensive mediastinal tumors that have a lateral component.

Most surgeons would agree that if a larger anterior mediastinal tumor is seen or malignancy is suspected, a median sternotomy with a more radical resection should be performed.

** Mediastinal Neoplasms**

**Thymic Hyperplasia.** Diffuse thymic hyperplasia was first described in children after successful chemotherapy for lymphoma. It has now been described in adults and is referred to as “rebound thymic hyperplasia.” It is most frequently reported after chemotherapy for lymphoma or germ cell tumors. Initially, atrophy of the thymic gland is seen with subsequent thymic gland enlargement, which can be dramatic. The usual time course for thymic hyperplasia is about 9 months after cessation of chemotherapy (range 2 weeks to 12 months). Benign hyperplasia must be clearly distinguished from recurrent lymphoma or germ cell tumors, which may be difficult since thymic hyperplasia is dramatic in some patients; careful follow-up with serial CT scans is the minimum requirement. The role of PET scanning is unclear. Thymic hyperplasia is a known cause of false-positive PET scans; in many patients, CT scan will show a triangular soft tissue density in the retrosternal space that has a characteristic bilobed anatomic appearance consistent with thymus gland. In addition, a low standardized uptake value of tracer on PET scan suggests a benign tumor. Biopsies may be required if the clinical index of suspicion is high.
**Thymoma.** While it is the most frequently encountered neoplasm of the anterior mediastinum in adults (seen most frequently between 40 and 60 years of age), thymoma is rare in children. Between 10% and 50% of patients with thymoma will have symptoms suggestive of myasthenia gravis or have circulating antibodies to acetylcholine receptor, but less than 10% of patients with myasthenia gravis have a thymoma. Most patients with thymoma are asymptomatic. Thymectomy leads to improvement or resolution of symptoms of myasthenia gravis in only about 25% of patients with thymomas. In contrast, in patients with myasthenia gravis and no thymoma, thymectomy results are superior: up to 50% of patients have a complete remission, and 90% improve. In 5% of patients with thymomas, other paraneoplastic syndromes, including red cell aplasia, hypogammaglobulinemia, systemic lupus erythematosus, Cushing’s syndrome, or SIADH, may be present. Large thymic tumors may present with symptoms related to a mass effect, which may include cough, chest pain, dyspnea, or SVC syndrome.

The diagnosis may be suspected based on CT scan and history, but imaging alone is not diagnostic. In most centers, the diagnosis is made after surgical resection because of the relative difficulty of obtaining a needle biopsy and the likelihood that removal will ultimately be recommended. Biopsy should be avoided in cases where imaging is highly suggestive of thymoma. In most patients, the distinction between lymphomas and thymomas can be made on CT scan since most lymphomas have marked lymphadenopathy and thymomas most frequently appear as a solitary encapsulated mass. PET scan may have a role in differentiating thymic cancer from thymoma, as thymic cancer tends to be very FDG avid.153 In addition, PET scan may facilitate identification of low-risk and minimally invasive thymoma; a standardized uptake value (SUV) <5 was associated with Masaoka stage I or II thymoma, whereas invasive thymoma and mediastinal lymphoma were more likely when the SUV was >5.154 In cases where the diagnosis is unclear, transmediastinal, not transpleural, CT-guided FNA biopsy has a diagnostic sensitivity of 87% and a specificity of 95% in specialized centers.

The most commonly accepted staging system for thymoma is that of Masaoka.155 It is based on the presence or absence of gross or microscopic invasion of the capsule and of surrounding structures, as well as on the presence or absence of metastases (Table 19-30). Histologically, thymomas are characterized by a mixture of epithelial cells and mature lymphocytes. Grossly, many thymomas remain well encapsulated. Even those with capsular invasion often lack histologic features of malignancy; they appear cytologically benign and identical to early-stage tumors. This lack of classic cellular features of malignancy is why most pathologists use the term “thymoma” or “invasive thymoma” rather than “malignant thymoma.” Thymic tumors with malignant cytopathic features are classified separately and referred to as “thymic carcinoma.”

The definitive treatment for thymoma is complete surgical removal; local recurrence rates and survival vary according to stage (Fig. 19-47). In centers with significant experience with VATS procedures, thymoma is not a contraindication to VATS approach, provided the principles of resection are adhered to, such as a complete resection without disrupting the capsule.156 Otherwise, resection is generally accomplished by median sternotomy with extension to hemi-clamshell in more advanced cases. Even advanced tumors with local invasion of resectable structures such as the pericardium, SVC, or innominate vessels should be considered for resection with reconstruction.

A multidisciplinary approach to nonresectable and more advanced lesions (stage ≥II) is mandatory to optimize patient care. The goal for surgical resection should be complete excision of the mass with total thymectomy. All contiguous and noncontiguous disease is removed at the same setting; this may include resection of the pericardium or pleura, adjacent adherent lung, phrenic nerve, major vascular structures, and pleural metastasis. Bilateral phrenic nerve resection should be avoided, however, due to the major respiratory morbidity associated with bilateral paralyzed hemidiaphragms.

The role of adjuvant or neoadjuvant therapies for advanced-stage tumors remains unclear. Traditionally, stage II thymomas have been treated by complete surgical resection followed by mediastinal radiation, but randomized trials have not been done. A recent retrospective review of a single-institution series of stage II thymoma patients showed no difference in survival or local recurrence after complete surgical resection alone, as compared with surgical resection with radiotherapy. Advanced thymomas have been shown to respond to platinum-based chemotherapy and to corticosteroids.157158 One summary of chemotherapy trials showed an overall response rate of about 70%. Cisplatin/doxorubicin-based regimens appear to yield the best results. The combination radiotherapy and chemotherapy for local progression appears to prolong survival in some small series.158 Radiation therapy in surgically resected stage III thymoma is likely beneficial in extending disease-specific survival; an analysis of the Surveillance, Epidemiology, and End Results (SEER) database identified 476 patients with stage III thymoma treated with primary surgery. Postoperative radiation was given

<table>
<thead>
<tr>
<th>Table 19-30</th>
<th>Masaoka staging system for thymoma</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage I</td>
<td>Encapsulated tumor with no gross or microscopic evidence of capsular invasion</td>
</tr>
<tr>
<td>Stage II</td>
<td>Gross capsular invasion or invasion into the mediastinal fat or pleura or microscopic capsular invasion</td>
</tr>
<tr>
<td>Stage III</td>
<td>Gross invasion into the pericardium, great vessels, or lung</td>
</tr>
<tr>
<td>Stage IVA</td>
<td>Pleural or pericardial dissemination</td>
</tr>
<tr>
<td>Stage IVB</td>
<td>Lymphogenous or hematogenous metastasis</td>
</tr>
</tbody>
</table>

Figure 19-47. Stage-specific survival for thymomas.
to 322 patients with a significant improvement in survival (127 months compared to 105 months, \( P = .038 \)) despite the fact that these patients were more likely to have had debulking rather than curative resection. In multivariate analysis, disease-specific survival was better in the adjuvant radiation group.\(^\text{156}\) Therefore, it is imperative that all patients with thymomas undergo a thorough evaluation for potential resection. Current guidelines recommend radiation for patients with unresectable thymoma who have failed induction chemotherapy or for patients with incompletely resected invasive thymoma or thymic cancer. Planning the radiation ports requires input from the surgeon; it is important for the surgeon to carefully document areas of adherence between the thymoma and adjacent structures during the operation, with clips or other radiopaque markers placed to guide radiation therapy postoperatively. Extracapsular extension and positive surgical margins should be noted by the pathologist and correlated anatomically so that the surgeon and radiation oncologist can ensure appropriate radiation treatment.

**Thymic Carcinoma.** Thymic carcinomas are unequivocally malignant at the microscopic level. Suster and Rosai classified thymic carcinomas into low-grade and high-grade tumors.\(^\text{160}\) Low-grade tumors are well differentiated with squamous cell, mucoepidermoid, or basaloid features. High-grade thymic carcinomas include those with lymphoepithelial, small cell neuroendocrine, sarcomatoid, clear cell, and undifferentiated or anaplastic features. Care must be taken to differentiate thymic carcinoma from lung cancer metastatic to the thymus gland as the histologic features can be similar between the two. Compared with thymomas, they are a more heterogeneous group of malignancies with a propensity for early local invasion and widespread metastases. Malignant pleural and pericardial effusions occur frequently.

Five-year survival rates are between 30% and 50%. Complete resection is occasionally curative and leads to improved survival, but most thymic carcinomas will recur and are refractory to chemotherapy.\(^\text{157}\) Management, therefore, depends on the completeness of the resection. Postoperative care includes radiation therapy, guided by residual gross disease or microscopically positive margins from the resection specimen. Chemotherapy may also be given, with carboplatin/paclitaxel recommended based on the best response rates with the least toxicity in clinical trials. The prognosis of patients with thymic cancer remains poor.

**Thymolipoma.** Thymolipomas are rare benign tumors that may grow to a very large size prior to diagnosis. On CT scan, their appearance can be dramatic, with a characteristic fat density dotted by islands of soft tissue density representing islands of thymic tissue (Fig. 19-48). Thymolipomas are generally well-encapsulated, soft, and pliable masses that do not invade surrounding structures. Resection is recommended for large masses.

**Neurogenic Tumors.** Most neurogenic tumors of the mediastinum arise from the cells of the nerve sheath, from ganglion cells, or from the paraganglionic system (Table 19-31). The incidence, cell types, and risk of malignancy strongly correlate with patient age. Tumors of nerve sheath origin predominate in adults. Most present as asymptomatic incidental findings, and most are benign. In children and young adults, tumors of the autonomic ganglia predominate, with up to two-thirds being malignant.\(^\text{161}\)

**Nerve Sheath Tumors.** Nerve sheath tumors account for 20% of all mediastinal tumors. More than 95% of nerve sheath tumors are benign neurilemomas or neurofibromas. Malignant neurosarcomas are much less common.

**Neurilemoma.** Neurilemomas, also called schwannomas, arise from Schwann cells in intercostal nerves. They are firm, well encapsulated, and generally benign. Two characteristic histologic components are referred to as Antoni type A and Antoni type B regions. Antoni type A regions contain compact spindle

---

**Figure 19-48.** Massive thymolipoma that was asymptomatic in an 18-year-old female.
cells with twisted nuclei and nuclear palisading. Antoni type B regions contain loose and myxoid connective tissue with hap hazard cellular arrangement. These characteristics distinguish neurilemoma from malignant fibrosarcomatous tumors, which lack encapsulation and have no Antoni features. If routine CT scan suggests extension of a neurilemoma into the intervertebral foramen, MRI is used to evaluate the extent of this “dumbbell” configuration (Fig. 19-49). Such a configuration may lead to cord compression and paralysis and requires a more complex surgical approach. Resection is recommended; VATS has been established as safe and effective for simple and, in experienced centers, even the more complex operations.\textsuperscript{162} It is reasonable to follow small, asymptomatic paravertebral tumors in older patients or in patients at high risk for surgery. In children, ganglioneuroblastomas or neuroblastomas are more common; therefore, all neurogenic tumors should be completely resected.

**Neurofibroma.** Neurofibromas consist of both nerve sheath and nerve cells and account for up to 25% of nerve sheath tumors. Up to 40% of patients with mediastinal fibromas have generalized neurofibromatosis (von Recklinghausen’s disease). About 70% of neurofibromas are benign, but malignant degeneration to neurofibrosarcoma occurs in 25% to 30% of patients.\textsuperscript{163} The risk of malignant degeneration increases with advancing age, von Recklinghausen’s disease, and exposure to previous radiation. Neurofibrosarcomas carry a poor prognosis because of rapid growth and aggressive local invasion along nerve bundles. Complete surgical resection is the mainstay of treatment. Adjuvant radiotherapy or chemotherapy does not confer a significant benefit, but may be added if complete resection is not possible.\textsuperscript{164} The 5-year survival rate is 53%, but it drops to 16% in patients with neurofibromatosis or with large tumors (>5 cm).

**Ganglion Cell Tumors** Ganglion cell tumors (ganglioneuromas, ganglioneuroblastomas, and neuroblastomas) arise from the sympathetic chain or from the adrenal medulla.

1. **Ganglioneuroma.** Well-differentiated, benign tumors characterized histologically by well-differentiated ganglion cells with a background of Schwann cells, these are most often found incidentally in asymptomatic young adults. Diarrhea related to secretion of a vasoactive intestinal peptide has been described in some patients. These tumors have a propensity for intraspinal canal extension, although they remain well-encapsulated; complete resection is curative, with a low risk of local recurrence.

---

**Table 19-31**

<table>
<thead>
<tr>
<th>TUMOR ORIGIN</th>
<th>BENIGN</th>
<th>MALIGNANT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nerve sheath</td>
<td>Neurilemoma, neurofibroma, melanotic schwannoma, granular cell tumor</td>
<td>Neurofibrosarcoma</td>
</tr>
<tr>
<td>Ganglion cell</td>
<td>Ganglioneuroma</td>
<td>Ganglioneuroblastoma, neuroblastoma</td>
</tr>
<tr>
<td>Paraganglionic cell</td>
<td>Chemodectoma, pheochromocytoma</td>
<td>Malignant chemodectoma, malignant pheochromocytoma</td>
</tr>
</tbody>
</table>

2. **Ganglioneuroblastoma.** Ganglioneuroblastomas contain a mixture of benign ganglion cells and malignant neuroblasts. The distribution of these cells within the tumor is predictive of the clinical course. The nodular pattern has a high incidence of metastatic disease, whereas the diffuse pattern rarely metastasizes. Gross examination typically reveals encapsulated tumor; histologically, there are focal calcifications around regions of neuroblasts. Ganglioneuroblastomas arise most frequently in infants and children <3 years old. The majority are resectable, with 80% 5-year survival.

3. **Neuroblastoma.** Highly malignant, neuroblastomas are the most common extracranial solid malignancy of childhood. The primary site is intrathoracic malignancy in 14%; extension into the spinal canal and osseous invasion commonly present. These thoracic tumors are not as recalcitrant to chemotherapy and surgical resection as other chest malignancies; they are more likely to be resectable, with less invasion of surrounding organs. More than half occur in children under 2 years old; 90% arise within the first decade of life, and thus, these malignancies are discussed in more detail in Chapter 39.

**Paraganglionic Tumors.** Paraganglionic tumors arising in the thoracic cavity include chemodectomas and pheochromocytomas. Only 10% of all pheochromocytomas are located in an extra-adrenal site. Intrathoracic pheochromocytomas are one of the rarest tumors. Approximately 10% of thoracic pheochromocytomas are malignant, a rate similar to that of adrenal tumors. The most common thoracic location is within the costovertebral sulcus, but paraganglionic tumors also arise within the visceral compartment of the mediastinum. These catecholamine-producing lesions can lead to life-threatening hemodynamic problems, so complete removal is important. Diagnosis is generally confirmed by measuring elevated levels of urinary catecholamines and their metabolites. Localization is by CT scan, aided by MIBG scintigraphy. Preoperative care includes α- and β-adrenergic blockade to prevent intraoperative malignant hypertension and arrhythmias. These tumors tend to be highly vascular and should be approached with care. Chemodectomas are rare tumors that may be located around the aortic arch, vagus nerves, or aortic syphymathetics. They rarely secrete catecholamines and are malignant in up to 30% of patients.

**Lymphoma.** Overall, lymphomas are the most common malignancy of the mediastinum. In about 50% of patients who have both Hodgkin’s and non-Hodgkin’s lymphoma, the mediastinum may be the primary site. The anterior compartment is most commonly involved, with occasional involvement of the middle compartment and hilar nodes. The posterior compartment is rarely involved. Chemotherapy and/or radiation results in a cure rate of up to 90% for patients with early-stage Hodgkin’s disease and up to 60% with more advanced stages.

**Mediastinal Germ Cell Tumors.** Germ cell tumors are uncommon neoplasms, but they are the most common malignancy in young men 15 to 35 years of age. Most germ cell tumors are gonadal in origin; primary mediastinal germ cell tumors comprise less than 5% of all germ cell tumors and less than 1% of all mediastinal tumors (usually occurring in the anterior compartment). If a malignant mediastinal germ cell tumor is found, it is important to exclude a gonadal primary tumor. Primary mediastinal germ cell tumors (including teratomas, seminomas, and nonseminomatous malignant germ cell tumors) are a heterogeneous group of benign and malignant neoplasms thought to originate from primitive pluripotent germ cells “misplaced” in the mediastinum during embryonic development. Previously, most mediastinal germ cell tumors were thought to be metastatic. However, two lines of evidence suggest that many mediastinal germ cell tumors are primary, developing from pluripotent primordial germ cells in the mediastinum: (a) several autopsy series showed that patients with extragonadal sites of germ cell tumors, presumed previously to have originated from the gonads, had no evidence of an occult primary tumor or of any residual scar of the gonads, even after an exhaustive search; and (b) patients treated by surgery or radiation for their mediastinal germ cell tumors had long-term survival with no late testicular recurrences.

About one-third of all primary mediastinal germ cell tumors are seminomas. Two-thirds are nonseminomatous tumors or teratomas. Treatment and prognosis vary considerably within these two groups. Mature teratomas are benign and can generally be diagnosed by the characteristic CT findings of multilocular cystic tumors, encapsulated with combinations of fluid, soft tissue, calcium, and/or fat attenuation in the anterior compartment. FNA biopsy alone may be diagnostic for seminomas, usually with normal serum markers, including hCG and AFP. In 10% of seminomas, hCG levels may be slightly elevated. FNA findings, along with high hCG and AFP levels, can accurately diagnose nonseminomatous tumors. If the diagnosis remains uncertain after assessment of FNA findings and serum marker levels, then core-needle biopsies or surgical biopsies may be required. Thoracoscopy is the most frequent diagnostic surgical approach.

1. **Seminoma.** Most patients with seminomas have advanced disease at the time of diagnosis and present with symptoms of local compression, including SVC syndrome, dyspnea, or chest discomfort. With advanced disease, the preferred treatment is combination cisplatin-based chemotherapy regimens with bleomycin and either etoposide or vindesine. Complete responses have been reported in over 75% of patients treated with these regimens. Surgical resection may be curative for small asymptomatic seminomas that are found incidentally with screening CT scans. Surgical resection of residual masses after chemotherapy may be indicated.

2. **Nonseminomatous germ cell tumors.** Nonseminomatous germ cell tumors include embryonal cell carcinomas, choriocarcinomas, endodermal sinus tumors, and mixed types. They are often bulky, irregular tumors of the anterior mediastinum with areas of low attenuation on CT scan because of necrosis, hemorrhage, or cyst formation. Frequently, adjacent structures have been involved, with metastases to regional lymph nodes, pleura, and lungs. Lactate dehydrogenase (LDH), AFP, and hCG levels are frequently elevated. Chemotherapy is the preferred treatment and includes combination therapy with cisplatin, bleomycin, and etoposide, followed by surgical resection of residual disease. With this regimen, survival is 67% at 2 years and 60% at 5 years. Surgical resection of residual masses is indicated, as it may guide further therapy. Up to 20% of residual masses contain additional tumors; in another 40%, mature teratomas; and the remaining 40%, fibrotic tissue.
It is important to note that oxygen toxicity can occur in patients who have been exposed to bleomycin; high levels of oxygen supplementation in the perioperative setting should be avoided in these patients as respiratory failure and death can ensue.\textsuperscript{166} Factors independently predictive of survival after induction chemotherapy followed by resection are elevated serum tumor markers after resection, postchemotherapy pathologic findings (complete necrosis vs. teratoma), and persistent germ cell or non-germ cell cancer in the pathologic specimen.\textsuperscript{166}

3. **Teratoma.** Teratomas are the most common type of mediastinal germ cell tumors, accounting for 60% to 70% of mediastinal germ cell tumors. They contain two or three embryonic layers that may include teeth, skin, and hair (ectodermal), cartilage and bone (mesodermal), or bronchial, intestinal, or pancreatic tissue (endodermal). Therapy for mature, benign teratomas is surgical resection, which confers an excellent prognosis. Rarely, teratomas may contain a focus of carcinoma; these malignant teratomas (or teratocarcinomas) are locally aggressive. Often diagnosed at an unresectable stage, they respond poorly to chemotherapy and in a limited manner to radiotherapy; prognosis is uniformly poor.

**Mediastinal Cysts**

Benign cysts account for up to 25% of mediastinal masses and are the most frequently occurring mass in the middle mediastinal compartment. A CT scan showing characteristic features of near water density in a typical location is virtually 100% diagnostic.\textsuperscript{167}

1. **Pericardial cyst.** Usually asymptomatic and detected incidentally in the right costophrenic angle, pericardial cysts typically contain a clear fluid and are lined with a single layer of mesothelial cells. For most simple, asymptomatic pericardial cysts, observation alone is recommended. Surgical resection or aspiration may be indicated for complex cysts or large symptomatic cysts.

2. **Bronchogenic cyst.** Developmental anomalies that occur during embryogenesis and occur as an abnormal budding of the foregut or tracheobronchial tree, bronchogenic cysts arise most often in the mediastinum just posterior to the carina or main stem bronchus. Approximately 15% occur within the pulmonary parenchyma. Thin-walled and lined with respiratory epithelium, they contain a protein-rich mucoid material and varying amounts of seromucous glands, smooth muscle, and cartilage. They may communicate with the tracheobronchial tree. In adults, over half of all bronchogenic cysts are found incidentally during workup for an unrelated problem or during screening. The natural history of an incidentally diagnosed, asymptomatic bronchogenic cyst is unknown, but it is clear that many such cysts do not lead to clinical problems. In one study of young military personnel, 78% of all bronchogenic cysts found on routine CXRs were asymptomatic. However, in other reports with more comprehensive follow-up, up to 67% of adults with incidentally found bronchogenic cysts eventually became symptomatic. Symptoms include chest pain, cough, dyspnea, and fever. If large (>6 cm) or symptomatic, resection is generally recommended since serious complications may occur if the cyst becomes larger or infected.

Complications include airway obstruction, infection, rupture, and, rarely, malignant transformation.\textsuperscript{168,169}

Traditionally, complete removal of the cyst wall has been via posterolateral thoracotomy.\textsuperscript{170} Resection of infected cysts may be quite difficult because of dense adhesions; elective removal is often recommended before infection has a chance to occur. Thoracoscopic exploration and resection are possible for small cysts with minimal adhesions. With increasing experience using video-assisted or robotic-assisted thoracoscopy, a greater proportion of these lesions are amenable to minimally invasive resection.

3. **Enteric cyst.** Generally asymptomatic, thymic cysts are often discovered incidentally. Simple cysts are of no consequence; however, the occasional cystic neoplasm must be ruled out. Cystic components occasionally are seen in patients with thymoma and Hodgkin’s disease.

4. **Thymic cyst.** Generally asymptomatic, thymic cysts are often discovered incidentally. Simple cysts are of no consequence; however, the occasional cystic neoplasm must be ruled out. Cystic components occasionally are seen in patients with thymoma and Hodgkin’s disease.

5. **Ectopic endocrine glands.** Up to 5% of all mediastinal masses are of thyroid origin; most are simple extensions of thyroid masses. Usually nontoxic, over 95% can be completely resected through a cervical approach. True ectopic thyroid tissue of the mediastinum is rare. About 10% to 20% of abnormal parathyroid glands are found in the mediastinum; most can be removed during exploration from a cervical incision. In cases of true mediastinal parathyroid glands, thoracoscopic or open resection may be indicated. Location can generally be pinpointed by a combination of CT scan and Sestamibi scans.

**Mediastinitis**

**Acute Mediastinitis.** Acute mediastinitis is a fulminating infectious process that spreads rapidly along the continuous fascial planes connecting the cervical and mediastinal compartments. Infections originate most commonly from esophageal perforations, sternal infections, and oropharyngeal or neck infections, but a number of less common etiologic factors can lead to this deadly process (Table 19-32). Clinical signs and symptoms include fever, chest pain, dysphagia, respiratory distress, and cervical and upper thoracic subcutaneous crepitus. In severe cases, the clinical course can rapidly deteriorate to florid sepsis, hemodynamic instability, and death. Thus, a high index of suspicion is required in the context of any infection with access to the mediastinal compartments.

A chest CT scan illuminates the extent of spread and guides selection of the best approach to surgical drainage. Acute mediastinitis is a true surgical emergency; treatment must be instituted immediately and aimed at correcting the primary problem, such as the esophageal perforation or oropharyngeal abscess, and debridement and drainage of the spreading infectious process within the mediastinum, neck, pleura, and other tissue planes. Antibiotics, fluid resuscitation, and other supportive measures are also important.
Debridement may need to be repeated and other planes and cavities explored depending on the patient’s clinical status. Blood cell counts and serial CT scans may also be required. Persistent sepsis or collections on CT scan may require further radical surgical debridement.

**Chronic Mediastinitis.** Sclerosing or fibrosing mediastinitis results from chronic mediastinal inflammation that originates in the lymph nodes, most frequently from granulomatous infections such as histoplasmosis or tuberculosis. Chronic, low-grade inflammation leads to fibrosis and scarring, which can, in some patients, result in entrapment and compression of the low-pressure veins (including the SVC and innominate and azygos veins), the esophagus, and pulmonary arteries. There is no definitive treatment. Surgery is indicated only for diagnosis or in specific patients to relieve airway or esophageal obstruction or to achieve vascular reconstruction. Reports of palliative success with less invasive procedures (such as dilation and stenting of airways, the esophagus, or the SVC) are promising. In one series of 22 patients, ketoconazole was effective in controlling progression. In another series of 71 patients, 30% died from disease-associated complications during long-term follow-up. Chronic mediastinitis is similar to retroperitoneal fibrosis, sclerosing cholangitis, and Riedel’s thyroiditis.

**Table 19-32**

<table>
<thead>
<tr>
<th>Etiologic factors in acute mediastinitis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Esophageal perforation</strong></td>
</tr>
<tr>
<td>Iatrogenic</td>
</tr>
<tr>
<td>Balloon dilatation (for achalasia)</td>
</tr>
<tr>
<td>Bougienage (for peptic stricture)</td>
</tr>
<tr>
<td>Esophagoscopy</td>
</tr>
<tr>
<td>Sclerotherapy (for variceal bleeding)</td>
</tr>
<tr>
<td><strong>Spontaneous</strong></td>
</tr>
<tr>
<td>Postemetic (Boehrhaave’s syndrome)</td>
</tr>
<tr>
<td><strong>Straining during:</strong></td>
</tr>
<tr>
<td>Elimination</td>
</tr>
<tr>
<td>Weight lifting</td>
</tr>
<tr>
<td>Seizure</td>
</tr>
<tr>
<td>Pregnancy</td>
</tr>
<tr>
<td>Childbirth</td>
</tr>
<tr>
<td>Ingestion of foreign bodies</td>
</tr>
<tr>
<td>Trauma</td>
</tr>
<tr>
<td>Blunt</td>
</tr>
<tr>
<td>Penetrating</td>
</tr>
<tr>
<td>Postsurgical</td>
</tr>
<tr>
<td>Infection</td>
</tr>
<tr>
<td>Anastomotic leak</td>
</tr>
<tr>
<td>Erosion by cancer</td>
</tr>
<tr>
<td><strong>Deep sternotomy wound infection</strong></td>
</tr>
<tr>
<td><strong>Oropharynx and neck infections</strong></td>
</tr>
<tr>
<td>Ludwig’s angina</td>
</tr>
<tr>
<td>Quinsy</td>
</tr>
<tr>
<td><strong>Retropharyngeal abscess</strong></td>
</tr>
<tr>
<td><strong>Cellulitis and supplicative lymphadenitis of the neck</strong></td>
</tr>
<tr>
<td><strong>Infections of the lung and pleura</strong></td>
</tr>
<tr>
<td><strong>Subphrenic abscess</strong></td>
</tr>
<tr>
<td><strong>Rib or vertebral osteomyelitis</strong></td>
</tr>
<tr>
<td><strong>Hematogenous or metastatic abscess</strong></td>
</tr>
</tbody>
</table>


**PLEURA AND PLEURAL SPACE**

**Anatomy**

Each hemithorax has a mesothelial lining that invaginates at the hilum of each lung and continues on to cover each lung. The portion lining the bony rib cage, mediastinum, and diaphragm is called the parietal pleura, whereas the portion encasing the lung is known as the visceral pleura. Between these two surfaces is the potential pleural space, which is normally occupied by a thin layer of lubricating pleural fluid. A network of somatic, sympathetic, and parasympathetic fibers innervates the parietal pleura. Irritation of the parietal surface by inflammation, tumor invasion, trauma, and other processes can lead to a sensation of chest wall pain. The visceral pleura have no somatic innervation.

**Pleural Effusion**

Pleural effusion refers to any significant collection of fluid within the pleural space. Normally, between 5 and 10 L of fluid enters the pleural space each day by filtration through microvessels supplying the parietal pleura (located mainly in the less dependent regions of the cavity). The net balance of pressures in these capillaries leads to fluid flow from the parietal pleural surface into the pleural space, and the net balance of forces in the pulmonary circulation leads to absorption through the visceral pleura. Normally, 15 to 20 mL of pleural fluid is present at any given time. Any disturbance in these forces can lead to imbalance and accumulation of pleural fluid. Common pathologic conditions in North America that lead to pleural effusion include congestive heart failure, bacterial pneumonia, malignancy, and pulmonary emboli (Table 19-33).

**Access and Drainage of Pleural Fluid Collections**

Most patients with pleural effusions of unknown cause should undergo thoracentesis with the following exceptions: effusions in the setting of congestive heart failure, hepatic failure or renal failure, or small effusions associated with an improving pneumonia. If the clinical history suggests congestive heart failure as a cause, particularly in the setting of bilateral effusions, a trial of diuresis may be indicated (rather than thoracentesis). Up to 75% of effusions due to congestive heart failure resolve within 48 hours with diuresis alone. Similarly, thoracentesis can be avoided in patients with small effusions associated with resolving pneumonia. These patients typically present with cough, fever, leukocytosis, and unilateral infiltrate, and the effusion is usually a result of a reactive, parapneumonic process. If the effusion is small and the patient responds to antibiotics, a diagnostic thoracentesis may be unnecessary. If the effusion is large and compromising respiratory efforts, or if the patient has a persistent white blood cell count despite improving signs of pneumonia, an empyema of the pleural space must be considered. In these patients, early and aggressive drainage with chest tubes is required, possibly with surgical intervention.

Once the decision is made to access a pleural effusion, the next step is to determine if a sample of the fluid or complete drainage of the pleural space is desired. This step is influenced by the clinical history, the type and amount of fluid present, the
nature of the collection (such as free-flowing or loculated), the cause, and the likelihood of recurrence. For small, free-flowing effusions, an outpatient diagnostic and/or therapeutic thoracentesis with a relatively small-bore needle or catheter (14- to 16-gauge) can be performed (Fig. 19-50). The appearance of the fluid is informative: clear straw-colored fluid is often transudative; turbid or bloody fluid is often exudative.

The site of entry for drainage of a pleural effusion or pneumothorax may be based on the CXR alone if the effusion is demonstrated to be free-flowing. For free-flowing effusions, a low approach at the eighth or ninth intercostal space in the posterior midclavicular line facilitates complete drainage. If the effusion is loculated, CT- or ultrasound-guided drainage may be indicated. If the goal is complete drainage, a small-bore pigtail catheter is inserted and connected to a closed drainage system with applied suction (typically \(-20 \text{ cm } \text{H}_2\text{O}\)). In general, the smallest-bore drainage catheter that will effectively drain the pleural space should be chosen. Smaller-diameter catheters significantly decrease the pain associated with the placement of chest tubes but are more prone to clogging and twisting.\textsuperscript{174,175} For clinical situations requiring biopsy or for potential interventions such as adhesiolysis or pleurodesis, minimally invasive surgery may be indicated, using a VATS approach.

Table 19-33

<table>
<thead>
<tr>
<th>CAUSE</th>
<th>ANNUAL INCIDENCE</th>
<th>TRANSUDATE</th>
<th>EXUDATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congestive heart failure</td>
<td>500,000</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>300,000</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Cancer</td>
<td>200,000</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Pulmonary embolus</td>
<td>150,000</td>
<td>Sometimes</td>
<td>Sometimes</td>
</tr>
<tr>
<td>Viral disease</td>
<td>100,000</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Coronary artery bypass surgery</td>
<td>60,000</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Cirrhosis with ascites</td>
<td>50,000</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

Complications of Pleural Drainage. The most common complications of invasive pleural procedures are inadvertent injury to adjacent organs, including lung, with air leakage and pneumothorax; subdiaphragmatic entry and damage to the liver, spleen, or other intra-abdominal viscera; intercostal vessel injury with subsequent bleeding or larger vessel injury; and even cardiac puncture. Sometimes, bleeding may be the result of an underlying coagulopathy or anticoagulant therapy. Other technical complications include loss of a catheter, guidewire, or fragment in the pleural space and infections. Occasionally, rapid drainage of a large effusion can be followed by shortness of breath, clinical instability, and a phenomenon referred to as postexpansion pulmonary edema. For this reason, it is recommended to drain only up to 1500 mL initially. Most complications can be avoided by consulting with a clinician experienced in pleural drainage techniques.

Pleural Fluid Analysis. Pleural fluid collections are generally classified as transudates and exudates (Table 19-34). Transudates are protein-poor ultrafiltrates of plasma that result from alterations in the systemic hydrostatic pressures or colloid...
osmotic pressures (for example, with congestive heart failure or cirrhosis). On gross visual inspection, a transudative effusion is generally clear or straw-colored. Exudates are protein-rich pleural fluid collections that generally result from inflammation or pleural invasion by tumor. Grossly, they are often turbid, bloody, or purulent. Absent trauma, grossly bloody effusions are frequently malignant, but they may also occur in the setting of a pulmonary embolism or pneumonia.

Transudates and exudates can be differentiated using Light’s criteria. An effusion is exudative if the pleural fluid-to-serum ratio of protein is greater than 0.5 and the LDH ratio is greater than 0.6 or the absolute pleural LDH level is greater than two-thirds of the normal upper limit for serum. If criteria suggest a transudate, a careful evaluation for congestive heart failure, cirrhosis, or conditions associated with transudates is undertaken. If criteria suggest an exudate, further diagnostic studies may be helpful. If total and differential cell counts reveal a predominance of neutrophils (>50% of cells), the effusion is likely associated with an acute inflammatory process (such as a parapneumonic effusion or empyema, pulmonary embolus, or pancreatitis). A predominance of mononuclear cells suggests a more chronic inflammatory process (such as cancer or tuberculosis). Gram stains and cultures should be obtained if possible, with inoculation into culture bottles at the bedside. Pleural fluid glucose levels are frequently decreased (<60 mg/dL) with complex parapneumonic effusions or malignant effusions. It is important to note that while the distinction between transudate and exudate can be diagnostically useful, the ultimate decision for prolonged chest tube drainage or surgery depends on the effusion size and adequacy of drainage, presence of loculations, adequacy of lung reexpansion after drainage, and recurrence after initial drainage.

A pleural effusion occurring in association with pleuritic chest pain, hemoptysis, or dyspnea out of proportion to the size of the effusion should raise concern for pulmonary embolism. These effusions may be transudative, but if an associated infarct near the pleural surface occurs, an exudate may be seen. If a pulmonary embolism is suspected in a postoperative patient, most clinicians would obtain a spiral CT scan. Alternatively, duplex ultrasonography of the lower extremities may yield a diagnosis of deep vein thrombosis, thereby indicating anticoagulant therapy and precluding the need for a specific diagnosis of pulmonary embolism. In some patients, a blood test for levels of D-dimer may be helpful; if a sensitive D-dimer blood test is negative, pulmonary embolism may be ruled out.

### Malignant Pleural Effusion

Malignant pleural effusions may occur in association with a number of different malignancies, most commonly lung cancer, breast cancer, and lymphomas, depending on the patient’s age and gender (Tables 19-35 and 19-36). Cytologic testing should be done on exudative effusions to rule out an associated malignancy; accuracy is 70% when associated with adenocarcinomas, but it is less sensitive for mesotheliomas (≤10%), squamous cell carcinomas (20%), or lymphomas (25% to 50%). If the diagnosis remains uncertain after drainage and fluid analysis, thoracoscopy and direct biopsies are indicated. Malignant effusions are exudative and often tinged with blood. An effusion in the setting of a malignancy means a more advanced stage; mean survival ranges from 3 to 11 months, depending on the primary tumor location.

Occasionally, effusions associated with a bronchogenic NSCLC are benign, and surgical resection may still be indicated.

#### Table 19-35

<table>
<thead>
<tr>
<th>PRIMARY SITE OR TUMOR TYPE</th>
<th>NO. OF MALE PATIENTS</th>
<th>PERCENTAGE OF MALE PATIENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung</td>
<td>140</td>
<td>49.1</td>
</tr>
<tr>
<td>Lymphoma/leukemia</td>
<td>60</td>
<td>21.1</td>
</tr>
<tr>
<td>Gastrointestinal tract</td>
<td>20</td>
<td>7.0</td>
</tr>
<tr>
<td>Genitourinary tract</td>
<td>17</td>
<td>6.0</td>
</tr>
<tr>
<td>Melanoma</td>
<td>4</td>
<td>1.4</td>
</tr>
<tr>
<td>Miscellaneous less common tumors</td>
<td>10</td>
<td>3.5</td>
</tr>
<tr>
<td>Primary site unknown</td>
<td>31</td>
<td>10.9</td>
</tr>
<tr>
<td>Total</td>
<td>285</td>
<td>100.0</td>
</tr>
</tbody>
</table>


#### Table 19-36

<table>
<thead>
<tr>
<th>PRIMARY SITE OR TUMOR TYPE</th>
<th>NO. OF FEMALE PATIENTS</th>
<th>PERCENTAGE OF FEMALE PATIENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast</td>
<td>70</td>
<td>37.4</td>
</tr>
<tr>
<td>Female genital tract</td>
<td>38</td>
<td>20.3</td>
</tr>
<tr>
<td>Lung</td>
<td>28</td>
<td>15.0</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>14</td>
<td>8.0</td>
</tr>
<tr>
<td>Gastrointestinal tract</td>
<td>8</td>
<td>4.3</td>
</tr>
<tr>
<td>Melanoma</td>
<td>6</td>
<td>3.2</td>
</tr>
<tr>
<td>Urinary tract</td>
<td>2</td>
<td>1.1</td>
</tr>
<tr>
<td>Miscellaneous less common tumors</td>
<td>3</td>
<td>1.6</td>
</tr>
<tr>
<td>Primary site unknown</td>
<td>17</td>
<td>9.1</td>
</tr>
<tr>
<td>Total</td>
<td>187</td>
<td>100.0</td>
</tr>
</tbody>
</table>


Effusion size and degree of associated dyspnea influence management. Symptomatic, moderate to large effusions should be drained by tunneled indwelling pleural catheter, tube thoracostomy (chest tube or pigtail catheter) with subsequent instillation of a sclerosing agent, or VATS with talc instillation. Management is based on patient preference, degree of known or anticipated lung reexpansion, and patient tolerance for operative intervention. Lung entrapment by tumor or adhesions limits reexpansion and generally predicts a poor result with pleurodesis; it is the primary indication for placement of indwelling pleural catheters. Patient preference is also considered, as is their life expectancy. Tunneled indwelling pleural catheters have dramatically changed the management of end-stage cancer treatment.
because they substantially shorten the amount of time patients spend in the hospital during their final weeks of life. If the lung is expected to fully expand and the patient has a longer life expectancy (e.g., malignant effusions in the setting of breast cancer), drainage with sclerosis is the preferred option. The choice of sclerosant includes mechanical pleurodesis or pleurectomy, talc, bleomycin, or doxycycline. Success rates range from 60% to 90% and are highest with talc. Typically, talc is administered as an aerosolized powder during video-assisted thoracoscopy, whereas doxycycline or a talc slurry is infused at the bedside through a previously placed pigtail catheter or larger bore chest tube. Figure 19-51 presents a decision algorithm for the management of malignant pleural effusion.

Empyema
Thoracic empyema is defined by a purulent pleural effusion. Patients of all ages can develop empyema, but the frequency is increased in older or debilitated patients. Common associated conditions include a pneumonic process in patients with pulmonary disorders and neoplasms, cardiac problems, diabetes mellitus, drug and alcohol abuse, neurologic impairments, postthoracotomy problems, and immunologic impairments. The mortality of empyema frequently depends on the degree of severity of the comorbidity; it may range from as low as 1% to over 40% in immunocompromised patients.

Pathophysiology. The most common causes are parapneumonic, but postsurgical, posttraumatic, and GI-associated (e.g., subphrenic or hepatic abscess, perforation of esophagus or other viscus) empyema is also common (Table 19-37). The spectrum of organisms involved in pneumonic processes that progress to empyema is changing. Pneumococci and staphylococci continue to be the most common, but gram-negative aerobic bacteria and anaerobes are becoming more prevalent. Cases involving mycobacteria or fungi are rare. Multiple organisms may be found in up to 50% of patients. Cultures may be sterile, however, if antibiotics were initiated before the culture or if the culture process was not efficient. It is also fairly common for Pneumococcus to grow in blood cultures but not to grow in the pleural fluid cultures. The choice of antibiotics, therefore, is guided by the clinical scenario and not just the organisms found on culture. Broad-spectrum coverage may be required even when cultures do not grow out an organism or if a single organism is grown when the clinical picture is more consistent with a multiorganism process. For example, a polymicrobial gram stain, particularly including yeast, is strongly suggestive of esophageal perforation. Common gram-negative organisms include Escherichia coli, Klebsiella, Pseudomonas, and Enterobacteriaceae. Anaerobic organisms may be fastidious and difficult to document by culture and are associated with periodontal diseases (especially Streptococcus species), aspiration syndromes, alcoholism, general anesthesia, drug abuse, or other functional associations with gastroesophageal reflux.

Organisms gain entry into the pleural cavity through contiguous spread from pneumonia, lung abscess, liver abscess, or another, adjacent infectious processes. Organisms may also enter the pleural cavity by direct contamination from thoracentesis, thoracic surgical procedures, esophageal injuries,
or trauma. As organisms enter the pleural space, an influx of polymorphonuclear cells and fluid occurs, with subsequent release of inflammatory mediators and toxic oxygen radicals. These mechanisms lead to variable degrees of endothelial injury and capillary instability. This process overwhelms the normal pleural lymphatic drainage. This early effusion is watery and free-flowing in the pleural cavity. Thoracentesis at this stage yields fluid with a pH typically above 7.3, a glucose level greater than 60 mg/dL, and a low LDH level (<500 U/L). At this stage, the decision to use antibiotics alone or perform a repeat thoracentesis, chest tube drainage, thoracoscopy, or open thoracotomy depends on the amount of pleural fluid, its consistency, the clinical status of the patient, the degree of expansion of the lung after drainage, and the presence of loculated fluid in the pleural space (vs. free-flowing purulent fluid). Early in the parapneumonic process, when the purulent fluid is relatively thin, complete drainage with simple large-bore thoracentesis is possible. If complete lung expansion is obtained and the pneumatic process is responding to antibiotics, no further drainage may be necessary. Pleural fluid with a pH lower than 7.2 and with a glucose level of less than 40 mg/dL means that a more aggressive approach to drainage should be pursued.

The pleural fluid may become thick and loculated over the course of hours to days and may be associated with fibrinous adhesions (the fibrinopurulent stage). At this stage, chest tube insertion with closed-system drainage or drainage with thoracoscopy may be necessary to remove the fluid and adhesions and facilitate complete lung expansion. Further progression of the inflammatory process leads to the formation of a pleural peel, which may be flimsy and easy to remove early on. However, as the process progresses, a thick pleural rind may develop, leaving a trapped lung; complete lung decortication by either thoracoscopic or thoracotomy would then be necessary.

The use of intrapleural fibrinolytic therapy for management of empyema has been investigated in several large prospective trials. Intrapleural infusion of tissue plasminogen activator (t-PA) alone did not improve outcomes, whereas combined intrapleural t-PA and DNase was associated with a reduction in hospital stay of nearly 7 days, 77% fewer referrals for surgical intervention at 3 months, and more than double the reduction in the infected pleural fluid collection by CXR imaging. In this trial, the medications were given twice daily by intrapleural injection; the dose was 5 mg for the DNase and 10 mg for t-PA. The chest drain was clamped for 1 hour after injection and released. This study suggests that the combination of fibrinolysis (t-PA) and cleavage of uncoiled DNA by DNase reduces fluid viscosity and facilitates pleural clearance.

Management. If there is a residual space, persistent pleural infection is likely to occur. A persistent pleural space may be secondary to contracted, but intact, underlying lung; or it may be secondary to surgical lung resection. If the space is small and well drained by a chest tube, a conservative approach may be possible. This requires leaving the chest tubes in place and attached to closed-system drainage until symphysis of the visceral and parietal surfaces takes place. At this point, the chest tubes can be removed from suction; if the residual pleural space remains stable, the tubes can be cut and advanced out of the chest over the course of several weeks. If the patient is stable, tube removal can frequently be done in the outpatient setting, guided by the degree of drainage and the size of the residual space visualized on serial CT scans. Larger spaces may require open thoracotomy and decortication in an attempt to reexpand the lung to fill this residual space. If reexpansion has failed or appears too high risk, then open drainage, rib resection, and prolonged packing may be required, with delayed closure with muscle flaps or thoracoplasty. Most chronic pleural space problems can be avoided by early specialized thoracic surgical consultation and complete drainage of empyema, allowing space obliteration by the reinflected lung.

**Chylothorax**

Chylothorax develops most commonly after surgical trauma to the thoracic duct or a major branch, but may be also associated with a number of other conditions (Table 19-38). It is generally unilateral; for example, it may occur after dissection of the distal esophagus where the duct lies in close proximity to the esophagus as it enters the right chest from its origin in the

### Table 19-38 Etiology of chylothorax

<table>
<thead>
<tr>
<th>Congenital</th>
<th>Traumatic and/or iatrogenic</th>
<th>Surgery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atresia of thoracic duct</td>
<td>Blunt injury</td>
<td>Cervical</td>
</tr>
<tr>
<td>Thoracic duct-pleural space fistula</td>
<td>Penetrating injury</td>
<td>Excision of lymph nodes</td>
</tr>
<tr>
<td>Birth trauma</td>
<td>Surgery</td>
<td>Radical neck dissection</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Thoracic</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Correction of patent ductus arteriosus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Correction of coarctation of the aorta</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vascular procedure involving the origin of the left subclavian artery</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Esophagectomy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sympathectomy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Resection of thoracic aneurysm</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Resection of mediastinal tumors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Left pneumonectomy</td>
</tr>
<tr>
<td>Abdominal</td>
<td></td>
<td>Abdominal</td>
</tr>
<tr>
<td>Sympathectomy</td>
<td></td>
<td>Sympathectomy</td>
</tr>
<tr>
<td>Radical lymph node dissection</td>
<td></td>
<td>Radical lymph node dissection</td>
</tr>
<tr>
<td>Diagnostic procedures</td>
<td></td>
<td>Diagnostic procedures</td>
</tr>
<tr>
<td>Translumbar arteriography</td>
<td></td>
<td>Translumbar arteriography</td>
</tr>
<tr>
<td>Subclavian vein catheterization</td>
<td></td>
<td>Subclavian vein catheterization</td>
</tr>
<tr>
<td>Left-sided heart catheterization</td>
<td></td>
<td>Left-sided heart catheterization</td>
</tr>
</tbody>
</table>

**Neoplasms**

**Infections**

- Tuberculous lymphadenitis
- Nonspecific mediastinitis
- Ascending lymphangitis
- Filarisis

**Miscellaneous**

- Venous thrombosis
- Left subclavian-jugular vein
- Superior vena cava
- Pulmonary lymphangiomatosis

abdomen at the cisterna chyli (Fig. 19-52). If the mediastinal pleura are disrupted on both sides, bilateral chylothoraces may occur. Left-sided chylothoraces may develop after a left-sided neck dissection, especially in the region of the confluence of the subclavian and internal jugular veins. Chylothorax may also follow nonsurgical trauma, including penetrating or blunt injuries to the chest or neck area, central line placements, and other surgical misadventures. It may be seen in association with a variety of benign and malignant diseases that generally involve the lymphatic system of the mediastinum or neck.

Pathophysiology. Most commonly, the thoracic duct originates in the abdomen from the cisterna chyli, which is located in the midline, near the level of the second lumbar vertebra. From this origin, the thoracic duct ascends into the chest through the aortic hiatus at the level of T10 to T12, and courses just to the right of the aorta (see Fig. 19-52). As the thoracic duct courses cephalad above the diaphragm, it most commonly remains in the right chest, lying just behind the esophagus, between the aorta and aygos vein. The duct continues superiorly, lying just to the right of the vertebral column. Then, at the fourth thoracic vertebra, it crosses behind the aorta and the aortic arch into the left posterior mediastinum and travels superiorly, staying near the esophagus and mediastinal pleura as it exits the thoracic inlet. As it exits the thoracic inlet, it passes to the left, just behind the carotid sheath and anterior to the inferior thyroid and vertebral bodies. Just medial to the anterior scalene muscle, it courses inferiorly and drains into the union of the internal jugular and subclavian veins. Given the extreme variability in the main duct and its branches, accumulation of chyle in the chest or flow from penetrating wounds may be seen after a variety of traumatic and medical conditions.

The main function of the duct is to transport fat absorbed from the digestive system along with variable amounts of protein and lymphatic material (Table 19-39). Given the high volume of chyle that flows through the thoracic duct, significant injuries can cause leaks in excess of 2 L per day; if left untreated, protein, lymphocyte, and volume depletion can lead to serious metabolic effects and death. Thoracentesis is usually grossly suggestive, revealing milky, nonpurulent pleural fluid. However, if the patient is taking nothing by mouth, the pleural fluid may not be grossly abnormal. Laboratory analysis of the pleural fluid shows the presence of chylomicrons, a high lymphocyte count and high triglyceride levels. If the triglyceride level is greater than 110 mg/100 mL, a chylothorax is almost certainly present (a 99% accuracy rate). If the triglyceride level is less than 50 mg/mL, there is only a 5% chance of chylothorax. In many clinical situations, the accumulation of chyle may be slow because of minimal digestive fat flowing through the gastrointestinal tract after major trauma or surgery, so the diagnosis may be more difficult to establish.

Management. The treatment plan for any chylothorax depends on its cause, the amount of drainage, and the patient’s clinical status (Fig. 19-53). In general, most patients are treated with a short period of chest tube drainage, nothing by mouth (NPO) orders, total parenteral nutrition (TPN), and observation. In centers with interventional radiology expertise, thoracic duct embolization as soon as possible after diagnosis should be considered.

![Thoracic Duct Diagram](https://example.com/ThoracicDuctDiagram.png)

**Figure 19-52.** Normal thoracic duct anatomy. The esophagus comes into close proximity with the thoracic duct as it enters the chest from its origin in the abdomen at the cisterna chyli.

---

<table>
<thead>
<tr>
<th>COMPONENT</th>
<th>AMOUNT (PER 100 ML)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total fat</td>
<td>0.4–5 g</td>
</tr>
<tr>
<td>Total cholesterol</td>
<td>65–220 mg</td>
</tr>
<tr>
<td>Total protein</td>
<td>2.21–5.9 g</td>
</tr>
<tr>
<td>Albumin</td>
<td>1.1–4.1 g</td>
</tr>
<tr>
<td>Globulin</td>
<td>1.1–3.1 g</td>
</tr>
<tr>
<td>Fibrinogen</td>
<td>16–24 g</td>
</tr>
<tr>
<td>Sugars</td>
<td>48–200 g</td>
</tr>
<tr>
<td>Electrolytes</td>
<td>Similar to levels in plasma</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cellular elements</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lymphocytes</td>
<td>400–6800/mm³</td>
</tr>
<tr>
<td>Erythrocytes</td>
<td>50–600/mm³</td>
</tr>
<tr>
<td>Antithrombin globulin</td>
<td>≥25% of plasma concentration</td>
</tr>
<tr>
<td>Prothrombin</td>
<td>≥25% of plasma concentration</td>
</tr>
<tr>
<td>Fibrinogen</td>
<td>≥25% of plasma concentration</td>
</tr>
</tbody>
</table>

Chest cavity drainage must be adequate to allow complete lung reexpansion. Somatostatin has been advocated by some authors, with variable results. If significant chyle drainage (>500 mL per day in an adult, >100 mL in an infant) continues despite TPN and good lung expansion, early surgical duct ligation or embolization is recommended (within 4–7 days following diagnosis). Ligation can be approached best by right thoracotomy, and in some experienced centers, by right thoracoscopy. Chylothoraces due to malignant conditions often respond to radiation and/or chemotherapy and less commonly require surgical ligation. Significant nutritional and immunologic depletion results from untreated chylothorax; associated mortality is in excess of 50%. With early recognition and aggressive medical management as well as early surgical ligation or embolization for persistent leaks, the mortality rate of chylothorax is now less than 10%.

**Tumors of the Pleura**

Malignant mesothelioma is the most common type of primary tumor of the pleura, with approximately 3000 cases per year in the United States. Other, less common tumors include benign and malignant fibrous tumors of the pleura, lipomas, and cysts.

**Malignant Mesothelioma.** The only known risk factor for mesothelioma is exposure to asbestos, identified in over 50% of cases. Exposure is typically work-related in industries using asbestos in the manufacturing process, such as shipbuilding and brake pad linings. The risk extends to family members who are
exposed to the dust of the clothing or to the work environment. Asbestos exposure and smoking synergistically increase the risk for lung cancer, but smoking does not increase risk for malignant mesotheliomas. Male predominance is 2:1, and it occurs most commonly after the age of 40.

Risk of developing mesothelioma after asbestos exposure differs depending on the physical characteristics of the asbestos and similar fibers (either serpentine or amphibole). The serpentine fibers are large and curly and are generally not able to travel beyond larger airways. However, the narrow, straight amphibole fibers, in particular the crocidolite fibers, may navigate distally into the pulmonary parenchyma and are most clearly associated with mesotheliomas. Like many carcinogens, the latency period between asbestos exposure and the development of mesothelioma is at least 20 years. The tumor generally is multicentric, with multiple pleural-based nodules coalescing to form sheets of tumor. This process initially involves the parietal pleura, generally with early spread to the visceral surfaces and with a variable degree of invasion of surrounding structures. Autopsy studies have shown that most patients have distant metastases, but the natural history of the disease in untreated patients culminates in death due to local extension and effective stranulation of the lung.

Clinical Presentation Most patients present with dyspnea and chest pain. Over 90% have a pleural effusion, but thoracentesis is diagnostic in less than 10% of patients. Frequently, a thoracoscopy or open pleural biopsy with special stains is required to differentiate mesotheliomas from adenocarcinomas (Table 19-40). Epithelial subtypes are associated with a more favorable prognosis, and long-term survival may be seen in rare patients with no treatment. Sarcomatous and mixed tumors share a more aggressive course.

Management The treatment of malignant mesotheliomas remains controversial. Treatment options include supportive care only, surgical resection, and multimodality approaches (using a combination of surgery, chemotherapy, and radiation therapy). Surgical approaches range from extrapleural pneumonectomy (removal of the lung, entire parietal pleural, ipsilateral pericardium, and hemidiaphragm with patch reconstruction), pleurectomy and decortication (removal of the visceral and parietal pleura only), and palliative procedures such as talc pleurodesis and insertion of long-term tunneled indwelling pleural catheters.

### Table 19-40 Differentiation of mesothelioma from adenocarcinoma

<table>
<thead>
<tr>
<th>Immunohistochemical results</th>
<th>MESOTHELIOMA</th>
<th>ADENOCARCINOMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carcinoembryonic antigen</td>
<td>Negative</td>
<td>Positive</td>
</tr>
<tr>
<td>Vimentin</td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>Low molecular weight cytokeratins</td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>Electron microscopic features</td>
<td>Long, sinuous villi</td>
<td>Short, straight villi with fuzzy glycocalyx</td>
</tr>
</tbody>
</table>

**Fibrous Tumors of the Pleura.** Fibrous tumors of the pleura are unrelated to asbestos exposure or malignant mesotheliomas. They generally occur as a single pedunculated mass arising from the visceral pleura but can occasionally arise from the parietal pleura. They can grow to be quite large, with most ranging from 5 to 10 cm and 100 to 400 g in size by the time they are discovered. Architecturally, the most common microscopic feature is the “patternless pattern.” This is characterized by randomly situated areas of hypercellularity, containing spindle cells with bland, vesicular, ovoidal nuclei and scarce cytoplasm, and hypocellularity, with fibrous connective tissue, hemorrhage, myxoid, or necrosis. They can also have an hemangiopericytoma-like appearance. The neoplastic cells are immunoreactive for CD34 and CD99 but negative for cytokeratins and desmin. Immunoreactivity for Bcl-2 is variably positive. They are frequently discovered incidentally on routine CXRs, without an associated pleural effusion. They occur with equal frequency in males and females and are most common in the sixth to seventh decade of life. Fibrous tumors of the pleura may be benign or malignant. Symptoms such as cough, chest pain, and dyspnea occur in 30% to 40% of patients but are found in 75% of patients with malignant tumors. Malignant tumors are differentiated from benign tumors based on high cellularity, more than four mitotic figures per 10 high-power fields, nuclear pleomorphism, tumor necrosis, and hemorrhage. They are more likely to arise from the parietal pleura of the chest wall, diaphragm, or mediastinum, or in the fissures or invading into the lung parenchyma. Hypoglycemia, associated pleural effusion, and hypertrophic pulmonary osteoarthropathy (clubbed digits, long bone ossifying periostitis, and arthritis) are associated with these lesions in approximately 25% of patients. Less common are fever and hemoptysis. Symptoms resolve with surgical resection. Given the well-circumscribed and often pedunculated nature of fibrous tumors of the pleura, all benign lesions and approximately 50% of malignant lesions are cured by complete surgical resection. Incompletely resected malignant tumors may recur locally or metastasize, and more than 50% of patients with malignant tumors will die from the disease; frequently, they are fatal within 2 to 5 years.

**REFERENCES**

Entries highlighted in bright blue are key references.

48. Warren WH, Faber LP. Segmentectomy versus lobectomy in patients with stage I pulmonary carcinoma. Five-year survival...
85. Casati A, Alessandri P, Nuzzi M, et al. A prospective, randomized, blinded comparison between continuous...
thoracic paravertebral and epidural infusion of 0.2% ropivacaine after lung resection surgery. Eur J Anaesthesiol. 2006;23(12):999-1004.


This page intentionally left blank
INTRODUCTION

Congenital heart surgery is a dynamic and evolving field. The last 20 years have brought about rapid developments in technology, emphasis on a multidisciplinary approach to treatment, and a more thorough understanding of both the anatomy and pathophysiology of congenital heart disease, leading to the improved care of these challenging patients.

These advancements have created a sustained paradigm shift in the field of congenital heart surgery. The traditional strategy of initial palliation followed by definitive correction at a later age, which had pervaded the thinking of most surgeons, began to evolve into emphasizing early repair. Defects such as hypoplastic left heart syndrome (HLHS) are now successfully managed with staged palliation, resulting in excellent survival outcomes for these children.

The goal in most cases of congenital heart disease (CHD) is appropriate timing of complete repair. Rather than subdividing lesions into cyanotic or noncyanotic lesions, a more appropriate classification divides defects into three categories based on the feasibility of achieving complete repair: (a) defects that have no reasonable palliation and for which repair is the only option; (b) defects for which repair is not possible and for which palliation is the only option; and (c) defects that can either be repaired or palliated in infancy. It bears mentioning that all defects in the second category are those in which the appropriate anatomic components either are not present, as in hypoplastic left heart syndrome, or cannot be created from existing structures, i.e., unguarded tricuspid orifice.

Eight out of every 1000 live births will have some form of CHD, most of which, however, are mild.1 In the United States nearly 40,000 infants are affected each year.2 As of 2010, it is estimated that there are about 2 million people living with CHD in the United States, and as of 2011 there are more adults (>18) than children.3 CHD is the most common birth defect and the most common cause of infant death related to birth defects, accounting for 28% of deaths due to birth defects in the first month of life. There are currently 127 centers in North America that perform congenital heart surgery. The Society for Thoracic Surgeons (STS) reports an overall national mortality of 3.1%.3

DEFECTS AMENABLE TO COMPLETE REPAIR

Atrial Septal Defect

An atrial septal defect (ASD) is defined as discontinuity of the interatrial septum that permits direct mixing of blood between the systemic venous and pulmonary venous circulations.

Embryology. The atrial and ventricular septa form between the third and sixth weeks of fetal development. After the paired heart tubes fuse into a single tube folded onto itself, the distal portion of the tube indents to form the roof of the common atrium. Near this portion of the roof, the septum primum originates and descends in a crescentic formation toward the atrioventricular (AV) junction. The ostium primum is situated superiorly to the crux of the heart at the atrioventricular junction. Prior to completion of endocardial cushion fusion with the septum primum, a sequence of fenestrations appear that coalesce into the
ostium secundum. During this coalescence, the septum secundum grows downward from the roof of the atrium, parallel to and to the right of the septum primum. The septum primum does not fuse, but creates an oblique pathway, called the foramen ovale, within the interatrial septum. After birth, the increase in left atrial pressure associated with an increase in systemic vascular resistance (SVR) typically closes this pathway in approximately 80% of the population, obliterating the interatrial communication.

**Anatomy.** ASDs can be classified into three different types (Fig. 20-1): (a) ostium secundum type defect (Fig. 20-1B,C) (deficiency of septum primum), which are the most prevalent subtype, comprising 80% of all ASDs; (b) ostium primum defects (Fig. 20-1A), which may also be described as partial or transitional AV canal defect; and (c) sinus venosus type defects, comprising approximately 5% to 10% of all ASDs.4

**Pathophysiology.** ASDs result in an increase in pulmonary blood flow secondary to primarily left-to-right shunting through the defect. The direction of the intracardiac shunt is predominantly determined by the compliance of the respective ventricles. In utero, the distensibility, or compliance, of the right and left ventricles is equal, but postnatally the left ventricle (LV) becomes less compliant than the right ventricle (RV). This shift occurs because the resistance of the downstream vascular beds changes after birth. The pulmonary vascular resistance falls with the infant’s first breath, decreasing RV pressure, whereas the systemic vascular resistance rises dramatically, increasing LV pressure. The increase in LV pressure promotes hypertrophy with a thicker muscle mass, which offers a greater resistance to diastolic filling than does the RV; thus, the majority of flow through the ASD occurs from left to right. The greater volume of blood returning to the right atrium causes volume overload in the RV, but because of its lower muscle mass and low-resistance output, it easily distends to accommodate the increased volume.

The long-term consequences of RV volume overload include hypertrophy with elevated RV end-diastolic pressure and a relative pulmonary stenosis across the pulmonary valve because it cannot accommodate the increased RV flow. Compliance gradually decreases as the right ventricular pressure approaches systemic pressure, and the size of the left-to-right shunt decreases. Patients at this stage have a balanced circulation and may deceptively appear less symptomatic.
Patients with large ASDs gradually develop progressive pulmonary vascular changes as a result of chronic overcirculation. The increased pulmonary vascular resistance in these patients leads to an equalization of left and right ventricular pressures, and their ratio of pulmonary (Qp) to systemic flow (Qs), Qp to Qs, will approach 1.\(^5\) This does not mean, however, that there is no intracardiac shunting, only that the ratio between the left-to-right component and the right-to-left component is equal.

The ability of the RV to recover normal function is related to the duration of chronic overload because those undergoing ASD closure before age 10 years have a better likelihood of achieving normal RV volumes and function in the postoperative period.\(^6\)

The physiology of sinus venosus ASDs is similar to that discussed earlier, except that these are frequently accompanied by anomalous pulmonary venous drainage. This often results in significant hemodynamic derangements that accelerate the clinical course of these infants.

The same increase in symptoms is true for those with ostium primum defects because the associated mitral insufficiency from the “cleft” mitral valve can lead to more atrial volume load and increased atrial level shunting.

**Diagnosis.** Patients with ASDs upon auscultation may reveal prominence of the first heart sound with fixed splitting of the second heart sound. This results from the relatively fixed left-to-right shunt throughout all phases of the cardiac cycle. A diastolic flow murmur indicating increased flow across the tricuspid valve may be discerned, and frequently, an ejection flow murmur can be heard across the pulmonary valve. A right ventricular heave and increased intensity of the pulmonary component of the second heart sound indicates pulmonary hypertension.

Chest radiographs in the patient with an ASD demonstrate increased pulmonary vascularity, with prominent hilar markings and cardiomegaly. The electrocardiogram shows right axis deviation with an incomplete bundle-branch block. When right bundle-branch block is associated with a leftward or superior axis, an AV canal defect should be strongly suspected.

Diagnosis is clarified by two-dimensional echocardiography (Fig. 20-1A,C), and use of color-flow mapping facilitates an understanding of the physiologic derangements created by the defects. Older children and adults with unrepaired ASDs may present with stroke or systemic embolism from paradoxical embolism or atrial arrhythmias from dilatation of the right atrium.

Echocardiography also enables the clinician to estimate the amount of intracardiac shunting, and it can demonstrate the degree of mitral regurgitation in patients with ostium primum defects. With the addition of an agitated saline injection (bubble study), it can also assist in the detection of sinus venosus defects.

The advent of two-dimensional echocardiography with color-flow Doppler has largely superseded the use of cardiac catheterization because the ASD can be well defined by echocardiography alone. However, in cases where the right ventricular systolic pressure is elevated, or patient is older than age 40 years, catheterization can quantify the degree of pulmonary hypertension because those with a fixed pulmonary vascular resistance greater than 12 U/mL may be considered inoperable.\(^7\)

Cardiac catheterization also can be useful in that it provides data that enable the calculation of Qp and Qs so that the magnitude of the intracardiac shunt can be determined. The ratio (Qp to Qs) can then be used to determine whether closure is indicated in equivocal cases, because a ratio of Qp to Qs greater than 1.5:1

---

**Figure 20-1.** A. Echocardiogram of a patient with primum type atrial septal defect (‘*’ points to the atrial septal defect). B. Echocardiogram of a large secundum type ASD (‘*’ points to the defect). C. Intra-operative picture during repair of atrial septal defect. A large fenestrated atrial septum is seen. Bicaval venous cannulation has been performed and a right atriotomy provides exposure to the atrial septum.
is generally accepted as the threshold for surgical intervention. Finally, in patients older than age 40 years, cardiac catheterization can be important to evaluate for the presence of coronary artery disease.

In general, ASDs are closed when patients are between 4 and 5 years of age. Children of this size can usually be operated on without the use of blood transfusion and have excellent outcomes. Patients who are symptomatic may require repair earlier, even in infancy. Some surgeons advocate routine repair in infants and children especially in cases where prematurity-related lung disease may accelerate damage to the pulmonary vascular bed, though this philosophy may not be widespread. In a review by Reddy and colleagues, 116 neonates weighing less than 2500 g who underwent repair of simple and complex cardiac defects with the use of cardiopulmonary bypass were found to have no intracerebral hemorrhages, no long-term neurologic sequelae, and a low operative mortality rate (10%). These results correlated with the length of cardiopulmonary bypass and the complexity of repair. These investigators also found an 80% actuarial survival at 1 year and, more importantly, that growth following complete repair was equivalent to weight-matched neonates free from cardiac defects.

**Treatment.** Simple secundum type ASDs can frequently be repaired via a transcatheter technique, and assessment for transcatheter closure with TTE assessment is generally indicated prior to consideration of a surgical repair. The most common surgical approach requires standard cardiopulmonary bypass (CPB) technique through a midline sternotomy approach. The details of the repair itself are generally straightforward. An oblique atriotomy is made, the position of the coronary sinus and all systemic and pulmonary veins are determined, and the rim of the defect is completely visualized. Closure of an ostium secundum defect is accomplished either by primary repair or by insertion of a patch that is sutured to the rim of the defect. The decision of whether patch closure is necessary can be determined by the size and shape of the defect as well as by the quality of the edges.

The type of repair used for sinus venosus ASDs associated with partial anomalous pulmonary venous connection is dictated by the location of the anomalous pulmonary vein. If the anomalous veins connect to the atria or to the superior vena cava caudal to where the cava is crossed by the right pulmonary artery, the ASD can be repaired by inserting a patch, with redirection of the pulmonary veins behind the patch to the left atrium. Care must be taken with this approach to avoid obstruction of the pulmonary veins or the superior vena cava, although usually the superior vena cava is dilated and provides ample room for patch insertion. If the anomalous vein connects to the superior vena cava cranial to the right pulmonary artery, an alternative technique, the Warden procedure, may be necessary. In this operation, the superior vena cava is transected cranial to the connection of the anomalous vein (usually the right superior pulmonary vein). The caudal end of the transected cava is oversewn. The cranial end of the transected cava is anastomosed to the auricle of the right atrium. Inside the atrium, a patch is used to redirect pulmonary venous blood flow to the left atrium. In contrast to the repair for a defect where the pulmonary veins enter the right atrium or the superior vena cava below the right pulmonary artery, the patch covers the superior vena cava right atrial junction so that blood from the anomalous pulmonary vein that enters the cava is directed to the left atrium. Blood returning from the upper body enters the right atrium via the anastomosis between the superior vena cava and the right atrial appendage.

**Results and Complications of Surgical ASD Closure.** Traditional operative strategies, such as pericardial or synthetic patch closure, have been well established, with a low complication rate and a mortality rate of zero among patients without pulmonary hypertension. The most frequently reported immediate complications include postpericardiotomy syndrome and atrial arrhythmias. Beyond immediate postoperative outcomes, long-term outcomes following surgical closure (up to 20 years) document the low mortality rates and durability of functional status benefit. Importantly, however, atrial arrhythmias, particularly atrial fibrillation, are not completely mitigated by closure and can occur in 10% to 40% of patients, especially in older patients (>40 years) or those with preexisting arrhythmias. Luo and colleagues performed 300 patients from their institution, 152 of whom had surgical closure. Late mortality at 10 years was 3%, and functional health status had declined in only 15 patients during follow-up. Recently, there have been an increasing number of reports regarding the results following surgical closure among elderly patients (>60 years of age), which demonstrate equivalent survival to younger patients, albeit with slightly higher complication rates. Hanninen and colleagues studied 68 patients between 68 and 86 years at their institution undergoing either surgical (n = 13) or device (n = 54) closure. Although the 23% incidence of major complications (including pneumothorax, heart failure, and pneumonia) was higher than that recently reported by Mascio et al using the Society of Thoracic Surgeons’ Congenital Database (20%) or a single-institution review by Hopkins et al (12%), there were no operative deaths among the elderly cohort. Moreover, after ASD closure, echocardiographic indices of right ventricular size and function were significantly improved from preoperative values, and functional capacity as measured by standardized survey instruments was also significantly improved.

**New and Future Approaches to Traditional Surgical ASD Closure.** Because of the uniformly excellent outcomes with traditional surgery, attention has shifted to improving the cosmetic result and minimizing hospital stay and convalescence. Multiple strategies have been described to achieve these aims, including the right submammary incision with anterior thoracotomy, limited bilateral submammary incision with partial sternal split, and limited midline incision with partial sternal split. Some surgeons use either video-assisted thoracic surgery (VATS) in conjunction with the submammary and transxiphoid approaches to facilitate closure within a constricted operative field or totally endoscopic repair in selected patients. Use of robotics has also been reported in a small series of 12 adult patients by Argenziano and colleagues. The morbidity and mortality of all of these approaches are comparable to those of the traditional median sternotomy; however, each has technical drawbacks. Operative precision must be maintained with limited exposure in any minimally invasive technique. Extended CPB and aortic cross-clamp times, coupled with increased cost, may limit the utility of totally endoscopic or robotic-assisted ASD closure except at specific centers. Moreover, certain approaches have a specific patient population in whom they are most applicable. For example, the anterolateral thoracotomy should not be employed in prepubescent girls because it will interfere with breast development. Most totally endoscopic approaches are not feasible in very young patients because of the size of the thoracoscopic ports. Despite these potential drawbacks, however, in carefully selected patients, minimally invasive techniques have demonstrated benefits. Luo and associates performed
a prospective randomized study comparing ministernotomy (division of the upper sternum for aortic and pulmonary lesions and the lower sternum for septal lesions) to full sternotomy in 100 consecutive patients undergoing repair of septal lesions.\textsuperscript{19} The patients in the ministernotomy group had longer procedure times (by 15 to 20 minutes) but had less bleeding and shorter hospital stays. Consistent with these initiatives, conversion of “low-risk” patients undergoing minimally invasive ASD closure to an ambulatory population (discharge from hospital within 24 hours) has recently been described.\textsuperscript{21}

First performed in 1976, transcatheter closure of ASDs with the use of various occlusion devices is gaining widespread acceptance.\textsuperscript{22} Certain types of ASDs, including patent foramen ovale, secundum defects, and some fenestrated secundum defects, are amenable to device closure, as long as particular anatomic criteria (e.g., an adequate superior and inferior rim for device seating and distance from the AV valve) are met. Since the introduction of percutaneous closure (Fig. 20-2A,B), there has been a dramatic rise in device closure prevalence to the point where device closure has supplanted surgical therapy as the dominant treatment modality for secundum ASD.\textsuperscript{23} A study from Karamlou et al\textsuperscript{23} found that ASD and patent foramen ovale closures per capita increased dramatically from 1.08 per 100,000 population in 1988 to 2.59 per 100,000 population in 2005, an increase of 139%. When analyzed by closure type, surgical closure increased by only 24% (from 0.86 per 100,000 population in 1988 to 1.07 per 100,000 in 2005), whereas transcatheter closure increased by 3475% (from 0.04 per 100,000 population in 1988 to 1.43 per 100,000 in 2005). Importantly, this study determined that the paradigm shift favoring transcatheter closure has occurred mainly due to increased prevalence of closure in adults over age 40 years rather than an increase in closure in infants or children.

Despite the simplicity of ASD repair, there are a myriad of options for patients and physicians who care for patients with CHD. The patient population that might benefit from closure (whether device or surgical) is likely to increase, challenging current ideas and treatment algorithms that optimize outcomes.

### Aortic Stenosis

**Anatomy and Classification.** The spectrum of aortic valve abnormality represents the most common form of CHD, with the great majority of patients being asymptomatic until midlife. Obstruction of the left ventricular outflow tract (LVOT) occurs at multiple levels: subvalvular, valvular, and supraavalvular (Fig. 20-3A-D). The critically stenotic aortic valve in the neonate or infant is commonly unicommissural or bicommissural, with thickened, dysmorphic, and myxomatous leaflet tissue and a reduced cross-sectional area at the valve level. Associated left-sided lesions are often present. In a review of 32 cases from the Children’s Hospital in Boston, 59% had unicommissural valves, and 40% had bicommissural valves.\textsuperscript{24} Associated lesions were frequent, occurring in 88% of patients, most commonly patent ductus arteriosus, mitral regurgitation, and hypoplastic LV. Endocardial fibroelastosis (EFE) also is common among infants with critical aortic stenosis (AS). In this condition, the LV is usually prohibitively hypoplastic and noncompliant, rendering these patients poor candidates for recruitment of the LV into the systemic circulation with techniques that can be utilized in those with more normal sized LVs. In some neonates with critical AS, a dilated LV with poor diastolic compliance rather than a hypertrophied LV is encountered.\textsuperscript{24}

Neonates with critical AS are a challenging population because one must make a decision about the suitability of the left-sided structures to support a biventricular circulation. There are recent approaches that include techniques, such as aortic valvotomy coupled with EFE resection and mitral valve intervention, that are directed at LV rehabilitation. The advent of fetal valvotomy for critical AS may also increase the number of infants who are candidates for biventricular repair.

**Pathophysiology.** The unique intracardiac and extracardiac shunts present in fetal life allow even neonates with critical AS to survive. In utero, left ventricular hypertrophy and ischemia cause left atrial hypertension, which reduces the right-to-left flow across the foramen ovale. In severe cases, a reversal of
Figure 20-3. A. Congenital aortic valve stenosis, en fosse echocardiographic view of the stenotic bicuspid aortic valve. Parasternal long axis view of the same valve with a gradient of 60 mm of Hg (‘*’ points to the valve). B. Parasternal long axis ecocardiographic view of a patient with discrete subaortic membrane (‘*’ points to the membrane). C. Parasternal long axis ecocardiographic view of a patient with diffuse tunnel like subvalvar aortic stenosis with membrane. Doppler revealed a gradient of 81 mm of Hg (‘*’ represents the area of diffuse narrowing). D. Appearance of supravalvar aortic stenosis on an aortogram performed in the cardiac catheterization lab (‘*’ points to the stenosis). E. Appearance after four patch reconstruction of the same patient shown in Figure 20.3 d. (Re-formatted images obtained from a CT angiogram).
flow may occur, causing right ventricular volume loading. The RV then provides the entire systemic output via the patent ductus arteriosus (ductal-dependent systemic blood flow). Although cardiac output is maintained, the LV suffers continued damage as the intracavitary pressure precludes adequate coronary perfusion, resulting in LV infarction and subendocardial fibroelastosis. The presentation of the neonate with critical AS is then determined by the morphology of the LV and other left-sided heart structures, the degree of left ventricular dysfunction, and the completeness of the transition from a parallel circulation to an in-series circulation (i.e., on closure of the foramen ovale and the ductus arteriosus). Those infants with mild-to-moderate AS in whom LV function is preserved are asymptomatic at birth. The only abnormalities may be a systolic ejection murmur and electrocardiogram (ECG) evidence of left ventricular hypertrophy. However, those neonates with severe AS and compromised LV function are unable to provide adequate cardiac output at birth and will present in circulatory collapse once the ductus closes, with dyspnea, tachypnea, irritability, narrowed pulse pressure, oliguria, and profound metabolic acidosis. If ductal patency is maintained, systemic perfusion will be provided by the RV via ductal flow, and cyanosis may be the only finding.

**Diagnosis.** Neonates and infants with severe valvar AS may have a relatively nonspecific history of irritability and failure to thrive. Angina, if present, is usually manifested by episodic, inconsolable crying that coincides with feeding. As discussed previously, evidence of poor peripheral perfusion, such as extreme pallor, indicates severe LVOT obstruction. Differential cyanosis is an uncommon finding, but it is present when enough antegrade flow occurs only to maintain normal upper body perfusion, while a large patent ductus arteriosus produces blue discoloration of the abdomen and legs.

Physical findings include a systolic ejection murmur, although a quiet murmur may paradoxically indicate a more severe condition with reduced cardiac output. A systolic click correlates with a valvular etiology of obstruction. As LV dysfunction progresses, evidence of congestive heart failure occurs.

The chest radiograph is variable but may show dilatation of the aortic root, and the ECG often demonstrates LV hypertrophy. Echocardiography with Doppler flow is extremely useful in establishing the diagnosis, as well as quantifying the transvalvular gradient. Furthermore, echocardiography can facilitate evaluation for the several associated defects that can be present in critical neonatal AS, including mitral stenosis, LV hypoplasia, LV endocardial fibroelastosis, subaortic stenosis, VSD, or coarctation. The presence of any or several of these defects has important implications related to treatment options for these patients. Although cardiac catheterization is not routinely performed for diagnostic purposes, it can be invaluable as part of the treatment algorithm if the lesion is amenable to balloon valvotomy. Magnetic resonance imaging (MRI) is another very useful technique for assessing the adequacy of the left-sided structures and is increasingly utilized to determine candidacy for biventricular repairs.

**Treatment.** As alluded to previously, the first decision that must be made in the neonate with critical LVOT obstruction is whether the patient is a candidate for biventricular or univentricular repair. Central to this decision is assessment of the degree of hypoplasia of the LV and other left-sided structures. Alsoufi and colleagues have described a rational approach to the neonate with critical LVOT obstruction. The options vary depending on whether the infant follows a single or a biventricular pathway. The options for a single ventricle include the Norwood operation, a hybrid strategy (initial ductal stent and bilateral pulmonary artery bands followed by later completion of the Norwood operation) or heart transplantation. The options for a biventricular heart include balloon valvuloplasty, surgical valvotomy, neonatal Ross operation, or a Yasui operation. Often valvotomy is accompanied by LV rehabilitation techniques, including EFE resection and mitral valve interventions. Fetal aortic valvotomy, which is now offered at specialized centers, is another promising strategy to decompress the LV in fetal life and potentially allow growth of the left-sided structures sufficient to permit a biventricular circulation. Regardless of whether the baby is triaged to a single or biventricular strategy, any infant with severe AS requires urgent intervention. Preoperative stabilization, however, has dramatically altered the clinical algorithm and outcomes for this patient population. The preoperative strategy begins with endotracheal intubation and inotropic support. Prostaglandin infusion is initiated to maintain ductal patency, and confirmatory studies are performed prior to operative intervention. Therapy is generally indicated in the presence of a transvalvular gradient of 50 mmHg with associated symptoms including syncope, CHF, or angina, or if a gradient of 50 to 75 mmHg exists with concomitant ECG evidence of LV strain or ischemia. In the critically ill neonate, a gradient across the aortic valve may not be present because of poor LV function. However, the decision regarding treatment options must be based on a complete understanding of associated defects. For example, in the presence of a hypoplastic LV (left ventricular end-diastolic volume <20 mL/m²) or a markedly abnormal mitral valve, isolated aortic valvotomy should not be performed because studies have demonstrated high mortality in this population following isolated valvotomy.

Patients who have an LV capable of providing systemic output are candidates for intervention to relieve AS, generally through balloon valvotomy. Occasionally, if catheter-based therapy is not an option, relief of valvar AS in infants and children can be accomplished with surgical valvotomy using standard techniques of CPB and direct exposure to the aortic valve. A transverse incision is made in the ascending aorta above the sinus of Valsalva, extending close to, but not into, the noncoronary sinus. Exposure is attained with placement of a retractor into the right coronary sinus. After inspection of the valve, the chosen commissure is incised to within 1 to 2 mm of the aortic wall (Fig. 20-4A,B).

Balloon valvotomy performed in the catheterization lab is generally the procedure of choice for reduction of transvalvular gradients in symptomatic infants and children without significant aortic insufficiency. Balloon valvotomy provides relief of the valvar gradient and allows future surgical intervention (which is generally required in most patients when a larger prosthesis can be implanted) to be performed on an unscarred chest. An important issue when planning aortic valvotomy, whether percutaneously or via open surgical technique, is the risk of inducing hemodynamically significant aortic regurgitation. Induction of more than moderate aortic regurgitation is risky and may result in hemodynamic instability. A gradient of 50 mmHg is generally the upper limit of safety. The decision regarding
The majority of infants who undergo aortic valvotomy will require further intervention on the aortic valve within 10 years following initial intervention.26

Neonates with severely hypoplastic LVs or significant LV endocardial fibroelastosis may not be candidates for biventricular repair and are treated the same as infants with the hypoplastic left heart syndrome (HLHS), which is discussed later (see “Hypoplastic Left Heart Syndrome”).

As mentioned previously, fetal intervention for the aortic valve has been described with the goal being to improve the growth of the left ventricle. The group at Boston Children’s Hospital have reported fairly favorable results in a small cohort.34

Many surgeons previously avoided aortic valve replacement for AS in early childhood because the more commonly used mechanical valves would be outgrown and require replacement later and the obligatory anticoagulation for mechanical valves resulted in a substantial risk for complications. In addition, prosthetic valves have an incidence of bacterial endocarditis or perivalvular leak requiring reintervention.

The use of allografts and the advent of the Ross procedure have largely obviated these issues and made early definitive correction of critical AS a viable option.23,27,28 Donald Ross first described transposition of the pulmonary valve into the aortic position with allograft reconstruction of the pulmonary outflow tract in 1967.27 The result of this operation is a normal trileaflet semilunar valve made of a patient’s native tissue with the potential for growth to adult size in the aortic position in place of the damaged aortic valve (Fig. 20-5). The Ross procedure has become a useful option for aortic valve replacement in children because it has improved durability and can be performed with acceptable morbidity and mortality rates. The placement of a pulmonary conduit, which does not grow and becomes calcified and stenotic over time, does obligate the patient to reintervention (either surgically or using transcatheter techniques) to replace the RV-to-pulmonary artery conduit. Karamlou and colleagues29 have reviewed the outcomes and associated risk factors for aortic valve replacement in 160 children from the Hospital for Sick Children in Toronto. They found that younger age, lower operative weight, concomitant performance of aortic root replacement or reconstruction, and use of prosthesis type other than a pulmonary autograft were significant predictors of death, whereas the use of a bioprosthetic or allograft valve type and earlier year of operation were identified as significant risk factors for repeated aortic valve replacement. Autograft use was associated with a blunted progression of the peak prosthetic valve gradient and a rapid decrease in the left ventricular end-diastolic dimension. In agreement with these findings, Lupinetti and colleagues29 have reviewed the outcomes and associated risk factors for aortic valve replacement with the Ross procedure and found a more significant transvalvular gradient reduction and regression of left ventricular hypertrophy in those patients who underwent the Ross procedure. In some cases, the pulmonary valve may not be usable because of associated defects or congenital absence. These children are not candidates for the Ross procedure and can be treated with cryopreserved allografts (cadaveric human aortic valves) or prosthetic aortic valve replacement. At times, there may be a size discrepancy between the right ventricular outflow tract (RVOT) and the LVOT, especially in cases of severe critical AS in infancy. For these cases, the pulmonary autograft is placed in a manner that also provides enlargement of the aortic annulus (Ross/Konno).

Subvalvular AS occurs beneath the aortic valve and may be classified as discrete or tunnel-like (diffuse). A thin,
fibromuscular diaphragm immediately proximal to the aortic valve characterizes discrete subaortic stenosis. This diaphragm typically extends for 180° or more in a crescentic or circular fashion, often attaching to the mitral valve as well as the interventricular septum. The aortic valve itself is usually normal in this condition, although the turbulence imparted by the subvalvular stenosis may affect leaflet morphology and valve competence.

Diffuse subvalvular AS results in a long, tunnel-like obstruction that may extend to the left ventricular apex. In some individuals, there may be difficulty in distinguishing between hypertrophic cardiomyopathy and diffuse subaortic stenosis. Operation for subvalvular AS is indicated with a gradient exceeding 30 mmHg, in the presence of aortic valve insufficiency, or when symptoms indicating LVOT obstruction are present. Given that repair of isolated discrete subaortic stenosis can be done with low rates of morbidity and mortality, some surgeons advocate repair in all cases of discrete AS to avoid progression of the stenosis and the development of aortic insufficiency, although more recent data demonstrate that subaortic resection should be delayed until the LV gradient exceeds 30 mmHg because most children with an initial LV gradient less than 30 mmHg have quiescent disease. Diffuse AS is a more complex lesion and often requires aortoventriculoplasty. Results are generally excellent, with operative mortality less than 5%.

Supravalvular AS occurs more rarely and also can be classified into a discrete type, which produces an hourglass deformity of the aorta, and a diffuse form that can involve the entire arch and brachiocephalic arteries. The aortic valve leaflets are usually normal, but in some cases, the leaflets may adhere to the supravalvular stenosis, thereby narrowing the sinuses of Valsalva in diastole and restricting coronary artery perfusion. In addition, accelerated intimal hyperplastic changes in the coronary arteries can be demonstrated in these patients because the proximal position of the coronary arteries subjects them to abnormally high perfusion pressures.

The signs and symptoms of supravalvular AS are similar to other forms of LVOT obstruction. An asymptomatic murmur is the presenting manifestation in approximately half of these patients. Syncope, poor exercise tolerance, and angina may all occur with nearly equal frequency. Supravalvar AS is associated with Williams’ syndrome, a constellation of elfin facies, mental retardation, and hypercalcemia. Following routine evaluation, cardiac catheterization should be performed in order to delineate coronary anatomy, as well as to delineate the degree of obstruction. A gradient of 50 mmHg or greater is an indication for operation. However, the clinician must be cognizant of any coexistent lesions, most commonly pulmonic stenosis, which may add complexity to the repair.

The localized form of supravalvular AS can be treated by creating an inverted Y-shaped aortotomy across the area of stenosis, straddling the right coronary artery. The obstructing shelf is then excised, and a pantaloon-shaped patch (Doty technique) or individual sinus patch enlargement (Brom technique) is used (Fig. 20-3E).

The diffuse form of supravalvular stenosis is more variable (Fig. 20-6), and the particular operative approach must be tailored to each specific patient’s anatomy. In general, either an aortic endarterectomy with patch augmentation can be performed or if the narrowing extends past the aorta arch, a prothetic graft can be placed between the ascending and descending aorta. Operative results for discrete supravalvular AS are generally good, with a hospital mortality of less than 1% and an actuarial survival rate exceeding 90% at 20 years. In contrast, however, the diffuse form is more hazardous to repair and carried a mortality of 15% in a recent series.

Patent Ductus Arteriosus

Anatomy. The ductus arteriosus is derived from the sixth aortic arch and normally extends from the main or left pulmonary artery to the upper descending thoracic aorta, distal to the left subclavian artery. In the normal fetal cardiovascular system, ductal flow is considerable (approximately 60% of the combined ventricular output) and is directed exclusively from the pulmonary artery to the aorta. In infancy, the length of the ductus may vary from 2 to 8 mm, with a diameter of 4 to 12 mm.

Locally produced and circulating prostaglandin E2 (PGE2) and prostaglandin I2 (PGI2) induce active relaxation of the ductal musculature, maintaining maximal patency during the fetal period. At birth, increased pulmonary blood flow metabolizes these prostaglandin products, and absence of the placenta removes an important source of them, resulting in a marked decrease in these ductal-relaxing substances. In addition, release of histamines, catecholamines, bradykinin, and acetylcholine all promote ductal contraction. Despite all of these complex...
interactions, the rising oxygen tension in the fetal blood is the main stimulus causing smooth muscle contraction and ductal closure within 10 to 15 hours postnatally. Anatomic closure by fibrosis produces the ligamentum arteriosum connecting the pulmonary artery to the aorta.

Delayed closure of the ductus is termed prolonged patency, whereas failure of closure causes persistent patency, which may occur as an isolated lesion or in association with more complex congenital heart defects. In many of these infants with more complex congenital heart defects, either pulmonary or systemic perfusion may depend on ductal flow, and these infants may decompensate if exogenous PGE is not administered to maintain ductal patency.

Natural History. The incidence of patent ductus arteriosus (PDA) is approximately 1 in every 2000 births; however, it increases dramatically with increasing prematurity. In some series, PDAs have been noted in 75% of infants of 28 to 30 weeks gestation. Persistent patency occurs more commonly in males, with a 2:1 ratio.

PDA is not a benign entity, although prolonged survival has been reported. The estimated death rate for infants with isolated, untreated PDA is approximately 30%. The leading cause of death is congestive heart failure, with respiratory infection as a secondary cause. Endocarditis is more likely to occur with a small ductus and is rarely fatal if aggressive antibiotic therapy is initiated early.

Clinical Manifestations and Diagnosis. After birth, in an otherwise normal cardiovascular system, a PDA results in a left-to-right shunt that depends on both the size of the ductal lumen and its total length. As the pulmonary vascular resistance falls 8 to 10 weeks postnatally, the shunt will increase, and its flow will ultimately be determined by the relative resistances of the pulmonary and systemic circulations.

The hemodynamic consequences of an unrestrictive ductal shunt are left ventricular volume overload with increased left atrial and pulmonary artery pressures and right ventricular strain from the augmented afterload. These changes result in increased sympathetic discharge, tachycardia, tachypnea, and ventricular hypertrophy. The diastolic shunt results in lower aortic diastolic pressure and increases the potential for myocardial ischemia and underperfusion of other systemic organs, while the increased pulmonary flow leads to increased work of breathing and decreased gas exchange. Unrestrictive ductal flow may lead to pulmonary hypertension within the first year of life. These changes will be significantly attenuated if the size of the ductus is only moderate, and they will be completely absent if the ductus is small.

Physical examination of the afflicted infant will reveal evidence of a hyperdynamic circulation with a widened pulse pressure and a hyperactive precordium. Auscultation demonstrates a systolic or continuous murmur, often termed a machinery murmur. Cyanosis is not present in uncomplicated isolated PDA.

The chest radiograph may reveal increased pulmonary vascularity or cardiomegaly, and the ECG may show LV strain, left atrial enlargement, and possibly RV hypertrophy. Echocardiogram with color mapping reliably demonstrates the patency of the ductus as well as estimates the shunt size. Cardiac catheterization is necessary only when pulmonary hypertension is suspected.

Therapy. The presence of a persistent PDA is sufficient indication for closure because of the increased mortality and risk of endocarditis. In older patients with pulmonary hypertension, closure may not improve symptoms and is associated with much higher mortality.

In premature infants, aggressive intervention with indomethacin or ibuprofen to achieve early closure of the PDA is beneficial unless contraindications such as necrotizing enterocolitis or renal insufficiency are present. Term infants, however, are generally unresponsive to pharmacologic therapy with indomethacin, so mechanical closure must be undertaken once the diagnosis is established. This can be accomplished either surgically (Fig. 20-7) or with catheter-based therapy. Currently, transluminal placement of various occlusive devices, such as the Rashkind double-umbrella device or embolization with Gianturco coils, is in widespread use. However, there are a number of complications inherent with the use of percutaneous devices, such as thromboembolism, endocarditis, incomplete occlusion, vascular injury, and hemorrhage secondary to perforation. In addition, these techniques may not be applicable in very young infants because the peripheral vessels do not provide adequate access for the delivery devices. Attempts are being made to develop such devices for premature infants with early successful results in study populations.

Surgical closure can be achieved via either open or video-assisted approaches. The open approach employs a muscle-sparing posterior lateral thoracotomy in the third or fourth intercostal space on the side of the aorta (generally the left). The lung is then retracted anteriorly. In the neonate, the PDA is singly ligated with a surgical clip or permanent suture. Care must be taken to avoid the recurrent laryngeal nerve, which courses around the aortic arch. Occasionally, a short, broad ductus, in which the dimension of

Figure 20-6. Reformatted image obtained after CT angiography of a child with diffuse supravalvar aortic stenosis (‘*’ points to the transverse aortic arch).
its width approaches that of its length, will be encountered. In this case, division between vascular clamps with oversewing of both ends is advisable (Fig. 20-8). In extreme cases, the use of CPB to decompress the large ductus during ligation is an option. Video-assisted thoracoscopic occlusion, using metal clips, also has been described, although it offers few advantages over the standard surgical approach. Preterm newborns and children may do well with a surgical technique, while older patients (older than age 5 years) and those with smaller ducts (<3 mm) do well with coil occlusion. In fact, Moore and colleagues recently concluded from their series that coil occlusion is the procedure of choice for ducts smaller than 4 mm.45 Complete closure rates using catheter-based techniques have steadily improved.

**Outcomes.** In premature infants, the surgical mortality is very low, although the overall hospital death rate is significant as a consequence of other complications of prematurity. In older infants and children, mortality is less than 1%. Bleeding, chylothorax, vocal cord paralysis, and the need for reoperation occur infrequently. With the advent of muscle-sparing thoracotomy, the risk of subsequent arm dysfunction or breast abnormalities is virtually eliminated.46

**Aortic Coarctation**

**Anatomy.** Coarctation of the aorta (COA) is defined as a luminal narrowing in the aorta that causes an obstruction to blood flow. This narrowing is most commonly located distal to the left subclavian artery. The embryologic origin of COA is a subject of some controversy. One theory holds that the obstructing shelf, which is largely composed of tissue found within the ductus, forms as the ductus involutes.47 The other theory holds that a diminished aortic isthmus develops secondary to decreased aortic flow in infants with enhanced ductal circulation.

Extensive collateral circulation develops, predominantly involving the intercostals and mammary arteries as a direct result of aortic flow obstruction. This translates into the well-known finding of “rib-notching” on chest radiograph, as well as a prominent pulsation underneath the ribs.

Other associated anomalies, such as ventricular septal defect, PDA, and ASD, may be seen with COA, but the most common is that of a bicuspid aortic valve, which can be demonstrated in 25% to 42% of cases.48

**Pathophysiology.** Infants with COA develop symptoms consistent with left ventricular outflow obstruction, including pulmonary overcirculation and, later, biventricular failure. In addition, proximal systemic hypertension develops as a result of mechanical obstruction to ventricular ejection, as well as hypoperfusion-induced activation of the renin-angiotensin-aldosterone system.
Interestingly, hypertension is often persistent after surgical correction despite complete amelioration of the mechanical obstruction and pressure gradient.\textsuperscript{49} It has been shown that early surgical correction may prevent the development of long-term hypertension, which undoubtedly contributes to many of the adverse sequelae of COA, including the development of circle of Willis aneurysms, aortic dissection and rupture, and an increased incidence of coronary arteriopathy with resulting myocardial infarction.\textsuperscript{50}

**Diagnosis.** COA is likely to become symptomatic either in the newborn period if other anomalies are present or in the late adolescent period with the onset of left ventricular failure.

Physical examination will demonstrate a hyperdynamic precordium with a harsh murmur localized to the left chest and back. Femoral pulses will be dramatically decreased when compared to upper extremity pulses, and differential cyanosis may be apparent until ductal closure.

Echocardiography will reliably demonstrate the narrowed aortic segment, as well as define the pressure gradient across the stenotic segment. In addition, detailed information regarding other associated anomalies can be gleaned. Aortography (Fig. 20-9) is reserved for those cases in which the echocardiographic findings are equivocal. Cross-sectional imaging with computed tomography (CT) scan or MRI is also increasing to facilitate definition of arch anatomy (i.e., transverse arch hypoplasia), assess intracardiac volumes, and associated defects.

**Therapy.** The routine management of hemodynamically significant COA in all age groups has traditionally been surgical. Transcatheter repairs (Fig. 20-10) are used with increasing frequency in older patients and those with recoarctation following surgical repair. Balloon dilatation of native coarctation in neonates generally is avoided because of the high recoarctation rate. However, in infants who present with severely depressed LV function and a closed ductus arteriosus, initial decompression with balloon dilation of the COA followed by later surgical intervention may be useful. The most common surgical techniques in current use are resection with end-to-end anastomosis or extended end-to-end anastomosis, taking care to remove all residual ductal tissue.\textsuperscript{51,52} Extended end-to-end anastomosis (Fig. 20-11) may also allow the surgeon to treat transverse arch hypoplasia, which is commonly encountered in infants with aortic coarctation.\textsuperscript{53,54} The subclavian flap

\textbf{Figure 20-9.} Reformatted images obtained from CT angiography of a baby showing a discrete coarctation of the aorta (\textsuperscript{*} points to the coarctation).

\textbf{Figure 20-10.} \textbf{A.} Reformatted images obtained from a CT angio-gram of a child with discrete coarctation of the aorta (\textsuperscript{*} points to the coarctation). \textbf{B.} Aortogram performed in the cardiac catheterization lab after stenting the coarctation (\textsuperscript{*} points to the stent).
Aortoplasty is another repair, although it is used less frequently in the modern era because of the risk of late aneurysm formation and possible underdevelopment of the left upper extremity ischemia. In this method, the left subclavian artery is transected and brought down over the coarcted segment as a vascularized patch. The main benefit of these techniques is that they do not involve the use of prosthetic materials, and evidence suggests that extended end-to-end anastomosis may promote arch growth, especially in infants with the smallest initial aortic arch diameters.

Despite the benefits, however, extended end-to-end anastomosis may not be feasible when there is a long segment of coarctation or in the presence of previous surgery because sufficient mobilization of the aorta above and below the lesion may not be possible. In this instance, prosthetic materials, such as a patch aortoplasty, in which a prosthetic patch is used to enlarge the coarcted segment, or an interposition tube graft must be employed. One of the most important decisions in infants and neonates with COA and some degree of transverse arch hypoplasia is whether the lesion should be approached with a sternotomy or a thoracotomy. Cross-sectional imaging with CT scan can be extremely helpful in assessing the adequacy of the transverse arch and any associated abnormalities with branching that may complicate repair from the side.

The most common complications after COA repair are late restenosis (Fig. 20-12) and aneurysm formation at the repair site. Aneurysm formation is particularly common after patch aortoplasty when using Dacron material. In a large series of 891 patients, aneurysms occurred in 5.4% of the total, with 89% occurring in the group who received Dacron-patch aortoplasty and only 8% occurring in those who received resection with primary end-to-end anastomosis. A further complication, although uncommon, is lower-body paralysis resulting from ischemic spinal cord injury during the repair. This dreaded outcome complicates 0.5% of all surgical repairs, but its incidence can be lessened with the use of some form of distal perfusion, preferably left heart bypass with the use of femoral arterial or distal thoracic aorta for arterial inflow and the femoral vein or left atrium for venous return. These techniques are generally reserved for older patients with complex coarctations that may need prolonged aortic cross clamp times for repair, often in the setting of large collateral vessels and/or previous surgery.

Hypertension is also well recognized following repair of COA. Bouchart and colleagues reported that in a cohort of 35 hypertensive adults (mean age, 28 years) undergoing repair, despite a satisfactory anatomic outcome, only 23 patients were normotensive at a mean follow-up period of 165 months. Likewise, Bhat and associates reported that in a series of 84 patients (mean age at repair, 29 years), 31% remained hypertensive at a mean follow-up of 5 years following surgery.

Although operative repair is still the gold standard, treatment of COA by catheter-based intervention has become more widespread for older children and adults. Both balloon dilatation and primary stent implantation have been used successfully. The most extensive study of the results of balloon angioplasty reported on 970 procedures: 422 native and 548 recurrent COAs. Mean gradient reduction was 74% ± 24% for native and 70% ± 31% for recurrent COA. This demonstrated that catheter-based therapy could produce equally effective results both in recurrent and in primary COA, a finding with far-reaching implications in the new paradigm of multidisciplinary treatment algorithms for CHD. In the Valvuloplasty and Angioplasty of Congenital Anomalies (VACA) report, higher preangioplasty gradient, earlier procedure date, older patient age, and the presence of recurrent COA were independent risk factors for suboptimal procedural outcome.

The gradient after balloon dilatation in most series is generally acceptable. However, there is a significant minority of patients (0%–26%) for whom the procedural outcome is suboptimal, with a postprocedure gradient of 20 mmHg or greater. These patients may be ideal candidates for primary stent placement. Deaths from the procedure also are infrequent (<1% of cases), and the main major complication is aneurysm formation,
which occurs in 7% of patients. With stent implantation, many authors have demonstrated improved resolution of stenosis compared with balloon dilatation alone, yet the long-term complications on vessel wall compliance remain largely unknown because only mid-term data are widely available.

In summary, children younger than age 6 months with native COA should be treated with surgical repair, while those requiring intervention at later ages may be ideal candidates for balloon dilatation or primary stent implantation. Additionally, catheter-based therapy should be employed for those cases of restenosis following either surgical or primary endovascular management.

**Truncus Arteriosus**

**Anatomy.** Truncus arteriosus is a rare anomaly, comprising between 1% and 2% of all live born cases of CHD. It is characterized by a single great artery that arises from the heart, overrides the ventricular septum, and supplies the pulmonary, systemic, and coronary circulations.

The two major classification systems are those of Collett and Edwards, described in 1949, and Van Praagh, described in 1965 (Fig. 20-13). The Collett and Edwards classification focuses mainly on the origin of the pulmonary arteries from the common arterial trunk, whereas the Van Praagh system is based on the presence or absence of a VSD, the degree of formation of the aortopulmonary septum, and the status of the aortic arch.

During embryonic life, the truncus arteriosus normally begins to separate and spiral into a distinguishable anterior pulmonary artery and posterior aorta. Persistent truncus, therefore, represents an arrest in embryologic development at this stage. Other implicated events include twisting of the dividing truncus because of ventricular looping, subinfundibular atresia, and abnormal location of the semilunar valve anlagen.

The neural crest may also play a crucial role in the normal formation of the great vessels, as experimental studies in chick embryos have shown that ablation of the neural crest results in persistent truncus arteriosus. The neural crest also develops into the pharyngeal pouches that give rise to the thymus and parathyroids, which likely explains the prevalent association of truncus arteriosus and DiGeorge’s syndrome.

The annulus of the truncal valve usually straddles the ventricular septum in a “balanced” fashion; however, it is not unusual for it to be positioned predominantly over the RV, which increases the potential for LVOT obstruction following surgical repair. In the great majority of cases, the leaflets are thickened and deformed, which leads to valvular insufficiency. There are usually three leaflets (60%), but occasionally a bicuspid (5%) or even a quadricuspid valve (25%) is present.

In truncus arteriosus, the pulmonary trunk bifurcates, with the left and right pulmonary arteries forming posteriorly and to the left in most cases. The caliber of the pulmonary arterial branches is usually normal, with stenosis or diffuse hypoplasia occurring in rare instances.

The coronary arteries may be normal; however, anomalies are not unusual and occur in 50% of cases. Many of these are relatively minor, although two variations are of particular importance because they have implications in the conduct of operative repair. The first is that the left coronary ostium may arise high in the sinus of Valsalva or even from the truncal tissue at the margin of the pulmonary artery tissue. This coronary artery can be injured during repair when the pulmonary arteries are removed from the trunk or when the resulting truncal defect is closed. The second is that the right coronary artery can give rise to an important accessory anterior descending artery, which often passes across the RV in the exact location where the right ventriculotomy is commonly performed during repair.

**Physiology and Diagnosis.** The main pathophysiologic consequences of truncus arteriosus are (a) the obligatory mixing of systemic and pulmonary venous blood at the level of the ventricular septal defect (VSD) and truncal valve, which leads to arterial saturations near 85% and (b) the presence of a nonrestrictive left-to-right shunt, which occurs during both systole and diastole, the volume of which is determined by the relative resistances of the pulmonary and systemic circulations. Additionally, truncal valve stenosis or regurgitation, the presence of important LVOT obstruction, and stenosis of pulmonary artery branches can further contribute to both pressure and volume-loading of the ventricles. The presence of these lesions often results in severe heart failure and cardiovascular instability early in life. Pulmonary vascular resistance may develop as early as 6 months of age, leading to poor results with late surgical correction.

Patients with truncus arteriosus usually present in the neonatal period, with signs and symptoms of congestive heart failure and mild to moderate cyanosis. A pansystolic murmur may be noted at the left sternal border, and occasionally a diastolic murmur may be heard in the presence of truncal regurgitation.

Chest radiography will be consistent with pulmonary overcirculation, and a right aortic arch can be appreciated 35% of the time. The thymus is prominent by its absence if associated with DiGeorge syndrome (Fig. 20-14). The ECG is usually nonspecific, demonstrating normal sinus rhythm with biventricular hypertrophy.

Echocardiography with Doppler color-flow or pulsed Doppler is diagnostic and usually provides sufficient information to determine the type of truncus arteriosus, the origin of the
coronary arteries and their proximity to the pulmonary trunk, the character of the truncal valves, and the extent of truncal insufficiency (Fig. 20-15). CT scan helps define the pulmonary arteries and the coronary anatomy (Fig. 20-16). Cardiac catheterization can be helpful in cases where pulmonary hypertension is suspected or to further delineate coronary artery anomalies prior to repair.

The presence of truncus is an indication for surgery. Repair should be undertaken in the neonatal period or as soon as the diagnosis is established.

**Repair.** Truncus arteriosus was first managed with pulmonary artery banding as described by Armer and colleagues in 1961. However, this technique led to only marginal improvements in 1-year survival rates because ventricular failure inevitably occurred. In 1967, however, complete repair was accomplished by McGoon and his associates based on the experimental work of Rastelli, who introduced the idea that an extracardiac valved conduit could be used to restore ventricular-to-pulmonary artery continuity. Over the next 20 years, improved survival rates led to uniform adoption of complete repair even in the youngest and smallest infants.

Surgical correction entails the use of CPB. Repair is completed by separation of the pulmonary arteries from the aorta, closure of the aortic defect (occasionally with a patch) to minimize coronary flow complications, placement of a valved cryopreserved allograft or jugular venous valved conduit (Contegra) to reconstruct the RVOT, and VSD closure. Important branch pulmonary arterial stenosis should be repaired at the time of complete repair and can usually be accomplished with longitudinal allograft patch arterioplasty. Severe truncal valve insufficiency occasionally requires truncal valve repair or even replacement, which can be accomplished with a cryopreserved allograft.

**Results.** The results of complete repair of truncus have steadily improved. Ebert reported a 91% survival rate in his series of 77 patients who were younger than 6 months of age; later reports by others confirmed these findings and demonstrated that excellent results could be achieved in even smaller infants with complex-associated defects.

Newer extracardiac conduits also have been developed and used with success, which has widened the repertoire of the modern congenital heart surgeon and improved outcomes. Severe truncal regurgitation, interrupted aortic arch, coexistent coronary anomalies, chromosomal or genetic anomalies, and age younger than 100 days are risk factors associated with perioperative death and poor outcome.

**Total Anomalous Pulmonary Venous Connection**

Total anomalous pulmonary venous connection (TAPVC) occurs in 1% to 2% of all cardiac malformations and is characterized by abnormal drainage of the pulmonary veins into the right heart, whether through connections into the right atrium or into its tributaries. Accordingly, the only mechanism by which oxygenated blood can return to the left heart is through an ASD, which is almost uniformly present with TAPVC.
Unique to this lesion is the absence of a definitive form of palliation. Thus, TAPVC with concomitant obstruction (Fig. 20-17) represents one of the only true surgical emergencies across the entire spectrum of congenital heart surgery.

**Anatomy and Embryology.** The lungs develop from an outpouching of the foregut, and their venous plexus arises as part of the splanchnic venous system. TAPVC arises when the pulmonary vein evagination from the posterior surface of the left atrium fails to fuse with the pulmonary venous plexus surrounding the lung buds. In place of the usual connection to the left atrium, at least one connection of the pulmonary plexus to the splanchnic plexus persists. Accordingly, the pulmonary veins drain to the heart through a systemic vein.

Darling and colleagues classified TAPVC (Fig. 20-18) according to the site or level of connection of the pulmonary veins to the systemic venous system: type I (45%), anomalous connection at the supracardiac level; type II (25%), anomalous connection at the cardiac level; type III (25%), anomalous connection at the infracardiac level; and type IV (5%), anomalous connection at multiple levels. Within each category, further subdivisions can be implemented, depending on whether pulmonary venous obstruction exists. Obstruction to pulmonary venous drainage is a powerful predictor of adverse natural outcome and occurs most frequently with the infracardiac type, especially when the pattern of infracardiac connection prevents the ductus venosus from bypassing the liver.

**Pathophysiology and Diagnosis.** Because both pulmonary and systemic venous blood returns to the right atrium in all forms of TAPVC, a right-to-left intracardiac shunt must be present in order for the afflicted infant to survive. This invariably occurs via a nonrestrictive patent foramen ovale. Because of this obligatory mixing, cyanosis is usually present, and its degree depends on the ratio of pulmonary to systemic blood flow. Decreased pulmonary blood flow is a consequence of pulmonary venous obstruction, the presence of which is unlikely if the right ventricular pressure is less than 85% of systemic pressure.

The child with TAPVC may present with severe cyanosis and respiratory distress, necessitating urgent surgical intervention if a severe degree of pulmonary venous obstruction is present. However, in cases where there is no obstructive component, the clinical picture is usually one of pulmonary overcirculation, hepatomegaly, tachycardia, and tachypnea with feeding. In a child with serious obstruction, arterial blood gas analysis reveals severe hypoxemia (partial pressure of oxygen \([P_{O_2}] < 20 \text{ mmHg}\)), with metabolic acidosis.

Chest radiography (Fig. 20-19) will show normal heart size with generalized pulmonary edema. Two-dimensional echocardiography is very useful in establishing the diagnosis and also can assess ventricular septal position, which may be leftward secondary to small left ventricular volumes, as well as estimate the right ventricular pressure based on the height of the tricuspid regurgitant jet. Echocardiography can usually identify the pulmonary venous connections (types I to IV), and it is rarely necessary to perform other diagnostic tests.

Cardiac catheterization is not recommended in these patients because the osmotic load from the intravenous contrast can exacerbate the degree of pulmonary edema. When cardiac catheterization is performed, equalization of oxygen saturations in all four heart chambers is a hallmark finding in this disease because the mixed blood returned to the right atrium gets distributed throughout the heart.

**Therapy.** Operative correction of TAPVC requires anastomosis of the common pulmonary venous channel to the left atrium, obliteration of the anomalous venous connection, and closure of the ASD.
All types of TAPVC are approached through a median sternotomy, and many surgeons use deep hypothermic circulatory arrest in order to achieve an accurate and widely patent anastomosis. The technique for supracardiac TAPVC includes early division of the vertical vein, retraction of the aorta and the superior vena cava laterally to expose the posterior aspect of the left atrium and the pulmonary venous confluence, and a side-to-side anastomosis between a long, horizontal bialtrial incision and a longitudinal incision within the pulmonary venous confluence. The ASD can then be closed with an autologous pericardial or synthetic patch.

In patients with TAPVC to the coronary sinus without obstruction, a simple unroofing of the coronary sinus can be performed through a single right atriotomy with concomitant closure of the ASD. If pulmonary venous obstruction is present, the repair should include generous resection of roof of the coronary sinus.79

Repair of infracardiac TAPVC entails ligation of the vertical vein at the diaphragm, followed by construction of a proximal, patulous longitudinal venotomy. This repair is usually performed by “rolling” the heart toward the left, thus exposing the left atrium where it usually overlies the descending vertical vein.

As originally described by Lacour-Gayet and colleagues at the Marie-Lannelongue Hospital, Paris, and Coles and colleagues at The Hospital for Sick Children, Toronto, the sutureless technique was developed for patients with anastomotic stenosis occurring after TAPVC repair.80,81 After determining that favorable outcomes were possible using this technique, it is currently used in selected patients upon initial presentation of TAPVC.81 Incisions are made in the venous confluence. Based on the surgeon’s discretion, the incisions are extended into both upper and lower pulmonary veins separately if judged to be important for an unobstructed pathway. An atrioatrial anastomosis is created using the pericardium adjacent to where the pulmonary veins enter the pericardium (Fig. 20-20). This anastomosis avoids direct contact with the incision site in the wall of the pulmonary veins and allows the free egress of blood from the lungs to the left atrium.

Results. Results of TAPVC in infancy have markedly improved in recent years, with an operative mortality of 5% or less in some series.79,82 This improvement is probably multifactorial, mainly as a consequence of early noninvasive diagnosis and aggressive perioperative management. The routine use of echocardiography; improvements in myocardial protection with specific attention to the RV; creation of a large, tension-free anastomosis with maximal use of the venous confluence and atrial tissue; use of a sutureless technique in selected cases; and prevention of pulmonary hypertensive events have likely played a major role in reducing operative mortality. The importance of risk factors for early mortality, such as venous obstruction at presentation, urgency of operative repair, and infradiaphragmatic anatomic type, has been debated.81,83

Bando and colleagues84 made the controversial statement that both preoperative pulmonary venous obstruction and anatomic type had been neutralized as potential risk factors beyond calendar year 1991. Hyde et al82 similarly reported that connection type was not related to outcome. However, a large single-institution report of 377 children with TAPVC by the author from the Hospital for Sick Children in Toronto85 found that, although outcomes had improved over time, patient anatomic factors were still important determinants of both survival and the need for subsequent reoperation. Risk factors for postrepair death were earlier operation year, younger age at repair, cardiac connection type, and postoperative pulmonary venous obstruction. Risk-adjusted estimated 1-year survival for a patient repaired at birth with unfavorable morphology in 2006 was 37% (95% confidence interval [CI], 8%–80%) compared with 96% (95% CI, 91%–99%) for a patient with favorable morphology repaired at age 1 year. Freedom from reoperation was 82% ± 6%.
at 11 years after repair, with increased risk associated with mixed connection and postoperative pulmonary venous obstruction. A study from the Hospital for Sick Children, Toronto, showed a lower incidence of reoperation in the sutureless technique compared to conventional pulmonary venous confluence–left atrial anastomosis. However, there was no statistically significant difference suggesting similar results between the strategies. Although the sutureless technique appears to have favorable outcomes at primary repair for TAPVC, long-term follow-up is necessary to evaluate the occurrence of arrhythmias, such as complete heart block and atrial tachycardia, since an incision on the atrial septum and atrial wall is more invasive compared to the conventional technique.

The most significant postoperative complication of TAPVC repair is pulmonary venous obstruction (Figure 20-21), which occurs 9% to 11% of the time, regardless of the surgical technique employed. Mortality varies between 30% and 45%, and alternative catheter interventions do not offer definitive solutions. Recurrent pulmonary venous obstruction can be localized at the site of the pulmonary venous anastomosis (extrinsic), which usually can be cured with patch enlargement or balloon dilatation, or it may be secondary to endocardial thickening of the pulmonary venous ostia frequently resulting in diffuse pulmonary venous sclerosis (intrinsic), which carries a 66% mortality rate because few good solutions exist. More commonly, postrepair left ventricular dysfunction can occur as the noncompliant LV suddenly is required to handle an increased volume load from redirected pulmonary venous return. This can manifest as an increase in pulmonary artery pressure but is distinguishable from primary pulmonary hypertension (another possible postoperative complication following repair of TAPVC) from the elevated left atrial pressure and LV dysfunction along with echocardiographic evidence of poor LV contractility. In pulmonary hypertension, the left atrial pressure may be low, the LV may appear “underfilled” (by echocardiography), and the RV may appear dilated. In either case, postoperative support for a few days with extracorporeal membrane oxygenation may be lifesaving, and TAPVC should be repaired in centers that have this capacity.

Some investigators have speculated that preoperative pulmonary venous obstruction is associated with increased medial thickness within the pulmonary vasculature, which may predispose these infants to intrinsic pulmonary venous stenosis despite adequate pulmonary venous decompression. The majority of studies demonstrating that preoperative pulmonary venous obstruction is a predictor of subsequent need for reoperation to correct recurrent pulmonary venous obstruction lend credence to this notion.

**Cor Triatriatum**

**Anatomy.** Cor triatriatum is a rare congenital heart defect characterized by the presence of a fibromuscular diaphragm that partitions the left atrium into two chambers: a superior chamber that receives drainage from the pulmonary veins, and an inferior chamber that communicates with the mitral valve and the LV (Fig. 20-22). An ASD frequently exists between the superior chamber and the right atrium, or, more rarely, between the right atrium and the inferior chamber.

**Pathophysiology and Diagnosis.** Cor triatriatum results in obstruction of pulmonary venous return to the left atrium. The degree of obstruction is variable and depends on the size of fenestrations present in the left atrial membrane, the size of the ASD, and the existence of other associated anomalies. If the communication between the superior and inferior chambers is <3 mm, patients usually are symptomatic during the first year of life. The afflicted infant will present with the stigmata of low cardiac output and pulmonary venous hypertension, as well as congestive heart failure and poor feeding.

Physical examination may demonstrate a loud pulmonary S2 sound and a right ventricular heave, as well as jugular venous distention and hepatomegaly. Chest radiography will show cardiomegaly and pulmonary vascular prominence, and the ECG will suggest right ventricular hypertrophy. Two-dimensional echocardiography provides a definitive diagnosis in most cases, with catheterization necessary only when echocardiographic evaluation is equivocal.

**Therapy.** Operative treatment for cor triatriatum is fairly simple. CPB and cardioplegic arrest are used. A right atriotomy usually
allows access to the left atrial membrane through the existing ASD because it is dilated secondary to communication with the pulmonary venous chamber. The membrane is then excised, taking care not to injure the mitral valve or the interatrial septum, and the ASD is closed with a patch. Alternatively, if the right atrium is small, the membrane can be exposed through an incision directly into the superior left atrial chamber, just anterior to the right pulmonary veins. Surgical results are uniformly excellent for this defect, with survival approaching 100%.

The utility of catheter-based intervention for this diagnosis remains controversial, although there have been some reports of successful balloon dilatation.

**Aortopulmonary Window**

**Embryology and Anatomy.** Aortopulmonary window (APW) is a rare congenital lesion, occurring in about 0.2% of patients, characterized by incomplete development of the septum that normally divides the truncus into the aorta and the pulmonary artery.

In the vast majority of cases, APW occurs as a single defect of minimal length, which begins a few millimeters above the semilunar valves on the left lateral wall of the aorta (Fig. 20-23). Coronary artery anomalies, such as aberrant origin of the right or left coronary artery from the main pulmonary artery, are occasionally present.

**Pathophysiology and Diagnosis.** The dominant pathophysiology of APW is that of a large left-to-right shunt with increased pulmonary flow and the early development of congestive heart failure. Like other lesions with left-to-right flow, the magnitude of the shunt is determined by both the size of the defect and the pulmonary vascular resistance.

Infants with APW present with frequent respiratory tract infections, tachypnea with feeding, and failure to thrive. Cyanosis usually is absent because these infants deteriorate prior to the onset of significant pulmonary hypertension. The rapid decline with this defect occurs because shunt flow continues during both phases of the cardiac cycle, which limits systemic perfusion and increases ventricular work.

The diagnosis of APW begins with the physical examination, which may demonstrate a systolic flow murmur, a hyperdynamic precordium, and bounding peripheral pulses. The chest radiograph will show pulmonary overcirculation and cardiomegaly, and the ECG will usually demonstrate either left ventricular hypertrophy or biventricular hypertrophy. Echocardiography (Fig. 20-24) can detect the defect and also provide information about associated anomalies. Retrograde aortography will confirm the diagnosis but is rarely necessary.

**Therapy.** All infants with APW require surgical correction once the diagnosis is made. Repair is undertaken through a median sternotomy and the use of CPB. The pulmonary arteries are occluded once the distal aorta is cannulated, and a transaortic repair using a prosthetic patch for pulmonary artery closure is then carried out. The coronary ostia must be carefully visualized and included on the aortic side of the patch. Alternatively, a two-patch technique can be used, which may eliminate recurrent fistulas from suture line leaks that occasionally occur with the single-patch method.

**Results.** Results are generally excellent, with an operative mortality in most large series of less than 5%.

**Vascular Rings and Pulmonary Artery Slings**

Vascular rings constitute a group of disorders derived from anomalies that result from abnormal development of the aortic arches resulting in compression of the trachea or esophagus. The surgical management of vascular rings dates back to 1945 when Dr. Gross described the surgical management of a kid with double aortic arch. Most children present with symptoms during the first few months of life. Vascular rings can be complete (e.g., double aortic arch, right aortic arch with left ligation) or partial (e.g., innominate artery compression syndrome, pulmonary artery sling).

**Anatomy.** The embryologic basis of vascular rings involves the development of six pairs of aortic arches and the dorsal and ventral aortae. The development of a specific type of vascular ring depends on the deletion or preservation of a specific segment of these structures. The persistence of the right and left fourth arches leads to the development of double aortic arch. Persistence of the fourth right aortic arch and the involution of the left fourth arch leads to the development of a right aortic arch system with various combinations of mirror imaging.

![Figure 20-23](image1.png)

*Figure 20-23.* Cartoon depicting the various types of aortopulmonary window. *(Used with permission from Nicholas Clarke MD.)*

![Figure 20-24](image2.png)

*Figure 20-24.* Echo demonstrating an aortopulmonary window (*`).
SPECIFIC CONSIDERATIONS

branching, aberrant subclavian arteries or with a left-sided ligamentum arteriosum. When the developing left lung captures its blood supply from the right sixth arch caudad to the tracheobronchial tree, it leads to the development of a pulmonary artery sling. The left pulmonary artery arises from the right pulmonary artery and then wraps around the trachea and esophagus forming a “sling.” The pathophysiology of innominate artery compression syndrome is not very well understood.

Pathophysiology and Diagnosis. The symptoms associated with vascular rings include respiratory distress, barking cough, stridor, apnea, and recurrent respiratory tract infections. The diagnosis often requires a high index of suspicion. Minor respiratory tract infections may precipitate serious respiratory distress. The work up includes chest X-rays, echocardiography, bronchoscopy, CT scan (Fig. 20-25), MRI (Fig. 20-26), and, rarely, cardiac catheterization. Chest X-rays show the relationship of the aortic arch to the trachea. Tracheal compression can be better evaluated using lateral films. Unilateral hyperinflation of the lung is sometimes seen and is often associated with a pulmonary artery sling (Fig. 20-27). PA slings (Fig. 20-28) are often associated with complete tracheal rings necessitating a bronchoscopy when this diagnosis is made (Fig. 20-29).

Patients with dysphagia require a barium esophagogram as a part of their work-up (Fig. 20-30).

Treatment. All symptomatic patients should undergo surgery. On close questioning nearly all patients are asymptomatic. The treatment varies depending on the type of vascular ring. A left posterolateral thoracotomy provides good exposure to most types. A right thoracotomy is often used for innominate artery compression syndrome, and a median sternotomy often with cardiopulmonary bypass is used to treat pulmonary artery slings with or without associated complete tracheal rings. The outcomes and results for vascular rings are excellent (Fig. 20-31). Video-assisted thoracoscopic approaches have been developed for the management of these conditions.94-96 The criticism often stated involves retraction of vascular structures into the mediastinum and losing control of the stumps prior to definitive control leading to exsanguination.96

DEFECTS REQUIRING PALLIATION

Tricuspid Atresia

Tricuspid atresia occurs in 2% to 3% of patients with CHD and is characterized by atresia of the tricuspid valve. This results in discontinuity between the right atrium and RV. The RV is generally hypoplastic, and left-heart filling is dependent on an ASD. Tricuspid atresia is the most common form of the single-ventricle complex, indicating that there is functionally only one ventricular chamber.

Anatomy. As mentioned, tricuspid atresia results in a lack of communication between the right atrium and the RV, and in the
majority of patients there is no identifiable valve tissue or remnant. The right atrium is generally enlarged and muscular, with a fibrofatty floor. An unrestrictive ASD is usually present. The LV is often enlarged as it receives both systemic and pulmonary blood flow, but the left AV valve is usually normal.

The RV, however, is usually severely hypoplastic, and there is sometimes a VSD in its trabeculated or infundibular portion. In many cases, the interventricular communication is a site of obstruction to pulmonary blood flow, but obstruction may also occur at the level of the outlet valve or in the subvalvular infundibulum. In most cases, pulmonary blood flow is dependent on the presence of a PDA, and there may be no flow into the pulmonary circulation except for this PDA.

Tricuspid atresia is classified according to the relationship of the great vessels and by the degree of obstruction to pulmonary blood flow. Because of the rarity of tricuspid atresia with transposed great arteries, we will restrict our discussion to tricuspid atresia with normally related great vessels.

Pathophysiology. The main pathophysiology in tricuspid atresia is that of a univentricular heart of left ventricular morphology. That is, the LV must receive systemic blood via the interatrial communication and then distribute it to both the pulmonary circulation and the systemic circulation. Unless there is a VSD (as is found in some cases), pulmonary flow is dependent on the presence of a PDA. As the ductus begins to close shortly after birth, infants become intensely cyanotic. Reestablishing ductal patency (with PGE$_1$) restores pulmonary blood flow and stabilizes patients for surgical intervention. Pulmonary hypertension is unusual in tricuspid atresia. However, occasional patients have a large VSD between the LV and the infundibular portion of the RV (just below the pulmonary valve). If there is no obstruction at the level of this VSD or at the valve, these infants may actually present with heart failure from excessive pulmonary blood flow. Regardless of whether these infants are “ductal-dependent” for pulmonary blood flow or have pulmonary blood flow provided across a VSD, they will be cyanotic since the obligatory right-to-left shunt at the atrial level will provide complete mixing of systemic and pulmonary venous return so that the LV ejects a hypoxic mixture into the aorta.

Diagnosis. The signs and symptoms of tricuspid atresia are dependent on the underlying anatomic variant, but most infants are cyanotic and hypoxic as a result of decreased pulmonary blood flow and the complete mixing at the atrial level. When pulmonary blood flow is provided through a VSD, there may be a prominent systolic murmur. Tricuspid atresia with pulmonary blood flow from a PDA may present with the soft, continuous murmur of a PDA in conjunction with cyanosis.

In the minority of patients with tricuspid atresia, symptoms of congestive heart failure will predominate. This is often related to excessive flow across a VSD. The natural history of the muscular VSDs in these infants is that they will close and the congestive heart failure will dissipate and transform into cyanosis with reduced pulmonary blood flow. Chest radiography will show decreased pulmonary vascularity. The ECG is strongly suggestive because uncharacteristic left axis deviation will be present, due to underdevelopment of the RV. Two-dimensional echocardiography readily confirms the diagnosis and the anatomic subtype. (Fig 20-32)

Treatment. The treatment for tricuspid atresia in the earlier era of palliation was aimed at correcting the defect in the pulmonary circulation. That is, patients with too much pulmonary flow received a pulmonary band, and those with insufficient flow received a systemic-to-pulmonary artery shunt. Systemic-to-pulmonary artery shunts, or Blalock-Taussig (BT) shunts, were first applied to patients with tricuspid atresia in the 1940s and 1950s. Likewise pulmonary artery banding was applied.
to patients with tricuspid atresia and congestive failure in 1957. However, despite the initial relief of either cyanosis or congestive heart failure, long-term mortality was high, as the single ventricle was left unprotected from either volume or pressure overload.

Recognizing the inadequacies of the initial repairs, Glenn described the first successful cavopulmonary anastomosis, an end-to-side right pulmonary artery-to-superior vena cava shunt in 1958, and later modified this to allow flow to both pulmonary arteries. This end-to-side right pulmonary artery-to-superior vena cava anastomosis was known as the bidirectional Glenn, and it is the first stage to final Fontan repair in widespread use today (Fig. 20-33). The Fontan repair was a major advancement in the treatment of CHD, as it essentially bypassed the right heart and allowed separation of the pulmonary and systemic circulations. It was first performed by Fontan in 1971 and consisted of a classic Glenn anastomosis, ASD closure, and direct connection of the right atrium to the proximal end of the left pulmonary artery using an aortic homograft. The main pulmonary artery was ligated, and a homograft valve was inserted into the orifice of the inferior vena cava.

Multiple modifications of this initial repair were performed over the next 20 years. One of the most important was the description by deLeval and colleagues of the creation of an interatrial lateral tunnel that allowed the inferior vena caval blood to be channeled exclusively to the superior vena cava. A total cavopulmonary connection could then be accomplished by dividing the superior vena cava and suturing the superior portion to the upper side of the right pulmonary artery and the inferior end to the augmented undersurface of the right pulmonary artery. Pulmonary flow then occurs passively, in a laminar fashion, driven by the central venous pressure. This repair became known as the modified Fontan operation.

Another important modification, the fenestrated Fontan repair, was introduced in 1988. In this procedure, a residual 20% to 30% right-to-left shunt is either created or left unrepaired at the time of cavopulmonary connection to help sustain systemic output in the face of transient elevations in the pulmonary vascular resistance postoperatively.

Figure 20-31. Bronchoscopy before and after repair of a vascular ring: right arch, left descending aorta, and left ligament.

Figure 20-32. Echo showing tricuspid atresia. The ‘*’ demonstrates the membranous tissue instead of the presence of a tricuspid valve.

Figure 20-33. Angiogram showing a widely patent Glenn. The SVC (‘*’) is seen draining into the central pulmonary artery.
The last notable variation on the original Fontan repair uses an extracardiac prosthetic tube graft (Fig. 20-34), usually 18 to 20 mm in diameter, as the conduit directing inferior vena cava blood to the pulmonary arteries. This technique has the advantages of decreasing atrial geometric alterations by avoiding intra-atrial suture lines and improving flow dynamics in the systemic venous pathway by maximizing laminar flow. Several investigators have shown a decrease in supraventricular arrhythmias, as well as an improvement in ventricular function, which may be secondary to decreased atrial tension and alleviation of chronic elevations in coronary sinus pressure.

One potential disadvantage of the extracardiac Fontan is that it delays performance of the Fontan in order to allow placement of a conduit of sufficient size. Despite these innovative approaches, the current strategy for operative management still relies on the idea of palliation. Patients are approached in a staged manner, to maximize their physiologic state so that they will survive to undergo a Fontan operation. The therapeutic strategy must begin in the neonatal period and should be directed toward reducing the patient’s subsequent risk factors for a Fontan procedure. Accordingly, small systemic pulmonary shunts, which are usually performed through a median sternotomy, should be constructed for palliation of ductus-dependent univentricular physiology. This can easily be replaced with a bidirectional Glenn shunt or hemi-Fontan operation at 6 months of life. In non–ductus-dependent univentricular physiology, the infant can be managed medically until primary construction of a bidirectional cavopulmonary anastomosis becomes feasible. This is possible in the majority of cases because the physiologically elevated pulmonary vascular resistance prevents pulmonary overcirculation during the neonatal period.

The Fontan is usually performed when the child is between 2 and 4 years of age, and it is generally successful if the infant was staged properly, with a protected single ventricle, and there is adequate pulmonary artery growth. The pulmonary vascular resistance should be below 4 Wood units, and the ejection fraction should be more than 45% to ensure success. In patients with high pulmonary artery pressure, fenestration of the atrial baffle may be helpful because their pulmonary vascular resistance may preclude adequate cardiac output postoperatively.

**Results.** Recent reports of the Fontan procedure for tricuspid atresia have been encouraging, with an overall survival of 86% and an operative mortality of 2%. The main complications following repair are atrial arrhythmias, particularly atrial flutter; conduit obstruction requiring reoperation; protein-losing enteropathy; and decreased exercise tolerance.

A prospective multi-institutional study from the Congenital Heart Surgeons Society reported the outcomes of 150 neonates with tricuspid atresia and normally related great vessels. Five-year survival was 86%, and by the age of 2 years, 89% had undergone cavopulmonary anastomosis, and 75% of those surviving cavopulmonary anastomosis underwent Fontan operation within 3 years. Competing risks methodology was used in this study to determine the rates of transition to end-states and their associated determinants (Fig. 20-35). Risk factors for death without cavopulmonary anastomosis in this study included the presence of mitral regurgitation and palliation with systemic-to-pulmonary artery shunts not originating from the innominate artery. Factors associated with decreased transition rate to cavopulmonary anastomosis included patient variables (younger age at admission to a participating institution and noncardiac anomalies) and procedural variables (larger systemic-to-pulmonary arterial shunt diameter and previous palliation).

**Hypoplastic Left Heart Syndrome**

HLHS comprises a wide spectrum of cardiac malformations, including hypoplasia or atresia of the aortic and mitral valves and hypoplasia of the LV and ascending aorta. HLHS has a reported prevalence of 0.2 per 1000 live births and occurs twice as often in boys as in girls. Left untreated, HLHS is invariably fatal and is responsible for 25% of early cardiac deaths in neonates. However, the recent evolution of palliative surgical procedures has dramatically improved the outlook for patients with HLHS, and an improved understanding of anatomic and physiologic alterations has spurred advances in parallel arenas such as intrauterine diagnosis and fetal intervention, echocardiographic imaging, and neonatal critical care.

**Anatomy.** As implied by its name, HLHS involves varying degrees of underdevelopment of left-sided structures (Fig. 20-36), including the LV and the aortic and mitral valves. Thus, HLHS can be classified into four anatomic subtypes based on the valvular morphology: (a) aortic and mitral stenosis; (b) aortic and mitral atresia; (c) aortic atresia and mitral stenosis; and (d) AS and mitral atresia. Aortic atresia tends to be associated with more severe degrees of hypoplasia of the ascending aorta than does AS.

Even in cases without frank aortic atresia, however, the aortic arch is generally hypoplastic and, in severe cases, may even be interrupted. There is an associated coarctation shelf in 80% of patients with HLHS, and the ductus itself is usually quite large, as is the main pulmonary artery.

The segmental pulmonary arteries, however, are small, secondary to reduced intravascular pulmonary blood flow, which is itself a consequence of the left-sided outflow obstruction (Fig. 20-36). The left atrial cavity is generally smaller than normal and is accentuated because of the leftward displacement of the septum primum. There is almost always an interatrial communication via the foramen ovale, which can be large, but more
commonly restricts right-to-left flow. In rare cases, there is no atrial-level communication, which can be lethal for these infants because there is no way for pulmonary venous return to cross over to the RV.

Associated defects can occur with HLHS, and many of them have importance with respect to operative repair. For example, if a VSD is present, the LV can retain its normal size during development even in the presence of mitral atresia. This is because a right-to-left shunt through the defect impels growth of the LV. This introduces the feasibility of biventricular repair for this subset of patients.

Although HLHS undoubtedly results from a complex interplay of developmental errors in the early stages of cardiogenesis, many investigators have hypothesized that the altered blood flow is responsible for the structural underdevelopment that characterizes HLHS. In other words, if the stimulus for normal development of the ascending aorta from the primordial aortic sac is high-pressure systemic blood flow from the LV through the aortic valve, then an atretic or stenotic aortic valve, which impedes flow and leads to only low-pressure diastolic retrograde flow via the ductus, will change the developmental signals and result in hypoplasia of the downstream structures (Fig. 20-37). Normal growth and development of the LV and mitral valve can be secondarily affected, resulting in hypoplasia or atresia of these structures.

**Pathophysiology and Diagnosis.** In HLHS, pulmonary venous blood enters the left atrium, but atrial systole cannot propel blood across the stenotic or atretic mitral valve into the LV. Thus, the blood is shunted across the foramen ovale into the right atrium, where it contributes to volume loading of the RV. The end result is pulmonary venous hypertension from outflow obstruction at the level of the left atrium, as well as pulmonary overcirculation and right ventricular failure. As the pulmonary vascular resistance falls postnatally, the condition is exacerbated because right ventricular output is preferentially directed away from the systemic circulation, resulting in profound underperfusion of the coronary arteries and the vital organs. Closure of the ductus is incompatible with life in these neonates.

Neonates with severe HLHS receive all pulmonary, systemic, and coronary blood flow from the RV. Generally, a child with HLHS will present with respiratory distress within the first day of life, and mild cyanosis may be noted. These infants must be rapidly triaged to a tertiary center, and echocardiography should be performed to confirm the diagnosis. Prostaglandin E$_1$ must be administered to maintain ductal patency, and the

---

**Figure 20-35.** Competing risks depiction of events after diagnosis in 150 patients with tricuspid atresia. All patients began alive and thereafter migrated to one of four mutually exclusive end states (death, bidirectional cavopulmonary anastomosis [BDCPA], single-stage Fontan completion, or remaining alive without BDCPA) at time-dependent rates defined by the underlying hazard functions. At any point in time, the sum of proportions of children in each state is 100%. For example, estimated prevalences after 2 years from diagnosis are as follows: 89% BDCPA, 6% dead without BDCPA, 4% alive without BDCPA, and 1% single-stage Fontan completion. Solid lines represent parametric point estimates; dashed lines enclose 70% confidence intervals; circles with error bars represent nonparametric estimates; numbers in parentheses indicate the estimated proportion of patients in each state at 2 years from diagnosis. *(Reproduced with permission from Karamlou T, Ashburn DA, Caldarone CA, et al: Matching procedure to morphology improves outcomes in neonates with tricuspid atresia, J Thorac Cardiovasc Surg. 2005 Dec;130(6):1503-1510.)*

**Figure 20-36.** Echo In a patient with HLHS. Note the extremely hypoplastic left ventricle (*`).
ventilatory settings must be adjusted to avoid excessive oxygenation and increase carbon dioxide tension. These maneuvers will maintain pulmonary vascular resistance and promote improved systemic perfusion.

Cardiac catheterization should generally be avoided because it is not usually helpful and might result in injury to the ductus and compromised renal function secondary to the osmotic dye load.

**Treatment.** In 1983, Norwood and colleagues described a two-stage palliative surgical procedure for relief of HLHS that was later modified to the currently used three-stage method of palliation. Stage 1 palliation, also known as the modified Norwood procedure (Fig. 20-38), bypasses the LV by creating a single outflow vessel, the neoaorta, which arises from the RV.

The current technique of arch reconstruction involves completion of a connection between the pulmonary root, the native ascending aorta, and a piece of pulmonary homograft used to augment the diminutive native aorta. There are several modifications of this anastomosis, most notably the Damus-Kaye-Stansel (DKS) anastomosis, which involves dividing both the aorta and the pulmonary artery at the sinotubular junction. The proximal aorta is anastomosed to the proximal pulmonary artery, creating a “double-barreled” outlet from the heart. This outlet is anastomosed to the distal aorta, which can be augmented with homograft material if there is an associated coarctation. At the completion of arch reconstruction, a 3.5- or 4-mm shunt is placed from the innominate artery to the right pulmonary artery. The interatrial septum is then widely excised, thereby creating a large interatrial communication and preventing pulmonary venous hypertension.

The DKS connection, as described earlier, might avoid postoperative distortion of the tripartite connection in the neoaorta, and thus decrease the risk of coronary insufficiency. It can be used when the aorta is 4 mm or larger. Unfortunately, in many infants with HLHS, especially if there is aortic atresia, the aorta is diminutive and often less than 2 mm in diameter. The alternate technique available to provide pulmonary blood flow instead of a shunt is a RV-PA conduit often referred to as a “Sano.” It is usually a 5 or 6 mm ribbed Gore-tex graft.

The postoperative management of infants following stage 1 palliation is complex because favorable outcomes depend on establishing a delicate balance between pulmonary and systemic perfusion. Recent literature suggests that these infants require adequate postoperative cardiac output in order to supply both the pulmonary and the systemic circulations and that the use of oxi-metric catheters to monitor mixed venous oxygen saturation (Svo₂) aids clinicians in both the selection of inotropic agents and in ventilatory management. Introduction of a shunt between the RV and the pulmonary artery (Sano shunt) diminishes the diastolic flow created by the modified BT shunt and may augment coronary perfusion, resulting in improved postoperative cardiac function. A recent prospective, randomized, multi-institutional trial sponsored by the National Institutes of Health, the Systemic Ventricle Reconstruction (SVR) trial, compared the outcomes of neonates having either a modified Blalock–Taussig shunt (MBTS) or a Sano shunt. The SVR trial demonstrated that transplantation-free survival 12 months after randomization was higher with the Sano shunt than with the MBTS (74% vs. 64%, P = .01). However, the Sano shunt group had more unintended interventions (P = .003) and complications (P = .002). Right ventricular size and function at the age of 14 months and the rate of nonfatal serious adverse events at the age of 12 months were similar in the two groups. Data collected over a mean (± standard deviation) follow-up period of 32 ± 11 months showed a nonsignificant difference in transplantation-free survival between the two groups (P = .06).

Since the initial SVR publications in 2010, the 3-year and 6-year results have been analyzed. At 3 years, the combined death and cardiac transplantation rates for the RVPAS vs. MBTS groups were 33% vs. 39% (P = .14). When all available data were examined by Kaplan-Meier analysis (mean follow-up 4.4 ± 1.0 years), there was also no difference between groups (log rank P = .11). Overall, there were 100 deaths and 10 transplantations in the MBTS cohort and 86 deaths and 11 transplantations in the RVPAS group. At 6 years, although the point averages continued to reflect a difference favoring the RVPAS (combined death/transplantation rate, 36%) in comparison with the MBTS (41%), the number of subjects was not sufficient to
demonstrate a statistically significant difference between the two groups (log rank *P* = 0.13). Similar to the 3-year results, RVPAS subjects had a higher incidence of any catheter intervention (0.38 vs. 0.23 interventions/patient-year, *P* <0.001), including balloon angioplasty (*P* = 0.014), stent (*P* = 0.009), and coiling (*P* <0.001).113,114 Currently, there remains an ongoing controversy regarding MBTS vs. RV-PA conduit as the source of pulmonary blood flow after the Norwood operation.119,120

Although surgical palliation with the Norwood procedure is still the mainstay of therapy for infants with HLHS, a combined surgical and percutaneous option (hybrid procedure), which consists of bilateral pulmonary artery banding and placement of a ductal stent, has emerged as a promising alternative that obviates the need for CPB in the fragile neonatal period.121,122 The hybrid procedure is performed in a “hybrid suite,” incorporating both advanced fluoroscopic imaging facilities combined with complete operating room capabilities. A 3- or 3.5-mm PTFE tube graft is cut to a width of 3 to 4 mm and used as the bands on the branch pulmonary arteries, placed just distal to the main pulmonary artery. The ductal stent is then positioned in order to cover all ductal tissue and is deployed through a purse-string suture in the main pulmonary artery. A reverse systemic-to-pulmonary shunt is considered in patients with aortic atresia and preductal coarctation to improve coronary perfusion; however, a recent study demonstrated no difference in survival between those with and without the shunt.123 The hybrid procedure can also be used as a bridge to heart transplantation in those infants with severe AV valve regurgitation or otherwise unsuitable single-ventricle anatomy.124

Following stage 1 palliation, the second surgical procedure is the creation of a bidirectional cavopulmonary shunt (Fig. 20-39) or hemi-Fontan, generally at 3 to 6 months of life when the pulmonary vascular resistance has decreased to normal levels. This is the first step in separating the pulmonary and systemic circulations, and it decreases the volume load on the single ventricle. The existing innominate artery-to-pulmonary shunt (or RV-to-pulmonary shunt) or MBTS is eliminated during the same operation.

The third stage of surgical palliation, known as the modified Fontan procedure, completes the separation of the systemic and pulmonary circulations and is performed between 18 months and 3 years of age, or when the patient experiences increased cyanosis (i.e., has outgrown the capacity to perfuse the systemic circulation with adequately oxygenated blood). This has traditionally required a lateral tunnel within the right atrium to direct blood from the inferior vena cava to the pulmonary artery, allowing further relief of the volume load on the RV and providing increased pulmonary blood flow to alleviate cyanosis. More recently, many favor using an extracardiac conduit (e.g., 18- to 20-mm tube graft) to connect the inferior vena cava to the pulmonary artery (Fig. 20-40).

Not all patients with HLHS require this three-stage palliative repair. Some infants afflicted with a milder form of HLHS, recently described as hypoplastic left heart complex (HLHC), have aortic or mitral hypoplasia without intrinsic valve stenosis and antegrade flow in the ascending aorta. In this group, a two-ventricle repair can be achieved with reasonable outcome. Tcherewenkov has published the results with 12 patients with HLHC who underwent biventricular repair at a mean age of 7 days.114 The operative technique consisted of a pulmonary homograft patch aortoplasty of the aortic arch and ascending aorta and closure of the interatrial and interventricular communications. The left heart was capable of sustaining systemic perfusion in 92% of patients, and early mortality was 15.4%. Four patients required reoperations to relieve LVOT obstruction, most commonly between 12 and 39 months following repair. The group from Boston Children’s Hospital has been very aggressive in left ventricular recruitment. These operations still carry a high burden of late death and several reoperations.

Although the Norwood procedure is the most widely performed initial operation for HLHS, transplantation can be used as a first-line therapy and may be preferred when anatomic or physiologic considerations exist that preclude a favorable outcome with palliative repair. Significant tricuspid regurgitation, intractable pulmonary artery hypertension, or progressive right ventricular failure are cases where cardiac replacement may be advantageous. Widespread adaptation of transplantation as

Figure 20-39. Cartoon depicting a bidirectional Glenn. (Used with permission from Kelly Rosso MD.)

Figure 20-40. Extra cardiac fenestrated Fontan. ‘*’ shows the fenestration. (Used with permission from Kelly Rosso MD.)
first-line treatment for HLHS has been limited by improved Norwood survival rates as the operation and pre- and postoperative management of the patient have evolved and by limited organ availability. Organ availability should be considered prior to electing transplantation, as 24% of infants died awaiting transplantation in the largest series to date.\textsuperscript{126,127}

**Results.** Outcomes for HLHS are still significantly worse than those for other complex cardiac defects. However, with improvements in perioperative care and modifications in surgical technique, the survival following the Norwood procedure now exceeds 90% in experienced centers.\textsuperscript{115-120} The outcome for low-birth-weight infants has improved, but low weight still remains a major predictor of adverse survival, especially when accompanied by significant tricuspid valve insufficiency, a restrictive interatrial communication, poor RV function, or extracardiac or chromosomal anomalies.

**DEFECTS THAT MAY BE PALLIATED OR REPAIRED**

**Ebstein’s Anomaly**

**Anatomy.** This is a rare defect, occurring in less than 1% of CHD patients. The predominant maldevelopment in this lesion is the inferior displacement of the tricuspid valve into the RV, although Bove\textsuperscript{128} and others have emphasized the fact that Ebstein’s anomaly is primarily a defect in right ventricular morphology rather than an isolated defect in the tricuspid valve. The anterior leaflet is usually attached in its normal position to the annulus, but the septal and posterior leaflets are displaced toward the ventricle. This effectively divides the RV into two parts: the inlet portion (atrialized RV) and the outlet portion (true or trabeculated RV) (Fig. 20-41). The atrialized RV is usually thin and dilated. Similarly, the tricuspid annulus and the right atrium are extremely dilated, and the tricuspid valve is usually regurgitant with a “sail-like” leaflet (Fig. 20-42). There is commonly an ASD present, which results in a right-to-left shunt at the atrial level. Occasionally, there is true anatomic pulmonary atresia or milder forms of RVOT obstruction.

A Wolff-Parkinson-White (WPW) syndrome (Fig. 20-43) type of accessory pathway with associated preexcitation is present in 15% of patients.\textsuperscript{128}

**Pathophysiology.** Right ventricular dysfunction occurs in patients with Ebstein’s anomaly because of two basic mechanisms: the inflow obstruction at the level of the atrialized ventricle, which produces ineffective RV filling and contractile dysfunction. Inflow obstruction and tricuspid regurgitation, which is exacerbated by progressive annular dilatation, both produce ineffective RV filling. Contractile dysfunction of the RV is a result of a decrease in the number of myocardial fibers, as well as the discordant contraction of the large atrialized portion.

The lack of forward flow at the right ventricular level may lead to physiologic or functional pulmonary atresia, and the infant is dependent on ductal patency for survival. All systemic venous return must be directed through an ASD to the left atrium, where it can be shunted through the ductus for gas exchange. However, the left ventricular function is usually compromised in infants with severe Ebstein’s anomaly as well because the enormous RV and the to-and-fro flow within the atrialized RV prevent adequate intracardiac mixing. Left ventricular function may also be severely compromised in Ebstein’s anomaly because the large RV causes left ventricular compression (Fig. 20-44A,B).

**Diagnosis.** There is a spectrum of clinical presentation in infants with Ebstein’s anomaly that mirrors the anatomic spectrum of this anomaly. Some infants with less severe forms may present with a mild degree of cyanosis, whereas the onset of clinical symptoms in patients surviving childhood is gradual, with the average age of diagnosis in the mid-teens.

However, the infant with severe atrialization and pulmonary stenosis will be both cyanotic and acidotic at birth. The chest radiograph may demonstrate the classic appearance, which

![Figure 20-41](image1.png)

*Figure 20-41. Echo showing a patient with Ebsteins anomaly. Note the inferiorly displaced tricuspid valve (‘*’) and the atrialized portion of the RV (arrow).*

![Figure 20-42](image2.png)

*Figure 20-42. Echo in a patient with severe Ebsteins anomaly showing the large ‘sail like’ anterior leaflet (‘*’).*
consists of a globular “wall-to-wall” heart (Fig. 20-45), similar to that seen with pericardial effusion. The ECG may show right bundle-branch block and right axis deviation. WPW syndrome, as mentioned earlier, is a common finding in these patients. Echocardiography will confirm the diagnosis and provide critical information including tricuspid valvular function, size of the atrialized portion of the RV, degree of pulmonary stenosis, and the atrial size.\textsuperscript{128}

The Great Ormond Street Score (GOSE) (Table 20-1),\textsuperscript{129} which consists of the area of the right atrium plus the area of the atrialized portion of the RV divided by the diastolic area of the remaining cardiac chambers, has been proposed as a useful prognostic tool to stratify neonates with Ebstein’s anomaly. A score of greater than 2 translates into uniformly fatal outcome. Electrophysiology study with radiofrequency ablation is indicated in patients with evidence of WPW syndrome or in children.

**Figure 20-43.** EKG of a newborn with Ebstein’s anomaly and WPW syndrome. Note the pre-excitation (arrow).

**Figure 20-44.** A. Echo (short axis view) of a patient with severe Ebsteins anomaly showing the large RV (‘*’) and small LV (arrow) in diastole. B. Echo (short axis view) of a patient with severe Ebsteins anomaly showing the large RV (‘*’) and small ‘pancaked’ LV (arrow) in systole.
with a history of supraventricular tachycardia, undefined wide-complex tachycardia, or syncope.

**Treatment.** Surgery is indicated for symptomatic infants and for older children and adults with arrhythmias, progressive cyanosis, or New York Heart Association class III or IV. However, the operative repair may be different, depending on the patient’s age, because older children usually are candidates for a biventricular or one-and-a-half ventricle repair, whereas moderate survival has been reported for neonates, using a procedure that converts the anatomy to a single-ventricle physiology, as described by Starnes and coworkers.130

The surgical approach in widespread use today for patients surviving infancy was described by Danielson and colleagues in 1992.128,131 This procedure entails excision of redundant right atrial tissue and patch closure of any associated ASD, plication of the atrialized portion of the ventricle with obliteration of the aneurysmal cavity, posterior tricuspid annuloplasty to narrow the tricuspid annulus, reconstruction of the tricuspid valve if the anterior leaflet is satisfactory, or replacement of the tricuspid valve if necessary.131 If the tricuspid valve is not amenable to reconstruction, valve replacement should be considered. Care must be taken when performing the posterior annuloplasty, or during the conduct of tricuspid valve replacement, to avoid the conduction system, because complete heart block can complicate this procedure. In addition, patients who demonstrated preoperative evidence of preexcitation should undergo electrophysiologic mapping and ablation.

Neonatal Ebstein’s anomaly is a separate entity. Results with surgical correction have been poor, and many neonates are not candidates for operative repair as previously described. Surgical options for the symptomatic neonate include palliative procedures, the one-and-a-half ventricle repair, or conversion to single-ventricle physiology.132 Arguably, the most favorable outcomes in symptomatic neonatal Ebstein’s anomaly or repair in slightly older infants have been achieved using the right ventricular exclusion premise. This technique, known as the “Starnes” procedure (Fig. 20-46),130 uses a fenestrated patch to close the tricuspid valve orifice coupled with systemic-to-pulmonary artery shunt. The patch must be fenestrated to allow decompression of the RV in instances of anatomic pulmonary atresia. Although Knott-Craig and colleagues132 have described tricuspid valve repair for the full spectrum of neonates and infants with excellent short- and mid-term results, these results have not been reproduced in other institutions.133 The one-and-a-half ventricle repair was first described by Billingsly and coworkers as an attempt to achieve a more physiologic “pulsatile” pulmonary circulation in patients with a hypoplastic or dysplastic RV.134 This is accomplished by diverting the superior vena caval blood directly into the pulmonary arterial system by a bidirectional cavopulmonary shunt while recruiting the RV to propel the inferior vena caval blood directly to the pulmonary arteries via the RVOT. Thus, the hemodynamics of the one-and-a-half ventricle repair are characterized by separate systemic and pulmonary circulations in series. The systemic circulation is fully supported by a systemic ventricle, and the pulmonary circulation is supported by both the bidirectional Glenn shunt and the hypoplastic (pulmonary) ventricle. Proponents of this approach report a decreased right atrial pressure and a decrease in inferior vena cava hypertension, which is theorized to be responsible for many of the dreaded complications of the Fontan circulation, including protein-losing encephalopathy, hepatic congestion, atrial arrhythmias, and systemic ventricular failure. In addition, the maintenance of pulsatile pulmonary blood flow, as opposed to continuous laminar flow as in the Fontan circulation, may be advantageous to the pulmonary microcirculation, although it has not been proven in any studies thus far.134,135 Certain criteria, most notably an adequate tricuspid valve Z score, as well as

---

**Table 20-1**

<table>
<thead>
<tr>
<th>The Great Ormond Street Score (GOSE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOSE Score: Area of RA + aRA/Area of RV + LA + LV</td>
</tr>
<tr>
<td><strong>GOSE Score</strong></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

---

**Figure 20-45.** CXR in a newborn with severe Ebstein’s anomaly showing a ‘wall-to-wall’ heart.

**Figure 20-46.** Echo appearance after a Starnes operation. Note the jet of flow across the fenestration in the patch.
the absence of severe pulmonary hypertension or concomitant defects requiring intricate intracardiac repair, should be satisfied prior to electing the one-and-a-half ventricle approach. Patients who do not fulfill these criteria may be approached with a two-ventricle repair and atrial fenestration or a Fontan repair.

In the infant with severe Ebstein’s anomaly, initial stabilization with prostaglandin to maintain ductal patency, mechanical ventilation, and correction of cyanosis is mandatory. Metabolic acidosis, if present from compromised systemic perfusion, must be aggressively treated with afterload reduction. Many of these infants will improve over 1 to 2 weeks as pulmonary vascular resistance falls and they are able to improve antegrade flow into the pulmonary circulation through their abnormal RV and tricuspid valve. When stabilization and medical palliation fail, surgical management remains an option, although its success depends on numerous anatomic factors (e.g., adequacy of the tricuspid valve, RV, and pulmonary outflow tract), and surgery for symptomatic neonates with Ebstein’s anomaly carries a high risk. Knott-Craig and associates reported three cases where two-ventricle repair was undertaken by subtotal closure of the ASD, extensive resection of the right atrium, and vertical plication of the atrialized chamber. Five-year follow-up revealed all patients to be asymptomatic and in sinus rhythm without medications. Recently, they have reported on their 20-year experience with treating 32 such neonates with an overall mortality of 40%. Surgical management of neonates with Ebstein’s anomaly remains challenging. For neonates with Ebstein’s anomaly and anatomical pulmonary atresia, single-ventricle palliation is associated with lower early mortality compared with two-ventricle repair.

Results. In the neonatal period, the most common postoperative problem, whether after a simple palliative procedure such as a BT shunt or following a more extensive procedure such as attempted exclusion of the RV, has been low cardiac output. Supraventricular tachycardia also has been problematic postoperatively. Complete heart block necessitating pacemaker implantation should be uncommon if the techniques described to avoid suturing between the coronary sinus and the tricuspid annulus are used.

There are few published reports of outcomes, due to the rarity of this defect. However, based on the natural history of this condition, which is remarkably benign for the majority of older patients, the outlook should be excellent for patients who have survived the neonatal period.

Transposition of the Great Arteries

Anatomy. Complete transposition is characterized by connection of the atria to their appropriate ventricles with inappropriate ventriculoarterial connections. Thus, the aorta arises anteriorly from the RV, while the pulmonary artery arises posteriorly from the LV. Van Praagh and coworkers introduced the term dextro-transposition of the great arteries (D-TGA) to describe this defect, whereas levo-transposition of the great arteries (L-TGA) describes a form of corrected transposition where there is concomitant AV discordance.

D-TGA requires an obligatory intracardiac mixing of blood, which usually occurs at both the atrial and the ventricular levels or via a patent ductus. Significant coronary anomalies occur frequently in patients with D-TGA. The most common pattern, occurring in 68% of cases, is characterized by the left main coronary artery arising from the leftward coronary sinus, giving rise to the left anterior descending and circumflex arteries. The most common variant is for the circumflex coronary artery to arise as a branch from the right coronary artery instead of from the left coronary artery.

Pathophysiology. D-TGA results in parallel pulmonary and systemic circulations, with patient survival dependent on intracardiac mixing of blood. After birth, both ventricles are relatively noncompliant, and thus, infants initially have higher pulmonary flow due to the decreased downstream resistance. This causes left atrial enlargement and a left-to-right shunt via the patent foramen ovale.

Postnatally, the LV does not hypertrophy because it is not subjected to systemic afterload. The lack of normal extraterine left ventricular maturation has important implications for the timing of surgical repair because the LV must be converted to the systemic ventricle and be able to function against systemic vascular resistance. If complete repair is done within the first few weeks of life, the LV usually adapts easily to systemic resistance since it is conditioned to high intrathoracic pulmonary vascular resistance. After a few weeks of life, the LV that is conditioned to the decrease in pulmonary resistance that occurs when the lungs inflate after birth may have difficulty adapting to systemic vascular resistance without preoperative preparation or postoperative support. Novel techniques of LV “preparation” using a pulmonary arterial band have been used in cases where complete repair has been delayed (Fig. 20-47A,B).

Clinical Manifestations and Diagnosis. Infants with D-TGA and an intact ventricular septum are usually cyanotic at birth, with an arterial Po2 between 25 and 40 mmHg. If ductal patency is not maintained, deterioration will be rapid with ensuing metabolic acidosis and death. Conversely, those infants with a coexisting VSD may be only mildly hypoxemic and may come to medical attention after 2 to 3 weeks, when the falling pulmonary vascular resistance leads to symptoms of congestive heart failure.

The ECG will reveal right ventricular hypertrophy, and the chest radiograph will reveal the classic egg-shaped configuration. Definitive diagnosis is made by echocardiography, which reliably demonstrates ventriculoarterial discordance and any associated lesions. Cardiac catheterization is rarely necessary, except in infants requiring surgery after the neonatal period, to assess the suitability of the LV to support the systemic circulation. Limited catheterization, however, is useful for performance of atrial septostomy in neonates with inadequate intracardiac mixing.

Surgical Repair. Blalock and Hanlon introduced the first operative intervention for D-TGA with the creation of an atrial septectomy to enhance intracardiac mixing. This initial procedure was feasible in the pre-CPB era, but carried a high mortality rate. Later, Rashkind and Causo developed a catheter-based balloon septostomy, which largely obviated the need for open septectomy.

These early palliative maneuvers, however, met with limited success, and it was not until the late 1950s, when Senning and Mustard developed the first “atrial repair,” that outcomes improved. The Senning operation consisted of rerouting venous flow at the atrial level by incising and realigning the atrial septum over the pulmonary veins and using the right atrial free wall to create a pulmonary venous baffle (Fig. 20-48).

Although the Mustard repair (Fig. 20-49) was similar, it made use of either autologous pericardium or synthetic material to create the interatrial baffle. These atrial switch procedures
resulted in a physiologic correction, but not an anatomic one, as the systemic circulation is still based on the RV. Still, survival rose to 95% in most centers by using an early balloon septostomy followed by an atrial switch procedure at 3 to 8 months of age.\(^{141,142}\)

Despite the improved early survival rates, long-term problems, such as superior vena cava or pulmonary venous obstruction, baffle leak, arrhythmias, tricuspid valve regurgitation, and right ventricular failure, prompted the development of the arterial switch procedure by Jatene in 1975.\(^{143}\) The arterial switch procedure involves the division of the aorta and the pulmonary artery, posterior translocation of the aorta (LeCompte maneuver), mobilization of the coronary arteries, placement of a pantaloon-shaped pericardial patch, and proper alignment of the coronary arteries on the neoaorta (Fig. 20-50).

The most important consideration is the timing of surgical repair because arterial switch should be performed within 2 weeks after birth, before the LV loses its ability to pump against systemic afterload. In patients presenting later than 2 weeks, the LV can be retrained with preliminary pulmonary artery banding.
and aortopulmonary shunt followed by definitive repair. Alternatively, the unprepared LV can be supported following arterial switch with a mechanical assist device for a few days while it recovers ability to manage systemic pressures. Echocardiography can be used to assess left ventricular performance and guide operative planning in these circumstances.

The subset of patients who present with D-TGA complicated by LVOT obstruction and VSD may not be suitable for an arterial switch operation. The Rastelli operation, first performed in 1968, uses placement of an intracardiac baffle to direct left ventricular blood to the aorta and an extracardiac valved conduit to establish continuity between the RV and the pulmonary artery, which has led to successful outcomes in these complex patients.144

Results. For patients with D-TGA, intact ventricular septum, and VSD, the arterial switch operation provides excellent long-term results with a mortality rate of less than 5%. Operative risk is increased when unfavorable coronary anatomic configurations are present or when augmentation of the aortic arch is required. The most common complication is supravalvular pulmonary stenosis, occurring 10% of the time, which may require ballooning or reoperation (Fig. 20-51).145

Results of the Rastelli operation have improved substantially, with an early mortality rate of 5%.146 Late mortality rate results were less favorable because conduit failure requiring reoperation, pacemaker insertion, or relief of LVOT obstruction was frequent.

Figure 20-50. The Arterial Switch Operation. A. The maneuver of Lecompte (positioning the pulmonary artery anterior to the aorta) is shown with aortic cross-clamp repositioning to retract the pulmonary artery during the neoaortic reconstruction. A and B. After the coronary patches are rotated for an optimal lie, they are sutured to the linearly incised sinuses of Valsalva at the old pulmonary artery (neoaorta) (C). (Reproduced with permission from Mavroudis C, Backer CL: Arterial Switch. Cardiac Surgery: State of the Art Review. Vol. 5, no. 1. Philadelphia, PA: Hanley & Belfus; 1991.)

Figure 20-51. Angiographic appearance of the pulmonary arteries before and after balloon dilation. The RV pressures dropped from “systemic” to “1/2 systemic” after dilation.
Double-Outlet Right Ventricle

Anatomy. Double-outlet RV (DORV) accounts for 5% of CHD and exists when both the aorta and pulmonary artery arise wholly, or in large part, from the RV (Fig. 20-52). DORV encompasses a spectrum of malformations because the incomplete shift of the aorta toward the LV is often associated with other abnormalities of cardiac development, such as ventricular looping and infundibular-truncal spiraling. The vast majority of hearts exhibiting DORV have a concomitant VSD, which varies in its size and spatial association with the great vessels. The VSD is usually nonrestrictive and represents the only outflow for the LV; its location relative to the great vessels dictates the dominant physiology of DORV, which can be analogous to that of a large isolated VSD, tetralogy of Fallot, or D-TGA. In 1972, Lev et al suggested considering DORV as a spectrum of hearts that “pass imperceptibly from tetralogy with VSD with overriding aorta into double-outlet right ventricle with subaortic VSD.” Thus, Lev and colleagues described a classification scheme for DORV based on the “commitment” of the VSD to either or both great arteries. The VSD can be subaortic, doubly committed, noncommitted, or subpulmonic.

The subaortic type is the most common (47%) and occurs when the VSD is located directly beneath the aortic annulus. Doubly committed VSD (4%) is present when the VSD lies beneath both the aorta and the pulmonary artery, which are usually side-by-side in this lesion. The noncommitted VSD (26%) exists when the VSD is remote from the great vessels. The subset of DORV hearts with the VSD located beneath the pulmonary valve also are classified as the Taussig–Bing syndrome (Fig. 20-53). This occurs in 23% of cases of DORV with VSD, and it occurs when the aorta rotates more anteriorly, with the pulmonary artery rotated more posteriorly.

Clinical Manifestations and Diagnosis. Patients with DORV typically present with one of the following three scenarios: (a) those with doubly committed or subaortic VSD present with congestive heart failure and a high propensity for pulmonary hypertension, much like infants with a large single VSD; (b) those with a subaortic VSD and pulmonary stenosis present with cyanosis and hypoxia, much like infants with tetralogy of Fallot; and (c) those with subpulmonic VSD present with cyanosis, much like those with D-TGA, because streaming directs desaturated systemic venous blood to the aorta and oxygenated blood to the pulmonary artery. Thus, the three critical factors influencing the clinical presentation and subsequent management of infants with DORV are the size and location of the VSD, the presence or absence of important RVOT obstruction, and the presence of other anomalies (especially associated hypoplasia of left-sided structures sometimes seen with subpulmonary VSD).

Echocardiography is the mainstay of diagnosis and can also provide valuable information regarding the feasibility of biventricular repair. Specific anatomic questions that should be resolved to assist in surgical planning in addition to those mentioned earlier include the coronary anatomy (presence of a conal branch or left anterior descending from the right coronary coursing across the conus), the presence of additional muscular VSDs remote from either great vessel, and the distance between the tricuspid and pulmonary valve. Cardiac catheterization is rarely necessary in neonates or infants, except to determine the degree of pulmonary hypertension and to determine the effects of previous palliative procedures on the pulmonary arterial anatomy.

Therapy. The goals of corrective surgery are to relieve pulmonary stenosis, to provide separate and unobstructed outflow pathways from each ventricle to the correct great vessel, and to achieve separation of the systemic and pulmonary circulations.

Double-Outlet Right Ventricle With Noncommitted Ventricular Septal Defect

The repair of hearts with DORV and noncommitted VSD can be accomplished by constructing an intraventricular tunnel connecting the VSD to the aorta, closing the pulmonary artery, and placing a valved extracardiac conduit from the RV to the pulmonary artery. In patients without pulmonary stenosis who have intractable congestive failure, a pulmonary artery band can be placed in the first 6 months to control pulmonary artery...
overcirculation and prevent the development of pulmonary hypertension.

Infants with pulmonary stenosis can be managed with a systemic-to-pulmonary shunt followed by biventricular repair as described by Belli and colleagues in 1999, or with a modified Fontan. There is no consensus on the timing of repair, but recent literature suggests that repair within the first 6 months is associated with better outcome. However, in cases where an extracardiac-valved conduit is necessary, it is better to delay definitive repair until the child is 2 to 3 years of age because this allows placement of a larger conduit and possibly reduces the number of future obligatory conduit replacements.

Double-Outlet Right Ventricle With Subaortic or Doubly Committed Ventricular Septal Defect Without Pulmonary Stenosis

This group of patients can be treated by creating an intracardiac baffle that directs blood from the LV into the aorta. Enlargement of the VSD may be necessary to allow ample room for the baffle; this should be done anterosuperiorly to avoid injury to the conduction system that normally lies inferoposteriorly along the border of the VSD. In addition, other important considerations in constructing the LV outflow tunnel include the prominence of the conal septum, the attachments of the tricuspid valve to the conal septum, and the distance between the tricuspid and pulmonary valves. In some instances, unfavorable anatomy may preclude placement of an adequate intracardiac baffle, necessitating single ventricle repair.

Double-Outlet Right Ventricle With Subaortic or Doubly Committed Ventricular Septal Defect With Pulmonary Stenosis

Repair of this defect is similar to the above except that concomitant RVOT reconstruction must be performed in addition to the intracardiac tunnel. The RVOT augmentation can be accomplished with the placement of a transannular patch or with placement of an extracardiac-valved conduit when an anomalous left anterior descending artery precludes use of a patch.

Taussig–Bing Syndrome Without Pulmonary Stenosis

These infants are best treated with a balloon septostomy during the neonatal period to improve mixing, followed by VSD closure baffling LV egress to the pulmonary artery and an arterial switch operation. The Kawashima procedure, in which an intraventricular tunnel is used to baffle LV egress directly to the aorta, may alternatively be used when the aorta is more posterior or when there is associated pulmonary stenosis.

Taussig–Bing Syndrome With Pulmonary Stenosis

This defect may be treated with a variety of techniques, depending on the specific anatomic details and the expertise of the treatment team. A Rastelli-type repair, which involves construction of an intraventricular tunnel through the existing VSD that connects the LV to both great vessels, followed by division of the pulmonary artery at its origin and insertion of a valved conduit from the RV to the distal pulmonary artery, can be performed. Alternatively, a Yasui procedure, which involves baffling the VSD to the pulmonary artery and creation of a DKS anastomosis between the pulmonary artery and the aorta with patch augmentation, can be accomplished concomitant with placement of an RV pulmonary artery conduit.

Results. The results of DORV repairs are generally favorable, especially for the tetralogy-type DORV with subaortic VSD. However, more complex types of DORV, including noncommitted VSD and Taussig–Bing type, still carry important morbidity and mortality. Furthermore, repeated interventions for RVOT reconstruction or staged operations for patients triaged to single-ventricle pathways pose late hazards for patients surviving initial repair. A single-institution series evaluated 393 patients with DORV. The authors found that the need for reintervention approached 37% at 15 years following repair. Arterial switch operation, as opposed to Rastelli-type repair, was associated with an increased risk of early postrepair mortality, but mitigated against the risk of late death. Patients with hypoplastic left-sided structures and a nonsubaortic VSD may fare better with a single-ventricle repair.

Tetralogy of Fallot

Anatomy. The original description of tetralogy of Fallot (TOF) by Etienne Louis Fallot, as the name implies, included four abnormalities: a large perimembranous VSD adjacent to the tricuspid valve; an overriding aorta; a variable degree of RVOT obstruction, which might include hypoplasia and dysplasia of the pulmonary valve as well as obstruction at the subvalvar and pulmonary artery level; and right ventricular hypertrophy. More recently, the Van Praagh et al pointed out that TOF could be more correctly termed monology of Fallot, since the four components are explained by the malposition of the infundibular septum. When the infundibular septum is displaced anteriorly and leftward, the RVOT is narrowed and its anterior displacement results in failure of fusion of the ventricular septum between the arms of the trabeculo-septo-marginalis (Fig. 20-54).

The morphology of TOF is markedly heterogeneous and includes an absent pulmonary valve, concomitant AV septal defects, and pulmonary atresia with major aortopulmonary collaterals. The present discussion will focus only on the so-called classic presentation of TOF without coexisting intracardiac defects.

Anomalous coronary artery patterns, related to either origin or distribution, have been described in TOF. However, the most surgically important coronary anomaly occurs when

![Figure 20-54. Tetrology of Fallot. (Used with permission from Kelly Rosso MD.)](image-url)
the left anterior descending artery arises as a branch of the right coronary artery. This occurs in approximately 3% of cases of TOF and may preclude placement of a transannular patch, as the left anterior descending coronary artery crosses the RVOT at varying distances from the pulmonary valve annulus.\textsuperscript{158}

**Pathophysiology and Clinical Presentation.** The initial presentation of a child afflicted with TOF depends on the degree of RVOT obstruction. Children with cyanosis at birth usually have severe pulmonary annular hypoplasia with concomitant hypoplasia of the peripheral pulmonary arteries. Most children, however, present with mild cyanosis at birth, which then progresses as the right ventricular hypertrophy further compromises the RVOT. Cyanosis usually becomes significant within the first 6 to 12 months of life, and the child may develop characteristic “tet” spells, which are periods of extreme hypoxemia. These spells are characterized by decreased pulmonary blood flow and an increase in systemic blood flow. They can be triggered by any stimulus that decreases systemic vascular resistance, such as fever, agitation, or vigorous physical activity. Cyanotic spells increase in severity and frequency as the child grows, and older patients with uncorrected TOF may often squat, which increases peripheral vascular resistance and relieves the cyanosis.

Evaluation in the older patient with TOF may demonstrate clubbing, polycythemia, hemoptysis, or brain abscesses. Chest radiography will demonstrate a boot-shaped heart (Fig. 20-55), and EKG will show the normal pattern of right ventricular hypertrophy. Echocardiography confirms the diagnosis because it demonstrates the position and nature of the VSD, defines the character of the RVOT obstruction, and often visualizes the branch pulmonary arteries and the proximal coronary arteries. Cardiac catheterization is rarely necessary and is actually risky in TOF since it can create spasm of the RVOT muscle and result in a hypercyanotic episode (tet spell). Occasionally, aortography (Fig. 20-56) is necessary to delineate the coronary artery anatomy.

**Treatment.** John Deanfield\textsuperscript{160} stated “…long follow-up inevitably means surgery in an earlier era: More recent surgery, at a younger age, with better preoperative, operative, and postoperative care, will improve long-term results. Data from the former (earlier) era will be overly pessimistic.” This statement is particularly pertinent as surgical correction of TOF has evolved from a staged approach of antecedent palliation in infancy followed by intracardiac repair to primary repair during the first few months of life without prior palliative surgery.

However, systemic-to-pulmonary shunts, generally an MBTS, may still be preferred with an unstable neonate younger than 3 months of age, when an extracardiac conduit is required because of an anomalous left anterior descending coronary artery, or when pulmonary atresia, significant branch pulmonary artery hypoplasia, or severe noncardiac anomalies coexist with TOF.

Traditionally, TOF was repaired through a right ventriculotomy, providing excellent exposure for closure of the VSD and relief of the RVOT obstruction, but concerns that the resultant scar would significantly impair right ventricular function or lead to lethal arrhythmias led to the development of a transatrial approach. Transatrial repair, except in cases when the presence of diffuse RVOT hypoplasia requires insertion of a transannular patch, is now being increasingly advocated by many, although its superiority has not been conclusively demonstrated.\textsuperscript{161}

The operative technique involves the use of CPB. All existing systemic-to-pulmonary arterial shunts, as well as the ductus arteriosus, are ligated. A right atriotomy is then made, and the anatomy of the VSD and the RVOT are assessed by retracting the tricuspid valve. The outflow tract obstruction is relieved by resecting the offending portion of the infundibular septum as well as any muscle trabeculations. If necessary, a pulmonary valvotomy or, alternatively, a longitudinal incision in the main pulmonary artery can be performed to improve exposure. The diameter of the pulmonary valve annulus is assessed by inserting Hegar dilators across the outflow tract; if the pulmonary artery/aorta diameter is less than 0.5, or the estimated RV/LV pressure is greater than 0.7, or the size of the pulmonary valve is less than a Z score of −2.5, a transannular patch is inserted. Patch closure of the VSD is then accomplished, taking...
care when placing sutures along the posteroinferior portion to avoid the conduction system.

**Results.** Operative mortality for primary repair of TOF in infancy is less than 5% in most series. Previously reported risk factors such as transannular patch insertion or younger age at time of repair have been eliminated secondary to improved intraoperative and postoperative care. According to the Society of Thoracic Surgeons Congenital Heart Surgery Database, discharge mortality from 3059 operations from 2002 to 2007 was 7.5% for initial palliation, 1.3% for primary repair, and 0.9% for staged repair, indicating similar outcomes for patients getting primary repair compared to staged repair. Nevertheless, for neonatal repair, discharge mortality increased to 6.2% with palliation and 7.8% with primary repair. This may be partly explained by a higher chance of postoperative complications in neonates.

A major complication of repaired TOF is the development of pulmonary insufficiency, which subjects the RV to the adverse effects of acute and chronic volume overload. This is especially problematic if residual lesions such as a VSD or peripheral pulmonary stenosis exist. Pulmonary valve regurgitation after repair of TOF is relatively well tolerated in the short term, partly because the hypertrophied RV usually adapts to the altered hemodynamic load. The detrimental effects of chronic pulmonary valve regurgitation are, however, numerous, and include progressive right ventricular dilatation and failure, tricuspid valve regurgitation, exercise intolerance, arrhythmia, and sudden death. Mechanoelectrical interaction, by which a dilated RV provides the substrate for electrical instability, might underlie the propensity toward ventricular arrhythmia. In support of this contention, Gatzoulis and colleagues found that the risk of symptomatic arrhythmia was high in patients with marked right ventricular enlargement and QRS prolongation on resting ECG of more than 180 ms. Karamlou et al have shown that similar structural and hemodynamic abnormalities, including a larger right atrial volume and right ventricular chamber size, are also related to atrial arrhythmias in patients following TOF repair. We found that prolongation of the QRS duration beyond a threshold of 160 ms increased the risk of atrial arrhythmias. Together, these data show that a similar mechanism could be responsible for both atrial and ventricular arrhythmias after repair in TOF patients.

When significant deterioration of ventricular function occurs, insertion of a pulmonary valve may be required, although this is rarely necessary in infants. Unfortunately, there are no universal criteria establishing the timing of pulmonary valve replacement. The current criteria for pulmonary valve replacement are the presence of two of the following criteria: RVEDD index >160 ml/m2, RVEDI >70 ml/m2, LVEDV index >65 ml/m2, RVEF <45%, RVOT aneurysm, and clinical symptoms or signs, including syncope or VT. PVR can be achieved with minimal morbidity and mortality.

The alternative to surgical PVR is percutaneous pulmonary valve implantation. The Melody valve system (Fig. 20-57) is the most popular of such systems. Following risk adjustment, no significant differences were observed between surgical or transcatheter PVR. However, transcatheter PVR was associated with a shorter hospitalization. Hospitalization costs are similar for both procedures.

Arrhythmias are potentially the most serious late complication following TOF repair. In a multicenter cohort of 793 patients studied by Gatzoulis et al, a steady increase was documented in the prevalence of ventricular and atrial tachyarrhythmia and sudden cardiac death in the first 5 to 10 years after intracardiac repair. Clinical events were reported in 12% of patients at 35 years after repair. Prevalence of atrial arrhythmias from other studies, however, ranges from 1% to 11%, which is a reflection of the strong time dependence of arrhythmia onset.

Underlying causes of arrhythmia following repair are complex and multifactorial, resulting in poorly defined optimum screening and treatment algorithms. Older repair age has been associated with an increased frequency of both atrial and ventricular arrhythmias. Impaired ventricular function secondary to a protracted period of cyanosis before repair might contribute to the propensity for arrhythmia in older patients.

**Ventricular Septal Defect**

**Anatomy.** VSD refers to a hole between the LV and RV. These defects are common, comprising 20% to 30% of all cases of CHD, and may occur as an isolated lesion or as part of a more
complex malformation. VSDs vary in size from 3 to 4 mm to more than 3 cm and are classified into four types based on their location in the ventricular septum: perimembranous (or paramembranous, conoventricular), AV canal (inlet), outlet or supracristal, and muscular (Fig. 20-58).

Perimembranous VSDs are the most common type requiring surgical intervention, comprising approximately 80% of cases. These defects involve the membranous septum and include the malalignment defects seen in tetralogy of Fallot. In rare instances, the anterior and septal leaflets of the tricuspid valve adhere to the edges of the perimembranous defect, forming a channel between the LV and the right atrium. These defects result in a large left-to-right shunt due to the large pressure differential between the two chambers.

AV canal defects, also known as inlet defects, occur when part or all of the septum of the AV canal is absent. The VSD lies beneath the tricuspid valve and is limited upstream by the tricuspid annulus, without intervening muscle.

The supracristal or outlet VSD results from a defect within the conal septum. Characteristically, these defects are limited upstream by the pulmonary valve and are otherwise surrounded by the muscle of the infundibular septum.

Muscular VSDs are the most common type and may lie in four locations: anterior, midventricular, posterior, or apical. These are surrounded by muscle and can occur anywhere along the trabecular portion of the septum. The rare “Swiss-cheese” type of muscular VSD consists of multiple communications between the RV and LV, complicating operative repair.

Pathophysiology and Clinical Presentation. The size of the VSD determines the initial pathophysiology of the disease. Large VSDs are classified as nonrestrictive and are at least equal in diameter to the aortic annulus. These defects allow free flow of blood from the LV to the RV, elevating right ventricular pressures to the same level as systemic pressure.

Consequently, the pulmonary-to-systemic flow ratio (Qp to Qs) is inversely dependent on the ratio of pulmonary vascular resistance to systemic vascular resistance. Nonrestrictive VSDs produce a large increase in pulmonary blood flow, and the afflicted infant will present with symptoms of congestive heart failure. However, if untreated, these defects will cause pulmonary hypertension with a corresponding increase in pulmonary vascular resistance. This will lead to a reversal of flow (a right-to-left shunt), which is known as Eisenmenger’s syndrome.

Small restrictive VSDs offer significant resistance to the passage of blood across the defect, and therefore right ventricular pressure is either normal or only minimally elevated and the ratio of Qp to Qs rarely exceeds 1.5. These defects are generally asymptomatic because there are few physiologic consequences. However, there is a long-term risk of endocarditis because endocardial damage from the jet of blood through the defect may serve as a possible nidus for colonization (Fig. 20-59A,B).

Diagnosis. The child with a large VSD will present with severe congestive heart failure and frequent respiratory tract infections. Children with Eisenmenger’s syndrome may be deceptively asymptomatic until frank cyanosis develops.

The chest radiograph will show cardiomegaly and pulmonary overcirculation, and the ECG will show signs of left ventricular or biventricular hypertrophy. Echocardiography provides definitive diagnosis and can estimate the degree of shunting as well as pulmonary arterial pressures. Cardiac catheterization has
SPECIFIC CONSIDERATIONS

PART II

largely been supplanted by echocardiography, except in older children where measurement of pulmonary resistance is necessary prior to recommending closure of the defect.

Treatment. VSDs may close or narrow spontaneously, and the probability of closure is inversely related to the age at which the defect is observed. Thus, infants at 1 month of age have an 80% incidence of spontaneous closure, whereas a child at 12 months of age has only a 25% chance of closure.170 This has an important impact on operative decision-making because a small or moderate-size VSD may be observed for a period of time in the absence of symptoms. Large defects and those in severely symptomatic neonates should be repaired during infancy to relieve symptoms and because irreversible changes in pulmonary vascular resistance may develop during the first year of life.

Repair of isolated VSDs requires the use of CPB with moderate hypothermia and cardioplegic arrest. The right atrial approach (Fig. 20-60) is preferable for most defects, except apical muscular defects, which often require a right ventriculotomy for adequate exposure. Supracristal defects may alternatively be exposed via a pulmonary arteriotomy or through an incision in the RV immediately beneath the pulmonary valve (Fig. 20-61). Regardless of the type of defect present, a right atrial approach can be used initially to inspect the anatomy, as this may be abandoned should it offer inadequate exposure for repair. After careful inspection of the heart for any associated malformations, a patch repair is employed, taking care to avoid the conduction system. Routine use of intraoperative transesophageal echocardiography should be used to assess for any residual defect.

Successful percutaneous device closure of VSDs using the Amplatzer device has been described.152 The device has demonstrated a 100% closure rate in a small series of patients with isolated or residual VSDs, or as a collaborative treatment strategy for the VSD component in more complex congenital lesions. Proponents of device closure argue that its use can decrease the complexity of surgical repair, avoid reoperation for a small residual lesion, or avoid the need for a ventriculotomy. The use of devices to close paramembranous defects can cause heart block because the defect is in close association to the conduction system (Fig. 20-62).171 The procedure can be performed percutaneously or through the percutaneous approach. Embolization of the device is an added risk.

Multiple or “Swiss-cheese” VSDs represent a special case, and many cannot be repaired during infancy. In patients in whom definitive VSD closure cannot be accomplished, temporary placement of a pulmonary artery band can be employed to control pulmonary flow. This allows time for spontaneous closure of many of the smaller defects, thus simplifying surgical repair.172

Some centers, however, have advocated early definitive repair of the Swiss-cheese septum, by using oversize patches, fibrin glue, and combined intraoperative device closure, as well as techniques to complete the repair transatrially.173

Results. Even in very small infants, closure of VSDs can be safely performed with hospital mortality near 0%. The main risk factor remains the presence of other associated lesions, especially when present in symptomatic neonates with large VSDs.

Figure 20-60. Intra-op picture during a VSD closure performed by interrupted suture technique with patch closure.

Figure 20-61. Echocardiographic appearance of a supracristal VSD (arrow). Note its location just beneath the pulmonary valve (‘*’).
Atrioventricular Canal Defects

Anatomy. AV canal defects result from failure of fusion of the endocardial cushions in the central portion of the heart, causing a lesion that involves the atrial and the ventricular septum, as well as the anterior mitral and septal tricuspid valve leaflets. Defects involving primarily the atrial septum are known as partial AV canal defects and frequently occur in conjunction with a cleft anterior mitral leaflet. Complete AV canal defects have a combined deficiency of the atrial and ventricular septum associated with a common AV orifice rather than separate tricuspid and mitral valves. The common AV valve generally has five leaflets, three lateral (free wall) and two bridging (septal) leaflets. The defect in the ventricular septum can lie either between the two bridging leaflets or beneath them. The relationship between the septal defect and the anterior bridging leaflet forms the basis of the Rastelli classification for complete AV canal defects (Fig. 20-63).174,175

Pathophysiology and Diagnosis. Partial AV canal defects, in the absence of AV valvular regurgitation, frequently resemble isolated ASDs. Left-to-right shunting predominates as long as pulmonary vascular resistance remains low. However, 40% of patients with partial AV canal defects have moderate-to-severe valve incompetence, and progressive heart failure occurs early in this patient population.175 Complete AV canal defects produce more severe pathophysiologic changes because the large intracardiac communication and significant AV valve regurgitation contribute to ventricular volume loading and pulmonary hypertension. Children with complete AV canal defects develop signs of congestive heart failure within the first few months of life.

Treatment. The management of patients with AV canal defects can be especially challenging. Timing of operation is individualized. Patients with partial defects can be electively repaired between 2 and 5 years of age, whereas complete AV canal defects should be repaired within the first year of life to prevent irreversible changes in the pulmonary circulation. Complete repair in infancy should be accomplished, with palliative procedures such as pulmonary artery banding reserved for only those infants with other complex lesions or who are too ill to tolerate CPB.

The operative technique requires the use of either continuous hypothermic CPB or, for small infants, deep hypothermic circulatory arrest. The heart is initially approached through an oblique right atriotomy, and the anatomy is carefully observed. In the case of a partial AV canal, the cleft in the mitral valve is repaired with interrupted sutures and the ASD is closed with a pericardial patch. Complete AV canal defects are repaired by patch closure of the VSD, separating the common AV valve into tricuspid and mitral components and suspending the neovalves from the top of the VSD patch and closing the ASD.

Results. Partial AV canal defects have an excellent outcome, with a mortality rate of 0% to 2% in most series.175 Complete AV canal defects are associated with an operative mortality of 3% to 4%.176

The most frequently encountered postoperative problems are complete heart block (1%–2%), right bundle-branch block (22%), arrhythmias (11%), RVOT obstruction (11%), and severe mitral regurgitation (13%–24%).175 The increasing use of intraoperative transesophageal echocardiography may positively

Figure 20-62. Intraoperative picture at the time of removal of a percutaneously placed VSD device causing severe TR and complete heart block. Note the close association of the device to the tricuspid valve leaflet (arrow) and cordae.

Figure 20-63. Rastelli classification of complete AVSD. (Used with permission from Kelly Rosso MD.)

Figure 20-64. Echo of an infant with complete AVSD. Note the prominent absence of the ‘crux’ (‘∗’) of the heart in this defect.
influence outcomes, as the adequacy of repair can be assessed and treated without need for subsequent reoperation.\textsuperscript{174-175}

**Interrupted Aortic Arch**

**Anatomy.** Interrupted aortic arch (IAA) is a rare defect, comprising approximately 1\% of all cases of CHD.\textsuperscript{177} It is defined as an absence of luminal continuity between the ascending and descending aorta and does not occur as an isolated defect in most cases because a VSD or PDA is usually present. IAA is classified based on the location of the interruption (Fig. 20-65 to Fig. 20-67).

**Clinical Manifestations and Diagnosis.** Infants with IAA have ductal-dependent systemic blood flow and will develop profound metabolic acidosis and hemodynamic collapse upon ductal closure. In the rare instance of failed ductal closure, the diagnosis may be missed during infancy, and the child will present with symptoms of congestive heart failure from a persistent left-to-right shunt.

Once definitive diagnosis is made in infants, usually with echocardiography, preparations are made for operative intervention, and prostaglandin E\textsubscript{1} is infused to maintain ductal patency and correct acidosis. The infant’s hemodynamic status should be optimized with mechanical ventilation and inotropic support. An effort should be made to increase pulmonary vascular resistance by decreasing the fractional inspired oxygen and avoiding hyperventilation because this will preferentially direct blood into the systemic circulation.

**Treatment.** Initial strategies for the management of IAA involved palliation through a left thoracotomy by using one of the arch vessels as a conduit to restore aortic continuity. Pulmonary artery banding can be simultaneously performed to limit left-to-right shunting because it is not feasible to repair the VSD or other intracardiac communications with this approach.

However, complete one stage surgical repair in infants with IAA is now preferable. The operative technique involves use of a median sternotomy and CPB with short periods of circulatory arrest. Aortic arch reconstruction can be accomplished with either direct anastomosis or patch aortoplasty followed by closure of the VSD.\textsuperscript{178}

In certain cases, the defect will involve hypoplasia of the left heart, precluding attempts at definitive repair. These infants should be managed with a Norwood procedure followed by a Fontan repair.

**Results.** Outcomes in infants with IAA have improved substantially over the last decades as a result of improved perioperative care. Operative mortality is now less than 10\% in most series.\textsuperscript{177,179} Some authors advocate the use of patch augmentation of the aorta to ensure adequate relief of LVOT obstruction and to diminish anastomotic tension, thus reducing the subsequent risk of restenosis and tracheobronchial compression.\textsuperscript{178}

**Pediatric Mechanical Circulatory Support**

Mechanical circulatory support has become standard therapy for adults with end stage heart failure. There has been a significant lag with development of similar devices for the pediatric population. This is probably related to the smaller market for these devices and the technical challenges associated with the anatomical constraints secondary to anatomy and size of the patients. Extracorporeal membrane oxygenation (ECMO)
has been the mainstay of mechanical support in many centers for the pediatric population. The adaptation of other adult devices to the pediatric population has led to the slow but steady development of pediatric durable mechanical devices. The Berlin Heart EXCOR (Berlin Heart AG, Berlin, Germany) device was approved by the FDA in 2011 in the United States as a paracorporeal device that can be used as a bridge to transplantation. This device has a 73% overall survival post implant at 12 months.\(^9\)

Infection, stroke and bleeding remain significant morbidities associated with it. Young age and small body surface area still remain poor prognostic factors. In 2010, the National Heart, Lung, and Blood Institute launched the Pumps for Kids, Infants, and Neonates (PumpKIN) program to promote development of new devices with the goal of clinical use.

ECMO remains the most commonly used form of mechanical support in the pediatric population in the United States. Per the ECLS Registry report released by the Extracorporeal Life Support Organization, as of January 2017, there were a total of 16,531 ECMO runs performed for cardiac causes, internationally.\(^1\)

The survival to discharge is about 40% in the neonatal population as opposed to 50% in the pediatric population. ECMO remains the only means of salvage for newborns and infants in many institutions. The biggest limitation remains the short duration it can be used. It is often used as a bridge to recovery and sometimes as a bridge to transplantation. The ability to place small infants on ECMO with peripheral cannulation continues to make it a very attractive first line option.

Ventricular assist devices can be either of the pulsatile or continuous types. The Berlin Heart EXCOR (Berlin Heart AG, Berlin, Germany) remains a classic example of a pulsatile device. The Impella 2.5 (Abiomed) (Fig. 20-68) has been used in the pediatric population as a temporary support device for recovering myocarditis, during treatment of acute rejection after heart transplantation and high-risk interventions in fragile patients with marginal function.\(^1\) Other continuous flow devices available for the pediatric patient include the Heartmate II and Heartmate III devices (Thoratec, Pleasanton, CA), DeBakey VAD Child (MicroMed Technology, Houston, TX), PediMag (Thoratec, Pleasanton, CA), Jarvik2015 and HeartWare HVAD (Fig 20-69) (HeartWare international Inc, Framingham, MA).\(^2\)

The total artificial heart (Syncardia Systems Inc, Tuscon, Az, USA) is an implantable biventricular device that replaces both ventricles. With the new introduction of the 50 ml pump, its popularity in the pediatric population has risen.

Posttransplant survival of patients bridged with and without mechanical circulatory support (ventricular assist device or total artificial heart) at 5 years post transplant remains the same. However, patients bridged to transplant with ECMO have a significantly worse survival.\(^3\) All in all, the field of pediatric heart surgery is very exciting and rapidly expanding.

Pediatric Heart Transplantation

Heart transplantation is currently an accepted mode of therapy in infants and children. Annually, about 600 pediatric heart transplants are performed worldwide,\(^4\) about 400 of which are performed in the United States.\(^5\) The common indications for heart transplant in the pediatric population are congenital heart disease, dilated cardiomyopathy, retransplantation, and other rare indications (e.g., arrhythmogenic right ventricular dysplasia, cancer, muscular dystrophy, and restrictive cardiomyopathy). The most common congenital heart defect requiring transplantation remains hypoplastic left heart syndrome. Although in the past some centers have advocated primary heart transplantation for this lesion, the improved outcomes with surgical palliation have eliminated this as an option. The first year post transplant remains the greatest risk for mortality. The overall median survival is 20.7 years for infants, 18.2 years for children age 1 to 5 years, 14 years for age 6 to 10 years, and 12.7 years for those age 11 to 17 years.\(^6\) Males seem to have a modestly superior overall survival compared with females. The causes of mortality include cardiac allograft vasculopathy, acute
rejection, infections, and graft failures. In the current era, the expected 1-year survival rate is 80% to 90%, the 2-year survival rate is 80% to 85%, and the 5-year survival rate is approximately 70% to 80% in experienced centers.interestingly, infants who undergo transplantation in the first month of life appear to have a survival advantage over infants who undergo transplantation during the remainder of the first year of life.

The two main techniques for performing the implant of the heart are the right atrial technique developed by lower and Shumway and the bicaval-left atrial technique described by Sievers and associates. In the latter technique, implantation consists of five anastomoses performed using a running prolene suture. These include the left atrial cuff, aorta, pulmonary artery, and the superior and inferior vena cava. One of the cornerstones of postoperative management remains immunosuppression. The triple drug regimen remains popular, corticosteroids, calcineurin inhibitor (cyclosporine or tacrolimus), and an antiproliferative agent (azathioprine or mycophenolate mofetil). Endomyocardial biopsy and coronary angiography are performed at regular intervals to monitor rejection. The field of pediatric heart transplantation has made huge strides since the days of “Baby Fae.”

Public Reporting and the STS Database in Congenital Heart Surgery

There has been a recent impetus in the field of congenital and pediatric cardiac surgery toward public reporting of outcomes. The advantages of this include promoting patient autonomy, shows a commitment to quality improvement, and also serves as a free marketing tool. The Society of Thoracic Surgeons Congenital Heart Surgery Database (STS-CHSD), is the largest clinical database in the world for congenital and pediatric cardiac surgery. It was founded in 1994. It contains data of about 394,980 operations as of September 9, 2016. These data are the foundation for assessment of performance by benchmark and comparison of individual programmatic outcomes to national aggregate data, development and subsequent application of sophisticated risk adjustment models, quality improvement initiatives, research, voluntary public reporting, development of reimbursement strategies, and governmental and regulatory collaborations.190 The database is currently in its 25th overall data harvest and records and represents data from 120 participants and 392 surgeons. Thus, this database has greater than 95% penetrance. STS CHSD public reporting started in January 2015, and participation is voluntary. Reporting is restricted to the hospital level and involves a rolling 4-year analytic window of data. Public reporting is based on the STS CHSD Operative Mortality Risk Model. Developed in 2014, this risk model calculates the operative mortality rate of hospitals performing such surgery, adjusting for procedural and patient level factors. The overall mortality rate over a 4-year period and the operative mortality rate for each of the five STAT (Society of Thoracic Surgeons—European Association for Cardio-Thoracic Surgery) categories is reported. The STAT categories are a multi-institutional, validated complexity stratification tool. They range from a score of 1 to 5, and the risk of mortality increases with each category. In addition, the STS star rating system was introduced, and every institution is rated as one, two, or three stars. This system is based on the confidence limits of the O/E (observed to expected) overall mortality for the institution (Fig. 20-70). One star equals higher than expected operative

<table>
<thead>
<tr>
<th>Rady Children’s Hospital San Diego</th>
</tr>
</thead>
<tbody>
<tr>
<td>San Diego CA</td>
</tr>
<tr>
<td>Website: <a href="http://www.rchsd.org/programs-services/cardiology">http://www.rchsd.org/programs-services/cardiology</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rady Children’s Hospital San Diego Surgeons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enc Devaney, MD</td>
</tr>
<tr>
<td>Daniel DiBardino, MD</td>
</tr>
<tr>
<td>John Lambert, MD</td>
</tr>
<tr>
<td>Peter Pastuszko, MD</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Overall Star Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>★★★</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Operative and Adjusted Operative Mortality, Last 4 Years (January 2012–December 2015)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Population:</strong> Neonates, Infants, Children &amp; Adults</td>
</tr>
<tr>
<td>#/Eligible</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>Overall</td>
</tr>
<tr>
<td>STAT Mortality Category 1</td>
</tr>
<tr>
<td>STAT Mortality Category 2</td>
</tr>
<tr>
<td>STAT Mortality Category 3</td>
</tr>
<tr>
<td>STAT Mortality Category 4</td>
</tr>
<tr>
<td>STAT Mortality Category 5</td>
</tr>
</tbody>
</table>

**Figure 20-70.** Program performance as currently reported by the STS-CHSD.
mortality (the 95% confidence interval for their risk-adjusted O/E mortality ratio was entirely above the number 1), two stars equals the same as expected operative mortality (the 95% confidence interval for their risk-adjusted O/E mortality ratio overlapped with the number 1), and three stars equals lower than expected operative mortality (the 95% confidence interval for their risk-adjusted O/E mortality ratio was entirely below the number 1). The Spring 2016 STS CHSD Feedback Report includes data from 117 participants in the STS-CHSD, including 14 one-star programs, 83 two-star programs, and 8 three-star programs. Twelve participants did not receive a star rating due to incomplete data. Public reporting increased from 23% to 57.6% (all three-star programs, 50 two-star and three one-star programs). The online public reporting portal can be accessed at www.sts.org/congenital-public-reporting-module-search.

There are several criticisms to the current methodology used for reporting. Important limitations of current publicly reported data (including the STS star rating system) will need to be addressed in future initiatives in order to completely engage parents of children with CHD and reassure providers that risk-adjustment models are optimized. There are four specific areas that should be considered when making decisions how to improve this methodology: (a) While the mortality risk-adjustment model on which the star rating system is based is mature now, there are not comparable models that provide risk-adjusted morbidity (complication) rates. The assessment of the quality of congenital heart disease care at different centers should include complication metrics and incorporate failure-to-rescue as an important discriminator; (b) the star rating system does not provide risk-adjusted outcomes for specific procedures or, more importantly, for specific diagnoses. This is mainly because of the exceptionally wide spectrum of diagnoses and procedures in pediatric cardiac surgery that preclude sufficiently large numbers in most procedure-specific categories; (c) the star rating system, although the “best” we have at present, may not be understood equally by all families. It will be critical to provide equivalent information to the large numbers of underresourced and non–English-speaking families; (d) finally, the current adjusted mortality rate reported by the STS is calculated from a statistical formula and refers to what the hospital’s mortality rate would be if the measured performance (in this case the mortality rate) were extrapolated to the overall case-mix or make-up of patients within the entire STS database. This is a critical point because a hospital’s case-mix is highly variable, and discrimination based on mortality is mostly related to outcomes of more complex procedures. In other words, if hospital A has excellent survival for less complex procedures and therefore performs very few highly complex procedures (i.e., choosing a case-mix consistent with its expertise), the application of an extrapolated mortality rate may not reflect the actual quality of care for that particular hospital. This issue is evident because the majority of experienced centers with arguably the highest complexity received a “middle star” rating of 2. This rating may reflect calibration issues with the current rating system, whereby centers are potentially penalized for high-complexity predominance.

Fortunately, there are efforts to correct these deficiencies. In 2016, the STS CHSD Task Force and STS Quality Measurement Task Force began to collaborate on an initiative to refine risk adjustment for chromosomal abnormalities, syndromes, and noncardiac congenital anatomic abnormalities and to then enhance the STS CHSD Mortality Risk Model with this additional information. Upon completion of this project, STS CHSD Task Force plans to collaborate with the STS Quality Measurement Task Force to study the relationship between volume (programmatic volume and surgeon volume) and outcome using this enhanced STS CHSD Mortality Risk Model. Also, currently under development is a multidomain quality metric that incorporates mortality, morbidity, postoperative length of stay, and the occurrence of complications. As the largest congenital and pediatric cardiac surgical clinical data registry in the world, containing data about nearly all pediatric cardiac operations performed in the United States, STS CHSD contains a truly representative sample of national aggregate data that is useful for multiple purposes.

**Future Directions**

The future of congenital heart surgery remains very bright and exciting. The development of novel technologies such as four-dimensional MRI flow studies (Fig. 20-71) and three-dimensional printing have offered this field several new tools to help understand complex anatomy and pathophysiology. Three-dimensional printing of complex congenital heart defects has helped surgeons in preoperative planning by allowing translation of two-dimensional cross-sectional imaging studies into a tangible and easily visualized model. The hollow nature of the human heart and the direct correlation of structure to disease in the congenital population allows this technology to be used in abundance in this field. Its utilization to train young surgeons is very appealing (Figs. 20-72 and 20-73). Current research in the field of genetics, device bioengineering and miniaturization, stem cell therapy, and fusion imaging technology is expected to further improve patient outcome. The improved outcomes and survival of these young and fragile patients with congenital heart disease has led to the development of a complex new field termed *adult congenital heart disease*. The field of congenital heart surgery is young and offers brilliant, motivated, and upcoming surgeons a very daunting challenge to better the future of these babies.

Figure 20-71. 4D MRI flow study obtained in a complex single ventricle patient for the evaluation of persistent hypoxia.
Figure 20-72. 3D printed models of complex heart defects which were very helpful for preoperative surgical planning and patient education.

Figure 20-73. Example of Pre-Interventional Planning Using 3D Printed Models. Transthoracic echocardiogram (A) confirms tetralogy of Fallot/pulmonary atresia/multiple aortopulmonary collateral arteries (MAPCAs) diagnosis. Three-dimensional (3D) reconstruction (B and C) illustrates spatial relationship of patient-specific geometry such as true pulmonary arteries (blue), aorta (red), and MAPCAs (green and yellow) for central aortopulmonary shunt placement and coil planning. Three-dimensional printing (D) provides absolute scaling for planning purposes, as well as patient/family education. Angiography (E and F) captured after central shunt and prior to placement of MAPCA embolization coils. (Reproduced with permission from Ryan JR, Moe TG, Richardson R, et al: A novel approach to neonatal management of tetralogy of Fallot, with pulmonary atresia, and multiple aortopulmonary collaterals, JACC Cardiovasc Imaging. 2015 Jan;8(1):103-104.)
REFERENCES

Entries highlighted in bright blue are key references.


116. Newburger JW, Sleeper LA, Frommelt PC, et al; Pediatric Heart Network Investigators. Transplantation-free survival and interventions at 3 years in the single ventricle reconstruction trial. *Circulation.* 2014;129:2013-2020. This article discusses the long-term outcomes of the Norwood procedure for patients with the Sano or BT shunt. This is a landmark article in the field.


Clinical Evaluation
As with any other field in medicine, the history and physical examination form the foundation for the evaluation of a patient with acquired heart disease requiring surgical intervention. Obtaining a complete history identifies comorbid conditions and assists in delineating the operative risks and prognosis after surgery. Physical examination reveals factors that may increase the complexity of surgery, such as previous surgery or the presence of peripheral arterial or cerebrovascular disease. These may influence the operative approach, but they also help guide the choice and sequence of diagnostic studies. A complete assessment of the patient allows the surgeon to make educated decisions regarding the optimal treatment strategy for the patient.

History
Symptoms suggestive of heart disease include: chest discomfort, fatigue, edema, dyspnea, palpitations, and syncope. Adequate definition of these symptoms calls for detailed history-taking, paying particular attention to onset, intensity, radiation, duration, and exacerbating or alleviating factors. The demands on the heart are determined by its loading conditions and the metabolic state of the patient. Cardiac symptoms are commonly accentuated with physical exertion or postural changes.

Angina pectoris is the hallmark of coronary artery disease (CAD), but may occur with other cardiac pathologies that result in ischemia from a mismatch between the supply of oxygen by the coronary circulation and the metabolic demand of the myocardium. Typically, angina is described as tightness, heaviness, or dull pain, frequently substernal, that lasts for a few minutes. This discomfort may radiate to the left arm, neck, mandible, or epigastrium. Angina is most often provoked by activities that increase metabolic demand on the heart such as exercise, eating, and states of intense emotion, and it is typically alleviated by rest or use of nitroglycerin. It is important to note that a significant number of patients with myocardial ischemia, particularly diabetics, females, and the elderly, may have “silent” angina or angina equivalents (dyspnea, diaphoresis, nausea, or fatigue). The overlap of these features with those of noncardiac etiologies...
such as costochondritis, biliary colic, gastroesophageal reflux disease, diffuse esophageal spasm, and peptic ulcer disease, to name a few, can sometimes lead to misdiagnosis.

Heart failure can occur from either left and/or right heart dysfunction, and respective symptoms arise from congestion of blood flow owing to an inadequate cardiac output. Left heart failure manifests as dyspnea, usually with exertion. Orthopnea, defined as dyspnea while lying flat, suggests worsened pulmonary congestion with increased venous return. Ascites, peripheral edema, and hepatomegaly reflect congestion in the systemic venous circulation and are prominent features of right heart failure. Peripheral edema can occur in right heart failure secondary to systemic venous congestion or in left heart failure due to salt and fluid retention as a result of impaired renal perfusion. Patients with chronic suboptimal perfusion and oxygenation can also have digital clubbing and cyanosis.

It is difficult to implicate cardiac disease based solely on the presence of fatigue, which is a very nonspecific symptom. However, most cardiac pathologies do result in fatigue or exercise intolerance to some degree. It is important to differentiate fatigue from exertional dyspnea which some patients may experience.

Dyspnea is another common symptom. Although generally a late symptom in patients with valvular heart disease or cardiomyopathy, it may be a relatively early complaint in some patients, particularly those with mitral stenosis. As stated previously, dyspnea is also an anginal equivalent and may signal a myocardial ischemic episode. Many primary pulmonary disorders feature dyspnea as their cardinal symptom and should be evaluated simultaneously as the physiology of the heart and lungs are intimately related and can have dramatic influences on one another.

Patients typically describe palpitations as a “skipped beat” or “racing heart.” Depending on the clinical context, such as occasional premature atrial or ventricular beats in otherwise healthy individuals, these may be benign. Clinically significant arrhythmias, however, require thorough investigation. Atrial fibrillation is the most common arrhythmia and can occur alone or with other cardiac pathologies. It results in an irregular, and at times, rapid heartbeat. Concurrent symptoms such as angina, lightheadedness, or syncope are particularly worrisome for life-threatening arrhythmias such as ventricular tachycardia or ventricular fibrillation, particularly in patients with preexisting heart failure.

Syncope associated with heart disease results from an abrupt reduction in cerebral perfusion. Many of the potential cardiac etiologies are serious, including sinus node dysfunction, atrioventricular conduction abnormalities, malignant arrhythmias, aortic stenosis, and hypertrophic obstructive cardiomyopathy. Noncardiac causes of syncope include, but are not limited to, neurologic causes (e.g., transient ischemic attacks [TIAs]), orthostatic hypotension, vasovagal events, and carotid sinus hypersensitivity. Any episode of syncope warrants a thorough evaluation and search for the root cause. In addition to a thorough inquiry regarding the aforementioned symptoms, it is important to obtain details about the patient’s medical and...
surgical history, family history, social habits (including alcohol and tobacco use), current medications, focused review of systems, as well as an assessment of the patient’s functional status and frailty. Frailty is often defined as a state of increased vulnerability to adverse health outcomes. Clinicians can use frailty index calculators to assess a patient’s risk for adverse outcomes following cardiac intervention. Specific attention should also be directed to the patient’s comorbidities which not only sheds light on their general health but also helps delineate expected risks from surgery. A strong family history of coronary artery disease, myocardial infarction, hypertension, or diabetes is of particular importance as they increase the individual’s risk for having an adverse cardiac event.

**Functional Disability and Angina.** With regard to heart failure, functional capacity is strongly correlated with mortality. The New York Heart Association (NYHA) functional class is a widely used classification system in categorizing patients based on their functional status (Table 21-1). The NYHA classification has become one basis by which to compare patient populations in many studies. Although less commonly used, the Canadian Cardiovascular Society (CCS) angina classification is also used to incorporate anginal symptoms into the functional assessment for prognostic value (Table 21-2).

### Physical Examination

The physical examination is an invaluable tool in directing further diagnostic studies in the management of a patient with suspected heart disease. The astute clinician may be able to detect subtle signs that may further characterize the underlying pathology.

The general appearance of a patient is important in the clinical assessment. A pale, diaphoretic, and obviously uncomfortable patient is more likely to be in a clinically critical condition than one who is conversing comfortably with an unremarkable demeanor. In addition to basic vital signs, particular attention should be directed to the patient’s mental status as well as the color and temperature of the skin, as these may be reflective of the general adequacy of perfusion. Overall frailty and dementia have also been shown to be predictors of operative and late mortality.

Palpation of the precordium may demonstrate deviation in the point of maximal impulse, indicative of ventricular hypertrophy, or parasternal heaves, seen in right ventricular overload. Auscultation should be performed in a quiet environment as critical murmurs, rubs, or gallops may be subtle. Murmurs are characterized by their location, timing, quality, and radiation. They are typically secondary to valvular or other structural pathology, and new findings require further investigation. A rub due to pericardial friction is indicative of endocarditis. A third heart sound (S₃) is generated by the rapid filling of a stiff ventricle and can be normal in young patients, but when present in older adults, is indicative of diastolic dysfunction and is pathologic. Increased contribution of the atrial pump function to ventricular filling may manifest as a fourth heart sound (S₄) and is also suggestive of ventricular dysfunction.

Palpation of peripheral pulses is important not only to assess the adequacy of perfusion, but also the burden of coronary artery disease often correlates with the degree of peripheral arterial disease. The presence of a carotid bruit may indicate clinically significant stenosis and may alter the course and timing of treatment, especially if symptomatic. Heart failure will frequently have extracardiac manifestations and examination of the other organ systems should not be neglected. For example, auscultation of the lung fields may reveal rales in patients with pulmonary edema. The use of accessory muscles of breathing may be present in patients with significant pleural effusions and volume overload. The presence of jugular venous distention and hepatosplenomegaly may signal right heart failure.

Additionally, clinicians should know the manifestation of other cardiac pathologies including endocarditis and rheumatic heart disease, although less commonly seen. Endocarditis is an inflammation of the endocardium, usually on the heart valves. A cardiac murmur is a common physical exam finding. Relatively uncommon but more specific clinical manifestations for infectious endocarditis are: Janeway lesions, which are nontender erythematous macules on the palms and soles; Osler nodes, which are described as tender subcutaneous nodules mostly on the pads of the fingers and toes; and Roth spots, which are exudative hemorrhagic lesions of the retina with pale centers.

### Table 21-1

<table>
<thead>
<tr>
<th>CLASS</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Physical activity not limited by symptoms: fatigue, palpitations, or dyspnea.</td>
</tr>
<tr>
<td>II</td>
<td>Comfortable at rest. Slight limitation of physical activity. Fatigue, palpitations, or dyspnea with ordinary physical activity.</td>
</tr>
<tr>
<td>III</td>
<td>Comfortable at rest. Marked limitation of physical activity. Fatigue, palpitations, or dyspnea with less than ordinary physical activity.</td>
</tr>
<tr>
<td>IV</td>
<td>Inability to carry out any physical activity. Symptoms may be present at rest and increase with activity.</td>
</tr>
</tbody>
</table>

### Table 21-2

<table>
<thead>
<tr>
<th>CLASS</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Ordinary physical activity (walking, climbing stairs) does not cause angina. Angina occurs with strenuous, rapid, or prolonged exertion during work or recreation.</td>
</tr>
<tr>
<td>II</td>
<td>Slight limitation of ordinary activity. Angina occurs with climbing stairs rapidly, walking uphill in the wind, under emotional stress, in the cold, or after meals. Walking more than 2 blocks or climbing one flight of stairs causes angina.</td>
</tr>
<tr>
<td>III</td>
<td>Marked limitation of ordinary physical activity (climbing a flight of stairs or walking 1 to 2 blocks at a normal pace).</td>
</tr>
<tr>
<td>IV</td>
<td>Inability to carry out any physical activity without discomfort. Angina may be present at rest.</td>
</tr>
</tbody>
</table>
Cardiac Risk Assessment in Noncardiac Surgery Patients

Cardiovascular complications occur in approximately 3% of patients undergoing inpatient noncardiac surgery. The American College of Cardiology (ACC) and American Heart Association (AHA) have formed a joint task force to publish a consensus statement, with periodic focused updates, on guidelines and recommendations that were revised in 2017. The aim of these guidelines is to incorporate surgery-specific risks and patient characteristics to guide perioperative decision-making in the management of patient with valvular heart disease.

Surgical procedures have been categorized based on cardiovascular risk into low risk, moderate risk, and vascular procedures. Vascular procedures, likely due to both the nature of the procedures themselves as well as the associated cardiovascular pathology in many of these patients, carry the highest reported risk of cardiac events at more than 7%. Low risk procedures, including endoscopic procedures, superficial operations, cataract surgery, breast surgery, and ambulatory surgeries, have a risk generally less than 1%. Intermediate risk procedures include intraperitoneal and intrathoracic surgery, head and neck surgery, orthopedic procedures, and prostate surgery.

Patient characteristics can be classified by the status of the patient’s cardiac disease, comorbid conditions, and functional capacity. Patients are considered to be at major perioperative clinical risk if they have one or more of the following active cardiac conditions: acute coronary syndrome, decompensated heart failure, significant arrhythmias, or severe valvular heart disease. In these patients, intensive evaluation and treatment prior to surgery (unless emergent) is warranted, prior to proceeding with noncardiac surgery.

If the patient does not have any of the previously mentioned active cardiac conditions, the perioperative risk of major adverse cardiac events (MACE) should be estimated. Both the operation performed and the patient’s risk factors are predictive of MACE, and the ACC/AHA guidelines recommend the use of either the American College of Surgeons’ NSQIP risk calculator or the Revised Cardiac Risk Index for the estimation of patient-specific risk. Patients at low (<1%) risk or patients at elevated risk with functional capacity greater than or equal to 4 metabolic equivalents (METs), should proceed to surgery without further testing. It is reasonable to perform pharmacologic stress testing in patients with poor or unknown functional capacity if this testing will impact decision making or perioperative care. Patients with abnormal stress test results should undergo confirmatory test such as coronary angiography, if indicated, before an elective noncardiac surgery. The previous guidelines included intermediate and low cardiovascular risk profiles, but this has been replaced by cardiovascular risk factors in the update. These risk factors are: history of ischemic heart disease, history of prior or compensated heart failure, history of cerebrovascular disease, diabetes mellitus, and renal insufficiency. Based on the number of present risk factors and the surgery-specific risk, the guidelines recommend pathways for further evaluation and risk management. The most recent guidelines from ACC/AHA were published in 2014 (Fig. 21-1). One important subgroup of patients at elevated risk are those who have recently undergone percutaneous coronary intervention. In these patients, elective noncardiac surgery should be delayed until the risk of stent thrombosis decreases (30 days for bare metal stents and 180 to 365 days for drug-eluting stents), and dual antiplatelet therapy should be continued unless the risk of bleeding exceeds the risk of stent thrombosis.

Diagnostic Studies

Electrocardiogram and Chest X-ray. Electrocardiograms (ECGs) and chest X-rays are noninvasive diagnostic studies that provide invaluable information in the preoperative assessment of patients with cardiac pathology. ECGs can be useful in detecting old myocardial infarction, dilation or hypertrophy of the cardiac chambers, arrhythmias, and conduction abnormalities. A stress ECG requires a patient to exercise to a target heart rate and is used to help diagnose ischemic pathologies that may not be evident at rest.

A plain film of the chest can detect pulmonary pathology, sequelae of heart failure (e.g., pulmonary edema, cardiac enlargement, pleural effusions), as well as presence of hardware (e.g., prosthetic heart valves, sternal wires, pacemakers, and defibrillators).

Echocardiography. Echocardiography utilizes reflected sound waves to image the heart. Transthoracic echocardiography (TTE) is used widely due to its noninvasive nature. It is the primary diagnostic tool used to evaluate structural diseases of the heart, including valvular pathology, septal defects, cardiomyopathies, and cardiac masses. Although more invasive, transesophageal echocardiography (TEE) can provide more information and better definition of some valvular and structural abnormalities. It is particularly useful in identifying left atrial thrombi in patients with atrial fibrillation. Echocardiography is indispensable in assessing surgical prostheses such as valves, leads, or mechanical circulatory support devices. These examinations can be performed with M-mode imaging (motion along a single line) as well as two-dimensional (2D) and three-dimensional (3D) imaging depending on the information required.

Doppler technology has become a standard addition to assess changes in flow patterns across dysfunctional valves. Velocity measurements can be obtained to estimate pressure gradients across structures using the continuity equation. A common example would be the estimation of pulmonary arterial systolic pressure calculated from the regurgitant tricuspid jet profile during right ventricular systole.

Transesophageal echocardiography requires no sedation and is generally performed with the patient in a slight left lateral decubitus position. Standardized views are obtained with the ultrasound probe placed in the apical, parasternal, subcostal, and suprasternal positions. The apical four-chamber view is a useful window for visualizing all four cardiac chambers simultaneously as well as the tricuspid and mitral valves. Other windows can be obtained to assess specific structures such as the individual valve anatomy or myocardial wall segments. Dobutamine-stress echocardiography is a study similar in idea to the stress ECG that utilizes a pharmacologic agent to assess the patient for ischemia or stress-induced valvular abnormalities.

Transesophageal echocardiography, on the other hand, is performed using a special endoscope with an ultrasound probe mounted on its end that is introduced orally into the esophagus under sedation. Posterior structures such as the mitral valve and left atrium are particularly well visualized. TEEs are frequently used intraoperatively during cardiothoracic surgery to assess global cardiac function, integrity of valve repairs and replacements, intracavitary thrombus and/or air, and aortic atherosclerosis or dissections that can have significant influences on operative strategy.

Echocardiography utilizes reflected sound waves to image the heart. Transthoracic echocardiography (TTE) is used widely due to its noninvasive nature. It is the primary diagnostic tool used to evaluate structural diseases of the heart, including valvular pathology, septal defects, cardiomyopathies, and cardiac masses. Although more invasive, transesophageal echocardiography (TEE) can provide more information and better definition of some valvular and structural abnormalities. It is particularly useful in identifying left atrial thrombi in patients with atrial fibrillation. Echocardiography is indispensable in assessing surgical prostheses such as valves, leads, or mechanical circulatory support devices. These examinations can be performed with M-mode imaging (motion along a single line) as well as two-dimensional (2D) and three-dimensional (3D) imaging depending on the information required.

Doppler technology has become a standard addition to assess changes in flow patterns across dysfunctional valves. Velocity measurements can be obtained to estimate pressure gradients across structures using the continuity equation. A common example would be the estimation of pulmonary arterial systolic pressure calculated from the regurgitant tricuspid jet profile during right ventricular systole.

Transesophageal echocardiography requires no sedation and is generally performed with the patient in a slight left lateral decubitus position. Standardized views are obtained with the ultrasound probe placed in the apical, parasternal, subcostal, and suprasternal positions. The apical four-chamber view is a useful window for visualizing all four cardiac chambers simultaneously as well as the tricuspid and mitral valves. Other windows can be obtained to assess specific structures such as the individual valve anatomy or myocardial wall segments. Dobutamine-stress echocardiography is a study similar in idea to the stress ECG that utilizes a pharmacologic agent to assess the patient for ischemia or stress-induced valvular abnormalities.

Transesophageal echocardiography, on the other hand, is performed using a special endoscope with an ultrasound probe mounted on its end that is introduced orally into the esophagus under sedation. Posterior structures such as the mitral valve and left atrium are particularly well visualized. TEEs are frequently used intraoperatively during cardiothoracic surgery to assess global cardiac function, integrity of valve repairs and replacements, intracavitary thrombus and/or air, and aortic atherosclerosis or dissections that can have significant influences on operative strategy.
There are some more recent additions to the echocardiographic armamentarium that capitalize on the strengths of ultrasound imaging. Three-dimensional TEE is playing an increasing role in the preoperative and intraoperative evaluation of patients with valvular heart disease and is especially helpful in percutaneous mitral intervention. Tissue Doppler imaging is based on principles akin to conventional Doppler echocardiography, but attention is directed to the myocardium itself as opposed to the motion of blood to quantify abnormalities in wall motion. Strain imaging with speckle-tracking echocardiography measures the actual deformation of the myocardium by following inhomogeneities inherent to the myocardium and is a useful measure of myocardial function.

**Radionuclide Studies.** Although ECGs are useful, inexpensive, and safe, baseline abnormalities in the ECG may limit its diagnostic capacity. In particular, ventricular rhythms, bundle-branch blocks, left ventricular hypertrophy, drug effects, and baseline ST-segment depressions can make stress ECGs difficult to interpret for the presence of myocardial ischemia. In this setting, myocardial perfusion imaging (MPI) using radionuclides can be utilized. Thallium 201 ($^{201}$Tl) was the initial radionuclide used for MPI, but due to its long half-life and relatively low photoprobe, it has largely been replaced by technetium-99m (sestamibi and tetrofosmin) because of its more favorable characteristics. In the past, planar imaging with three separate 2D views of the heart were obtained. Currently, it is more common to have the images acquired by single-photon emission computed tomography (SPECT) technology, which detects emitted photons from 180° to 360° around the patient. The signals are then processed to reconstruct multiple slices that together provide a 3D image. The distribution of radionuclides depends on perfusion, and therefore areas that show uptake at rest, but not during stress, are concerning for ischemia. The amount of uptake at both rest and stressed states is compared to assess ischemia and viability of the myocardium. Territories that do not show uptake at rest or during stress are likely to be nonviable scar. The sensitivity and specificity of exercise SPECT are 87% and 73%, respectively.8,9

The image acquisition may also be gated to a simultaneously obtained ECG to assess global ventricular function. The endocardial and epicardial borders (as delineated by radionuclide uptake) are detected throughout the cardiac cycle and the ejection fraction, along with end-systolic and end-diastolic volumes, can be calculated. This study is also useful in revealing hypokinetic segments of the myocardium.

One significant drawback of SPECT imaging is that it shows regional ischemia well, but it does not adequately detect global or “balanced” ischemia that can occur with diffuse CAD. Positron emission tomography (PET) has been used due to its ability to obtain absolute quantitative data on both myocardial perfusion and metabolism. Tracers used in PET scans can be divided into those that assess perfusion (oxygen-15, nitrogen-13, and rubidium-82) and those that assess metabolism (carbon-11 and fluorine-18). The specificity of PET in detecting CAD is better than SPECT at 86% due to its superior spatial resolution.10

**Magnetic Resonance Imaging.** Magnetic resonance imaging (MRI) has a wide variety of uses in cardiac imaging depending on the pulse sequence and signal weighting. Cine-loop imaging of the heart throughout the cardiac cycle can yield information on global chamber function, geometry, and valvular pathologies.

The differential response of normal and ischemic myocardium to certain pulse sequences allows imaging of myocardial perfusion using MRI. Use of contrast agents such as gadolinium can enhance scar tissue and are very useful in viability assessment. Myocardial strain imaging can also be performed taking advantage of radio-frequency tagging of the myocardium, which deforms with the tissue and can be followed throughout the cardiac cycle.

**Cardiac Catheterization.** Cardiac catheterization involves access to the cardiac chambers, coronary arteries, and great vessels with a peripherally inserted catheter under fluoroscopic guidance. It is a versatile tool used to investigate cardiac chamber pressures, valvular abnormalities, wall motion, and coronary artery anatomy. While some of these roles are being replaced by less invasive techniques mentioned previously, cardiac catheterization continues to be widely performed and is the gold standard for the assessment of coronary artery disease.11

Left heart catheterization is performed by percutaneous access of the femoral, or radial, artery. Under fluoroscopic guidance, the catheter is threaded into the ascending aorta where a contrast aortogram may be performed. Coronary angiography requires manipulation of this catheter into the coronary ostia where contrast is directly injected. With advancement of the catheter retrograde through the aortic valve, left ventricular pressures can be obtained. This measurement is used to calculate direct pressure gradients across the aortic valve, in contrast to echocardiography that indirectly measures pressure, and can be used to confirm severe aortic stenosis. Again, contrast injection into the left ventricle can be used to estimate ejection fraction and visualize hypokinetic segments of the myocardium. Inappropriate retrograde leakage of contrast may indicate insufficiency of the aortic and/or mitral valves.

Coronary angiography provides information on hemodynamically significant stenoses in the coronary circulation as well as an anatomical roadmap for surgeons to plan revascularization (Fig. 21-2A,B). A stenosis is considered to be significant if it narrows the lumen of the artery by 70% (or 50% in the case of left main coronary artery). Borderline lesions or complex lesions may be assessed by fractional flow reserve (FFR) or instantaneous wave-free radio (iFR), which obviates the need of adenosine.12 This has been shown to very helpful in guiding revascularization strategies in recent clinical trials.13 Additional assessment can also be done using intravascular ultrasound (IVUS) inside the coronary circulation. There is some variability in the coronary arterial anatomy with the posterior descending artery being supplied by the right coronary artery in approximately 80% of patients (right dominant) or the left coronary artery in approximately 15% of patients (left dominant). The remaining patients have a codominant circulation where the posterior descending artery is supplied by both the right and left coronaries.

Right heart catheterization is performed by the introduction of catheter through a peripheral vein that is advanced into the right side of the heart.14 Right-sided pressures and structures are assessed in a similar fashion as in the left heart. Extension of the catheter into the pulmonary artery allows measurement of pulmonary artery pressures as well as pulmonary capillary wedge pressure (reflecting left ventricular end diastolic pressure) with an occlusive balloon. In addition to these measurements, cardiac output can be measured using thermodilution or by the Fick method using oxygen saturations of blood sampled from the various locations during the procedure.
ACQUIRED HEART DISEASE

CHAPTER 21

Figure 21-2. Cardiac catheterization angiography. A. Stenosis of right coronary artery indicated by the arrow. B. Still image of a normal left ventriculogram.

An advantage of cardiac catheterization is that it offers an opportunity for interventional therapy of coronary artery disease, arrhythmias, valvular abnormalities, and other structural defects of the heart. Cardiac catheterization is generally safe, but being an invasive procedure, it is associated with rare complications. The overall mortality is 0.11%, and total rate of major complications, including MI, stroke, arrhythmia, vascular injury, contrast reaction including allergic reaction and contrast-induced nephropathy, hemodynamic instability, and cardiac perforation is usually <2%.15

Cardiac Computed Tomography. Multislice computed tomography (CT) imaging can be used to assess the coronary vasculature. The coronary calcium score is an index developed to quantify the degree of coronary atherosclerotic burden by measuring Hounsfield units in a noncontrast cardiac CT. Although this technique is quite sensitive for angiographic stenoses (>50%) it remains fairly nonspecific as calcification often precedes significant luminal narrowing.16 CT coronary angiography using intravenous contrast is also utilized to assess coronary pathology and is particularly useful in the emergency room to perform a “triple rule-out” for acute coronary events, pulmonary embolism, and aortic dissection in patients who present with undifferentiated chest pain. LV ejection fraction may be measured by this technique, and, together with the degree of coronary stenosis, it has been shown to have incremental prognostic value for the presence of coronary artery disease and in the prediction of adverse coronary events.17

EXTRACORPOREAL PERFUSION

History

Prior to the development of extracorporeal perfusion, heart surgery was rarely performed and was limited to brief periods of asystole and/or hypothermia. The need for obtaining a bloodless operating field, while maintaining perfusion of heart and other organs, was evident.

John Gibbon’s motivation to develop a means for extracorporeal perfusion came from a desire to safely open the pulmonary artery in a patient who suffered from a pulmonary embolus following a cholecystectomy. After numerous experimental iterations, Gibbon’s cardiopulmonary bypass machine was first used clinically in 1953 to repair an atrial septal defect in an 18-year-old woman.18 Although Gibbon is credited for its invention, the development of modern cardiopulmonary bypass (CPB) is a culmination of the work of many investigators throughout the world. The early bubble oxygenators have evolved into the currently used membrane oxygenators. The search for an ideal biocompatible material that minimizes the inflammatory cascade initiated by the contact of blood with the circuit components continues to this day.

Technique

The basic CPB circuit consists of the venous cannulae, a venous reservoir, pump, oxygenator, filter, and the arterial cannula. Anticoagulation is required during CPB, and 300 to 400 units/kg of heparin are given to increase the activated clotting time (ACT) to greater than 450 seconds. Once an adequate level of anticoagulation is achieved, arterial cannulation is performed through a purse-string suture or through a side graft sewn onto the native artery. The distal ascending thoracic aorta is the most common site of cannulation. Other routinely utilized sites of cannulation include the femoral artery, the axillary artery, the innominate artery, or the distal aortic arch, and they are altered based on the indicated surgical repair and the presence of native arterial disease. Venous cannulation is performed through purse-string sutures placed on the right atrium either for a single cannula or for two separate cannulae placed into the superior and inferior vena cava, respectively. Alternatively, the venous cannula may be inserted from the femoral vein and advanced into the right atrium and superior vena cava using TEE guidance. This technique is frequently used in minimally invasive cardiac surgery. Effective communication between the surgeon, the anesthesiologist, and the perfusionist is mandatory for effective
Cardiopulmonary bypass. Once the appropriate cannulations and connections are complete, CPB is commenced. Venous return is initiated followed by arterial flow while monitoring systemic blood pressure. At normothermia, the flow required is approximately 2.4 L/min/m², but with hypothermia, oxygen consumption is reduced by 50% for every 10°C drop in temperature, and a flow of only 1 L/min/m² is required at 18°C. Once the heart is decompensated and hemodynamics are acceptable, ventilation is stopped. The oxygenator is adjusted to maintain a Pao₂ of 150 mmHg and normocarbia. Blood can also be filtered and returned through vents that are placed in chambers of the heart (such as the left ventricle or pulmonary artery) or through the cardiotomy suction used to aspirate blood from the surgical field.

When the cardiac procedure is complete, the patient is rewarmed, the lungs ventilated, and the heart defibrillated, if needed. The venous return to the CPB machine is gradually reduced allowing the heart to fill. The pump is also slowed while hemodynamics and global cardiac function are assessed with a TEE probe. Inotropic and vasopressor support may be used to augmen cardiac function and treat hypotension. Once CPB has been weaned and stable hemodynamics achieved, the cannulae are removed. The heparin anticoagulation is reversed with 1 mg protamine per 100 units of heparin and hemostasis is achieved.19

Adverse Effects

Cardiopulmonary bypass has a number of deleterious effects as various intertwining processes result in derangements in hemostasis, an enhanced systemic inflammatory response, and end-organ function.

Anticoagulation prior to the commencement of CPB is required as contact of blood with the artificial surfaces of the circuit can initiate a thrombogenic cascade. Generation of thrombin plays a major role in both thrombotic and bleeding phenomena during CPB. The endothelium that normally regulates the fine balance between procoagulant and anticoagulant pathways is perturbed. Fibrinogen is consumed rapidly as thrombin converts fibrinogen to fibrin while fibrinolytic mechanisms (initiated by the activated endothelium) degrade the fibrin macromolecules. Platelets are activated by the converging hemostatic pathways and are consumed.

The response of the humoral and cellular immune systems partly overlap with the hemostatic pathways. The classic and alternative complement pathways are activated by CPB generating powerful chemotactic molecules and anaphylatoxins.20 Monocytes, platelets, and neutrophils are activated releasing acute inflammatory mediators and cytokines that persist even after conclusion of CPB.21 These inflammatory cells also produce reactive oxidants that may have cytotoxic and cardiovascular effects such as vasoconstriction and hypotension.

The large quantity of unfractionated heparin used during cardiac surgery predisposes patients to developing heparin-induced thrombocytopenia (HIT) with an incidence of 1% to 5%.22 Platelet factor-4 (PF-4) is produced by platelets and avidly binds to heparin to form a heparin-PF-4 complex that can be antigenic in some patients binding IgG. The IgG-heparin-PF-4 complex can bind to platelets, which causes release of more PF-4, perpetuating the process. The earliest sign is a sudden drop of more than 50% in the platelet count, usually seen from several hours to days after surgery. HIT can be confirmed with an enzyme-linked immunosorbent assay (ELISA) or serotonin release assay (SRA). Of the patients with HIT, 20% to 50% of patients develop thromboses in arterial or venous beds, designated as heparin-induced thrombocytopenia and thrombosis (HITT), which can be life-threatening.23 Treatment is anticoagulation with nonheparin anticoagulant (e.g., argatroban, bivalirudin).

The etiology of end-organ dysfunction resulting from extracorporeal circulation can mostly be categorized into one of three mechanisms: hypoperfusion, embolization, and whole-body inflammatory response. Although cardiac output and blood pressure are monitored carefully during CPB, they are surrogates for regional perfusion and cannot detect end-organ hypoperfusion directly. This can be a problem particularly with the cerebral, renal, and mesenteric circulations. With manipulation of diseased vessels and dysregulation of the native coagulation system, macroscopic and microscopic emboli are a concern. Activated cells and circulating cytotoxic products of the immune response may cause microvascular injury and edema of other organs manifesting as neurocognitive deficits, respiratory failure, and renal injury.24

Myocardial Protection

During CPB, pharmacologic agents in cardioplegic solutions may be delivered into the coronary circulation to arrest the heart, allowing for a still operating target and improved myocardial protection. The most common cardioplegia consists of potassium-rich solutions that can be mixed with autologous blood and are delivered into the coronary circulation.25 Antegrade cardioplegia is delivered into the root of a cross-clamped aorta or directly into the individual coronary ostia using specialized catheters. A retrograde cardioplegia catheter is a balloon-cuffed catheter that is placed through the right atrium into the coronary sinus and is used to perfuse the coronary circulation in the opposite direction through the venous circulation. This has the advantage of more uniform distribution in patients with diffuse coronary artery disease and is not dependent on a competent aortic valve for delivery.

There is continued debate regarding the best method (antegrade vs. retrograde vs. both), type (crystalloid vs. blood), temperature (cold vs. warm vs. tepid), and interval (continuous vs. intermittent) of cardioplegia delivery. The optimal combination is beyond the scope of this text. However, most cardiac surgeons in the United States favor cold blood potassium cardioplegia.

CORONARY ARTERY DISEASE

History

Aortocoronary bypass for myocardial ischemia was first proposed and performed in laboratory animals by Carrel in 1910.26 The Vineberg operation, one of the initial attempts at surgical revascularization of the myocardium, was introduced in 1951.27 This procedure involved implantation of the internal thoracic artery directly into the myocardium itself. While some patients were relieved of their anginal symptoms, this resulted in virtually no increase in coronary flow and was soon supplanted by methods to restore flow directly. Coronary endarterectomy was introduced by Longmire during this time period but had high rates of restenosis and occlusion.28 The use of vein patches to repair the arteriotomy sites was described by Senning in 1961.29 The first saphenous vein coronary artery bypass grafting (CABG) was performed by Sabiston in 1962,30 but was popularized by Favalaro and Sones in 1967.31 In 1968, the internal thoracic artery was introduced as a bypass conduit by Green, who used it to bypass the left anterior descending coronary artery.32
Etiology and Pathogenesis
Atherosclerotic stenoses are the primary mechanism of CAD. The pathophysiologic process is initiated with vascular endothelial injury and is potentiated by inflammatory mechanisms, circulating lipids, toxins, and other vasoactive agents in the blood. Macrophages and platelets are attracted to this area of endothelial dysfunction inciting a local inflammatory response. During this process, macrophages infiltrate into the intimal layers and accumulate cholesterol-containing low-density lipoproteins. The growth factors secreted promote proliferation of smooth muscle cells within the intima and media of the arteries. Together with the accumulation of the lipid-laden macrophages, the smooth muscle hyperplasia results in an atheroma and subsequent stenosis of the vessel. These atheromas have a fibrous cap that may rupture, exposing the underlying cells and extracellular matrix, which are very prothrombotic. Acute plaque rupture and thrombus formation is thought to be the main pathophysiologic mechanism responsible for acute coronary syndromes.33-35

Risk Factors and Prevention
Prior to the establishment of modern management strategies, the annual mortality rate from ischemic heart disease was 482 out of 100,000 persons.36 Since the peak of coronary heart disease mortality in 1968, modern primary and secondary prevention strategies such as risk factor modification, percutaneous and surgical revascularization, use of medications (e.g., aspirin, HMG-CoA reductase inhibitors [statins], and β-blockers), has decreased mortality from coronary artery disease by 74%.36

The major risk factors for atherosclerosis include advanced age, cigarette smoking, hypertension, dyslipidemias, sedentary lifestyle, obesity, and diabetes. Likely due to increased public awareness and aggressive medical management, these risk factors (with the exception of glucose intolerance and obesity) have recently been on the decline.

Current guidelines outlined in the AHA/ACC consensus statement summarize the secondary prevention recommendations.37 Class I recommendations include smoking cessation and avoidance of environmental tobacco exposure, blood pressure control to under 140/90 mmHg (under 130/80 mmHg in those with diabetes or chronic kidney disease), LDL cholesterol levels less than 100 mg/dL, aspirin therapy in all patients without contraindications, a BMI target of less than 25 kg/m², diabetes management with target HbA1c <7%, and encouragement of daily moderate-intensity aerobic exercise. β-Blockers should be used in all patients with LV dysfunction and following MI, ACS, or revascularization, unless a specific contraindication is present. Renin-angiotensin-aldosterone system blockade in patients with hypertension, LV dysfunction, diabetes, or chronic kidney disease should also be considered.

Clinical Manifestations
Patients with CAD may have a spectrum of presentations, including angina pectoris, myocardial infarction, ischemic heart failure, arrhythmias, and sudden death.

Angina pectoris is the pain or discomfort caused by myocardial ischemia and is typically substernal and may radiate to the left upper extremity, neck, or epigastrium. The variety of presentations can make myocardial ischemia challenging to diagnose. Characteristics of chest pain that make myocardial ischemia less likely include pleuritic chest pain, pain reproducible by movement or palpation, or brief episodes lasting only seconds. Typical angina is relieved by rest and/or use of sublingual nitroglycerin. Differential diagnoses to be considered include, but are not limited to, musculoskeletal pain, pulmonary disorders, esophageal spasm, pericarditis, aortic dissection, gastroesophageal reflux, neuropathic pain, and anxiety.

Myocardial infarction is a serious consequence of CAD occurring when ischemia results in myocardial necrosis. This may be silent and need not be preceded by angina. Necrosis may result in disruption of the myocardial integrity leading to devastating conditions such as intracardiac shunts from ventricular septal defects, acute valvular regurgitation from rupture of necrotic papillary muscles, and cardiac aneurysms, which have the potential for fatal rupture.

Ischemic insults from CAD may lead to congestive heart failure. The initial myocardial damage sets off a cascade of both local and systemic responses. Over time, these changes can cause deleterious myocardial loading and abnormal neurohumoral responses that result in pathologic remodeling of the heart. Heart failure should be suspected in patients who present with dyspnea, orthopnea, fatigue, and edema.

Arrhythmias may also be sequelae of CAD. Ischemic etiologies should be investigated in patients who present with new arrhythmias. CAD may result in arrhythmias following an acute MI or as the result of ultrastructural and electrophysiologic remodeling secondary to chronic ischemic heart disease. Ischemia of the electrical conduction system may be present in the form of new onset complete or partial atrioventricular conduction blocks.

Preoperative Evaluation
A focused history and physical examination is essential with particular attention directed to the signs, symptoms, and clinical manifestations mentioned previously. The patient’s functional status is of importance not only because it is a component of preoperative risk assessment, but also because quality of life improvement and symptomatic relief are both goals of surgical therapy.

Coronary angiography is the primary diagnostic tool. The coronary anatomy and degrees of stenoses are delineated allowing for planning of surgical revascularization.

Noninvasive diagnostic studies, in combination with provocative maneuvers (exercise or pharmacologic agents) offer information regarding the functional significance of ischemic disease. A stress ECG is frequently used as a screening tool with 50% sensitivity and 90% specificity for coronary artery with a threshold of 1 mm of ST-segment depression.39 This test, however, requires patients to achieve a certain elevation in their heart rate and is therefore not suitable for those that cannot achieve this goal. Furthermore, baseline ECG abnormalities may render it impossible to detect typical ischemic changes with stress.

Echocardiography and nuclear imaging may be performed under pharmacologic stress (with dobutamine or dipyridamole) to assess reversible ischemia and myocardial viability. Technetium-99m or thallium-201 perfusion scans and stress echocardiography are more sensitive than stress ECG.39 These studies also have the ability to assess global ventricular function in terms of left ventricle ejection fraction, which can be used to determine operative risk. Please refer to the diagnostic studies section for more details.

CORONARY ARTERY BYPASS GRAFTING
Indications
A joint committee established by the ACC/AHA have published guidelines for surgical revascularization (CABG) in
Table 21-3
Algorithm set forth by ACC/AHA guidelines for preoperative cardiovascular evaluation before noncardiac surgery for patients who are scheduled for nonemergent, non-low risk surgery, no active cardiac disease, and less than 3 METs

<table>
<thead>
<tr>
<th>NUMBER OF RISK FACTORS</th>
<th>RECOMMENDATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Proceed with planned surgery.</td>
</tr>
<tr>
<td>1–2</td>
<td>Control HR and proceed with planned surgery or pursue further testing if it will change management.</td>
</tr>
<tr>
<td>3–5</td>
<td>Pursue further testing if it will advance management.</td>
</tr>
</tbody>
</table>

*Risk factors are history of ischemic heart disease, history of prior or compensated heart failure, history of cerebrovascular disease, diabetes mellitus, and renal insufficiency.

Table 21-4
Data from ACC/AHA guidelines for CABG in CAD to improve survival

<table>
<thead>
<tr>
<th>ANATOMY</th>
<th>CLASS OF RECOMMENDATION</th>
<th>LEVEL OF EVIDENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>• 3-vessel +/– proximal LAD</td>
<td>I</td>
<td>B</td>
</tr>
<tr>
<td>• 2-vessel + proximal LAD</td>
<td>I</td>
<td>B</td>
</tr>
<tr>
<td>• 2-vessel – proximal LAD</td>
<td>Ia – with extensive ischemia</td>
<td>B</td>
</tr>
<tr>
<td>• Multivessel disease with DM</td>
<td>Ia (CABG preferred over PCI)</td>
<td>B</td>
</tr>
<tr>
<td>• Proximal LAD only</td>
<td>Ia – with LITA for long-term benefit</td>
<td>B</td>
</tr>
<tr>
<td>• 1-vessel – proximal LAD</td>
<td>III – Harm</td>
<td>B</td>
</tr>
<tr>
<td>• LV dysfunction</td>
<td>IIa – LVEF 35%–50%</td>
<td>B</td>
</tr>
<tr>
<td>• Survivor of ischemia-mediated VT</td>
<td>Ilb – LVEF &lt;35% without LM disease</td>
<td>B</td>
</tr>
</tbody>
</table>

DM = Diabetes mellitus; LITA = Left internal thoracic/mammary artery; LM = Left main coronary artery; LV = left ventricle; VT = ventricular tachycardia. Class of recommendation: I – Benefit far outweighs risks and procedure should be performed; Ia – Benefit outweighs risks and procedure is considered to be reasonable; Iib – Potential benefits may exceed risks and procedure may be considered; II – Procedure not helpful and may cause harm. Level of evidence: A – Strong; multiple supporting randomized controlled trials or meta-analyses. B – Limited; data based on a single randomized trial or nonrandomized trials. C – Very limited; based on expert consensus, case studies or standards of care.
modest with an incremental cost-effectiveness ratio of $30,454 per quality-adjusted life year gained.47

**Everolimus-Eluting Stents or Bypass Surgery for Left Main Coronary Artery Disease (EXCEL Trial, 2016).** This study randomized 1905 patients with left main coronary artery disease of low or intermediate complexity to PCI or CABG. At 3 years, no difference was observed in the primary endpoint of death, stroke, myocardial infarction or ischemia-driven revascularization. However, it should be noted that the PCI group exhibited a greater increase in these events between 30 days and three years than the CABG group (11.5% vs. 7.9%).48

**Summary**

PCI technology has improved over time, and rates of periprocedural adverse events have decreased significantly. Management strategies must be tailored to the individual patient’s clinical status and context, but CABG maintains improved long-term outcome and remains the standard of care for patients with left-main, multivessel coronary artery disease and patients with diabetes. Recent and upcoming trials will continue to clarify which patient populations benefit most from either revascularization strategy.

**Operative Techniques and Results**

**Bypass Conduit Selection.** The most important criterion in conduit selection is graft patency. The conduit with the highest patency rate (98% at 5 years and 85%–90% at 10 years) is the internal thoracic artery, which is most commonly left attached proximally to the subclavian artery (although occasionally used as a free graft) and anastomosed distally to the target coronary artery.49,50 The use of both internal thoracic arteries has been shown to increase event-free survival in a number of studies.51-52

The greater saphenous vein can be harvested using an open or endoscopic technique. In the open technique, the initial incision is made along the course of the vein on the medial aspect of the lower extremity. The vein is harvested with meticulous attention directed toward minimizing manipulation of the vein itself. The incision may be continuous or bridged in an attempt to decrease the size of the incision, but multiple bridged incisions may have the potential risk of increased conduit manipulation during harvest. Endoscopic harvest is performed by making a small incision just above and medial to the knee where the endoscope is inserted. Side branches are cauterized under endoscopic visualization using bipolar electrocautery until dissection is carried proximally until the required length of vein is mobilized. A proximal counterincision is then made to extract the venous conduit, which is prepared in the standard fashion.

The radial artery is another frequently used conduit. After confirmation of ulnar collateral flow to the hand by the clinical Allen’s test or a duplex ultrasound study, an incision is made from a point just proximal to the radial styloid process and ending just medial and distal to the biceps tendon on the nondominant hand. With lateral retraction of the brachioradialis muscle, the radial artery is dissected sharply with care to avoid injury to the cutaneous nerves in this area and minimize manipulation of the artery itself. This artery can also be harvested using an endoscopic technique.

Many studies have looked at the patency rates of the radial artery graft in comparison to the saphenous vein graft. Although some studies have resulted in equivocal data, general consensus favors the use of radial arterial grafts over vein grafts with 5-year patency rates of 98% and 86%, respectively.53-54

From a historical perspective, the anterior circulation (left anterior descending artery) is generally bypassed using the internal thoracic artery, and the lateral (circumflex artery) or inferior (right coronary artery) territories are bypassed using a saphenous vein or radial artery graft. These conduits may be combined to form a composite T- or Y-graft, or sewn to multiple targets as sequential grafts. Since patency is best with arterial grafts, recent data have suggested that the best long-term results are achieved using a multiple or all-arterial revascularization strategy, particularly in patients >70 years of age and patients with diabetes.55-57 Other conduits such as the gastroepiploic arteries, lesser saphenous veins, and cephalic veins have been described, but these are not widely used and will not be discussed here.

**Conventional Coronary Artery Bypass Grafting.** Traditionally, CABGs are performed with the patient lying supine through a median sternotomy. Left internal thoracic artery and other conduit harvests are performed. After the patient is heparinized, cardiopulmonary bypass is initiated. The aorta is cross-clamped, and cardioplegia is delivered. Once adequate myocardial protection has been achieved, coronary arteriotomies are made, and distal anastomoses are performed using polypropylene suture (Fig. 21-3A,B). The proximal anastomoses are then performed directly onto the ascending aorta or onto preexisting grafts. It is important to note that significant coronary stenoses can cause differential distribution of cardioplegia and myocardial protection. It is therefore recommended to use retrograde cardioplegia or to revascularize the area with the most concern for ischemia first and give cardioplegia down the completed graft. The left internal thoracic artery to left anterior descending (LAD) graft is frequently performed last to avoid kinking or disruption of this important conduit. Once all grafts

---

**Table 21-5**

Data from ACC/AHA guidelines for CABG in CAD to improve symptoms

<table>
<thead>
<tr>
<th>ANATOMY ASSOCIATED SYMPTOMS</th>
<th>CLASS OF RECOMMENDATION</th>
<th>LEVEL OF EVIDENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Unacceptable angina with presence of ≥1 stenoses amenable to revascularization despite medical treatment</td>
<td>I</td>
<td>A</td>
</tr>
<tr>
<td>• Complex 3-vessel CAD +/- proximal LAD involvement</td>
<td>IIa (CABG preferred over PCI)</td>
<td>B</td>
</tr>
<tr>
<td>• Unacceptable angina with presence of ≥1 stenoses amenable to revascularization but medical treatment is not possible</td>
<td>IIa</td>
<td>C</td>
</tr>
<tr>
<td>• Previous CABG with ≥1 stenoses associated with ischemia and angina despite medical treatment</td>
<td>IIb</td>
<td>C</td>
</tr>
</tbody>
</table>
Conventional CABG Results. Several early randomized trials showed improved survival in patients who received a CABG as opposed to medical therapy. A propensity-matched study identified that CABG greatly benefited patients with LV dysfunction and left main stenosis >50% compared to medical management. The Bypass Angioplasty Revascularization Investigation (BARI) trial demonstrated impressively superior results with CABG compared to PCI in terms of 5-year cardiac mortality (5.8% vs. 20.6%) in patients with diabetes in addition to CAD. In a study examining the benefits of CABG over medical management for specific CAD distributions, survival was better in patients with proximal LAD stenoses, regardless of the number of diseased vessels. In general, these studies show survival rates of over 90% at 5 years and approximately 75% at 10 years following CABG.

The mortality and morbidity of the procedure itself has changed over time. Data from the Society of Thoracic Surgeons (STS) database accounts for 1,497,254 patients who underwent solitary CABG from 2000 to 2009. The mortality rate of CABGs has improved significantly from 2.4% in 2000 to 1.9% in 2009, despite the relatively constant predicted mortality rate of around 2.3%. In parallel with this, postoperative complication rates have also decreased: stroke (1.6%–1.2%), bleeding requiring reoperation (2.4%–2.2%), and deep sternal wound infection (0.59%–0.37%).

Off-pump Coronary Artery Bypass. To avoid the adverse consequences of cardiopulmonary bypass, off-pump coronary artery bypass (OPCAB) was developed and has been adopted in some centers over the past two decades.

With OPCAB, the heart is left beating. Performing anastomoses on the beating heart requires the use of myocardial stabilization devices, which help portions of the epicardial surface to remain relatively immobile while the anastomoses are being performed (Fig. 21–4).

Apical suction devices are used to aid in exposure, particularly of the lateral and inferior vessels. Many creative maneuvers have been developed, including patient repositioning, opening the right pleural space to allow for cardiac displacement, and creation of a pericardial cradle to minimize compromise of cardiac function while exposing the various surfaces of the heart. Temporary proximal occlusion of the target coronary artery, or the use of an intracoronary shunt, are necessary to...
provide adequate exposure of the anastomosis. Occlusion causes temporary ischemia, and if not tolerated during a test occlusion, coronary shunts can be employed.

**OPCAB Results.** The superiority of OPCAB over on-pump CABG remains a controversy despite the large body of literature on this topic. A pooled analysis of two randomized trials, the Beating Heart Against Cardioplegic Arrest Studies (BHACAS 1 & 2), is one of several studies that have touted lower short-term mortality rates with the off-pump compared to the on-pump technique.\(^\text{65-67}\) Other studies, however, have demonstrated equivocal or contrary results.\(^\text{68-70}\) Furthermore, the prospective and much larger ROOBY (Randomized On/Off Bypass) trial showed increased rates of adverse cardiac events with OPCAB compared to conventional CABG.\(^\text{71}\) Despite the initial enthusiasm for the theoretical advantages of avoiding cardiopulmonary bypass, consistent benefits in clinical outcome have not been observed. There does seem to be a more or less uniform trend towards decreased perioperative blood product transfusions with OPCAB compared to on-pump CABG. In terms of other measures of early outcome, postoperative renal failure, stroke, and acute MI, the superiority of OPCAB has been unclear.\(^\text{69,72,73}\) A more recent Cochrane review by Moller et al did not demonstrate any significant benefit of off-pump compared with on-pump CABG regarding mortality, stroke, or myocardial infarction. In contrast, better long-term survival in the group of patients undergoing on-pump CABG with the use of cardiopulmonary bypass and cardiopulmonary arrest was observed.\(^\text{74}\)

The higher cardiac morbidity in the ROOBY trial was associated with decreased 1-year angiographic patency rates.\(^\text{71}\) However, studies with contrasting findings exist, quoting equivalent rates of graft patency for OPCAB usage.\(^\text{75,76}\) The broad variety in results may be suggestive that other factors (e.g., surgeon skill, technical difficulty, patient factors) may be dominating the outcome rather than the use or avoidance of cardiopulmonary bypass.\(^\text{77}\) After almost two decades, OPCAB has not been widely adopted and remains less than 20% of all CABG procedures in the United States.

**Minimally Invasive Direct Coronary Artery Bypass.** As an extension of the off-pump coronary revascularization technique, minimally invasive direct coronary artery bypass (MIDCAB) has been described. MIDCAB is performed using a left anterior mini-thoracotomy through which mobilization of the left internal thoracic and direct in situ anastomosis to the left anterior descending artery (or its diagonal branches) is performed. This technique is primarily applicable to single-vessel disease, although reports of multivessel revascularizations do exist.

**MIDCAB Results.** A review of 411 patients undergoing MIDCAB quotes an operative mortality >1%. In this study, all patients received revascularization of the LAD only, regardless of the number of diseased vessels. The 3-year mortality in patients with single-vessel disease following a MIDCAB was 3.1%, which was not surprisingly, lower than those with multivessel disease (8.7%).\(^\text{78}\)

There is an inherent selection bias in retrospective reviews comparing MIDCAB to OPCAB or conventional CABG as MIDCAB patients tend to have less extensive disease. Because of this, there have been multiple randomized controlled trials looking at the efficacy of MIDCAB compared to PCI. A meta-analysis of 12 randomized prospective trials comparing PCI to MIDCAB revascularization of isolated proximal left anterior descending artery demonstrated comparable results in terms of mortality and MI but a lower revascularization requirement in the MIDCAB group.\(^\text{79}\)

A recent meta-analysis by Lee et al revealed CABG, as compared with PCI with DES, reduced long-term rates of the composite of all-cause death, myocardial infarction, or stroke in patients with left main or multivessel CAD. Compared to PCI with DES, CABG was found to be superior in patients with multivessel CAD (\(P = 0.001\)), but no between-group differences in those with left main CAD (\(P = 0.427\)).\(^\text{80}\) Similar conclusions have been made by multiple other studies.

**Total Endoscopic Coronary Artery Bypass.** With the advent of robotic surgical technology allowing stereoscopic visualization and increased instrument dexterity, total endoscopic coronary artery bypass (TECAB) has become possible. In July 2004, the da Vinci robotic surgical system received FDA approval for use in coronary anastomoses. Extracorporeal circulation with peripheral cannulation has been used in earlier reports, but the development of mechanical stabilizers has provided the ability to perform the internal thoracic artery harvest and coronary anastomosis off-pump with use of the robotic arms only. Several studies have looked at the feasibility of TECAB and have shown acceptable results, but this procedure has not been adopted by most surgeons because of its steep learning curve, longer operative times, and lack of demonstrable clinical benefit.\(^\text{81-83}\) Although the volume of robotic-assisted CABG is increasing, such procedures constituted <1% of all CABG procedures performed in the United States in 2012.\(^\text{84}\)

**Hybrid Coronary Revascularization.** With the increasing collaboration between cardiothoracic surgeons and interventional cardiologists, hybrid coronary revascularization (HCR) combining a minimally invasive surgical technique (MIDCAB or TECAB) with PCI has become a reality. This capitalizes on a major advantage of both treatments, utilizing the durable left internal thoracic artery to left anterior descending coronary artery bypass graft while treating other stenoses with PCI, obviating the need for a large surgical incision or cardiopulmonary bypass. HCR is not without its downsides as there are some concerns with this approach because aggressive anti-platelet therapy is required with PCI and may increase the hemorrhagic complications of surgical revascularization. A small study comparing HCR to OPCAB showed comparable graft patency and decreased hospital stay with HCR without an increase in complication rates.\(^\text{85}\) There are, however, some studies that have reported increased rates of requirement for reintervention in patients undergoing HCR, and this requires further study.\(^\text{86,87}\) A recent multicenter prospective observational study showed equivalent outcomes between HCR and multivessel PCI.\(^\text{88}\) HCR has not gained widespread acceptance, and its clinical value remains a matter of debate.

**Transmyocardial Laser Revascularization.** Despite the advancement of technology and revascularization strategies, patients with end-stage coronary artery disease may not be amenable to complete revascularization. Transmyocardial laser revascularization (TMR) relies on a CO\(_2\) or holmium:yttrium-aluminum-garnet (Ho:YAG) laser to create multiple transmural channels (1 mm in diameter) through the myocardium. The initial concept was that these channels would serve as conduits for direct perfusion from the ventricle, but evidence suggests that the resultant angiogenesis is primarily responsible for the improved perfusion. A meta-analysis of seven randomized controlled trials comparing TMR to medical therapy for chronic...
angina has shown higher rates of angina improvement in the TMR but was not able to show a difference in mortality between the two groups.90

TMR is also being used as an adjunct to CABG in the treatment of extensive CAD that is not amenable to surgical revascularization alone. In a study looking at the benefits of TMR in addition to CABG, Allen et al concluded that TMR decreases angina burden when added to CABG in patients who cannot be revascularized by CABG alone.90 The current STS guidelines support the consideration of TMR in patients with ischemic myocardial territories that cannot be revascularized by PCI or CABG.91 Because of equivocal late results at most centers, this therapeutic strategy has not gained widespread acceptance, and ACC/AHA guidelines give TMR a class IIb recommendation for treatment of refractory angina.90

New Developments

Regenerative Medicine and Tissue Engineering. Provocative investigations are being performed on the level of signaling molecules, gene therapy, stem cells, and tissue engineering to regenerate or replace damaged tissue in patients with ischemic heart disease. Growth factors, such as fibroblast growth factor (FGF) and vascular endothelial growth factor (VEGF), are receiving attention due to their ability to induce growth of new vessels. Although concerns regarding systemic administration of these pleiotropic signaling molecules exist, early placebo-controlled clinical trials have shown some promising results with administration of these agents.92,93 Adenoviral transfection of diseased tissue with transgenes for growth factors and transcription factors has been attempted with variable results.

Research in tissue engineering has been directed at creation of vascular conduits that are resistant to atherosclerosis. Stem cells have also been infused directly into the site of injury or in the generation of new tissue around a biodegradable scaffold.94 Despite their potential, these technologies are still in their infancy, and significant progress will be needed before more widespread clinical adoption.

VALVULAR HEART DISEASE

General Principles

The number of patients undergoing surgical management of valvular heart disease has increased over the last decade, from a total of 26,547 isolated aortic or mitral valve procedures reported to the STS Adult Cardiac Surgery Database in 2006 to 45,253 such procedures in 2015.95 In 2016, valve procedures represented over 50% of the cases performed at our institution. Although congenital and inherited etiologies represent important clinical entities, age-associated and acquired conditions still represent the primary causes of valvular heart disease and are the focus of this section.

The most common screening method for valvular heart disease is cardiac auscultation, with murmurs classified based primarily on their timing in the cardiac cycle, but also on their configuration, location and radiation, pitch, intensity, and duration (Table 21-6).96 Although some systolic murmurs are related to normal physiologic increases in blood flow, some may indicate cardiac disease, such as valvular aortic stenosis (AS), that are important to diagnose, even when asymptomatic. Diastolic and continuous murmurs, on the other hand, are frequently

<table>
<thead>
<tr>
<th>Table 21-6 Classification of cardiac murmurs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MURMUR</strong></td>
</tr>
<tr>
<td>------------------------------------------------</td>
</tr>
<tr>
<td><strong>SYSTOLIC MURMURS</strong></td>
</tr>
<tr>
<td>Holosystolic (pansystolic)</td>
</tr>
<tr>
<td>Mid-systolic (systolic ejection)</td>
</tr>
<tr>
<td>Early systolic</td>
</tr>
<tr>
<td>Mid to late systolic</td>
</tr>
<tr>
<td><strong>DIASTOLIC MURMURS</strong></td>
</tr>
<tr>
<td>Early high-pitched</td>
</tr>
<tr>
<td>Mid-diastolic</td>
</tr>
<tr>
<td>Presystolic</td>
</tr>
<tr>
<td><strong>CONTINUOUS MURMURS</strong></td>
</tr>
<tr>
<td>PDA</td>
</tr>
</tbody>
</table>

AI = aortic insufficiency; ASD = atrial septal defect; MR = mitral regurgitation; MS = mitral stenosis; MVP = mitral valve prolapse; PDA = patent ductus arteriosus; PR = pulmonic regurgitation; TI = tricuspid insufficiency; TS = tricuspid stenosis; VSD = ventricular septal defect.
pathologic in nature. Dynamic cardiac auscultation provides further evidence as to the significance and origin of many murmurs (Table 21-7).96

Although auscultation may provide initial evidence to the existence of valvular disease, associated signs and symptoms may help narrow the diagnosis. Abnormalities in the splitting of the heart sounds and additional heart sounds should be noted, as should the presence of pulmonary rales. Peripheral pulses should be checked for abnormal intensity or timing, and the presence of a jugular venous wave should be documented. Additionally, symptoms of syncope, angina pectoris, heart failure, and peripheral thromboembolism are important and may help guide diagnosis and management.

Several imaging examinations are also available to aid in the diagnosis and classification of various valvular disorders. Electrocardiograms may provide information regarding ventricular hypertrophy, atrial enlargement, arrhythmias, conduction abnormalities, prior myocardial infarction, and evidence of active ischemia that would prompt further workup. Posteroanterior and lateral chest X-rays are also easy to obtain and may yield information regarding cardiac chamber size, pulmonary blood flow, pulmonary and systemic venous pressure, and cardiac calcifications. The gold standard for the evaluation of valvular heart disease is transthoracic echocardiography (TTE), which is helpful in the noninvasive evaluation of valve morphology and function, chamber size, wall thickness, ventricular function, pulmonary and hepatic vein flow, and pulmonary artery pressures. Specialized examinations based on the specific findings of TTE examinations are discussed in the following sections.

Regardless of the etiology, valvular heart disease can produce a myriad of hemodynamic derangements. Left untreated, valvular stenosis and insufficiency can produce significant pressure and volume overload on the affected cardiac chamber, respectively, with mixed disease consequently causing mixed pathology. Although the heart can initially compensate for alterations in cardiac physiology, cardiac function eventually deteriorates, leading to heart failure, decreased patient functional status, ventricular dysfunction, and eventually death. In order to optimize long-term survival, surgery or transcatheter therapeutics are recommended in various forms of valvular heart disease and in an increasing number of elderly and high-risk patients.

### Surgical Options

Although valve repair is increasingly indicated, especially in patients with aortic, mitral or tricuspid insufficiency, valve replacement is appropriate in certain patient populations. Valve replacement can be accomplished with either mechanical or biological prostheses, and the choice of valve depends on many patient-specific factors such as age, health status, and desire for future pregnancy. Preexisting indications or contraindications to anticoagulation therapy also influence the choice of mechanical versus tissue valve prosthesis.

Current options for mechanical valve replacement include either tilting disc valves or bileaflet valves. Although mechanical valves are highly durable, they require permanent anticoagulation to mitigate the risk of valve thrombosis and thromboembolic sequelae.97 Due to the concordant risk of hemorrhagic complications, patient characteristics such as debility, lifestyle, and contraindications to systemic anticoagulation therapy may preclude mechanical valve replacement. Moreover, young women who are planning future pregnancies cannot take warfarin due to its teratogenic potential. Conversely, patients with other indications for systemic anticoagulation, such as other risk factors for thromboembolism (i.e., atrial fibrillation), or the presence of a mechanical prosthetic valve in place in another position, may benefit from mechanical valve replacement. Current ACC/AHA guidelines recommend a shared-decision-making process between patient and physician when determining the choice of valve prosthesis, with the use of

<table>
<thead>
<tr>
<th>INTERVENTION</th>
<th>EFFECT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respiration</td>
<td>Right-sided murmurs increase with inspiration. Left-sided murmurs increase with expiration.</td>
</tr>
<tr>
<td>Valsalva maneuver</td>
<td>Most murmurs decrease in length and intensity. The murmur of HCM becomes louder, and the murmur of MVP becomes louder and longer.</td>
</tr>
<tr>
<td>Exercise</td>
<td>Benign flow murmurs and murmurs caused by stenotic valves become louder with isotonic and isometric exercise. The murmurs of MR, VSD, and AI also increase with isometric exercise.</td>
</tr>
<tr>
<td>Positional changes</td>
<td>Most murmurs decrease with standing; the murmur of HCM becomes louder, and the murmur of MVP becomes louder and longer. Brisk squatting and passive leg raising increases most murmurs; the murmurs of HCM and MVP diminish.</td>
</tr>
<tr>
<td>Postventricular premature beat or atrial fibrillation</td>
<td>Benign flow murmurs and stenosis at the semilunar valves increase in intensity following a ventricular premature beat or a long cycle length in atrial fibrillation. Systolic murmurs of ativoventricular valve regurgitation do not change.</td>
</tr>
<tr>
<td>Pharmacologic interventions</td>
<td>The initial hypotensive phase following inhalation of amyl nitrate decreases the murmurs of MR, VSD, and AI, and increases the murmur of AS. The late tachycardic phase following inhalation of amyl nitrate increases right-sided murmurs and the murmur of MS. The response in MVP is biphasic (softer then louder than control).</td>
</tr>
<tr>
<td>Transient arterial occlusion</td>
<td>Transient external compression of the upper extremity increases the murmurs of MR, VSD, and AI.</td>
</tr>
</tbody>
</table>

AI = aortic insufficiency; AS = aortic stenosis; HCM = hypertrophic cardiomyopathy; MR = mitral regurgitation; MS = mitral stenosis; MVP = mitral valve prolapse; VSD = ventricular septal defect.
bioprosthetic valves in all patients who have a contraindication to lifelong anticoagulation or who are unwilling to receive it.98

The potential to avoid the hazards of serious bleeding complications spurred the development of valve prostheses using biological materials, which obviates the need for systemic anticoagulation therapy. As tissue valves are naturally less thrombogenic, the attendant yearly risks of both thromboembolic and anticoagulation-related complications are considerably less than with mechanical valves.99 Consequently, tissue valve replacement is generally recommended for patients averse to systemic anticoagulation therapy, with potential concerns regarding compliance or follow-up while taking anticoagulant medications, and in the case of reoperation for a thrombosed mechanical valve. However, biological valves are more prone to degeneration, especially when implanted in the mitral position, and in younger patients, and those in renal failure, on hemodialysis, or with hypercalcemia.99 Improved manufacturing methods have made currently available tissue valves more durable than previous versions, and valve replacement with a biological prosthesis is generally preferred in patients without other indications for anticoagulation therapy who are >60 years of age for the aortic position and >70 years of age for the mitral position. The availability of transcatheter aortic valves for re-replacement, has resulted in patients and surgeons preferring bioprosthetic aortic valves in patients even <60 years of age in many centers.

**Mechanical Valves.** The first bileaflet mechanical valve was introduced in 1977. Bileaflet valves are comprised of two semi-circular leaflets that open and close, creating one central and two peripheral orifices (Fig. 21-5). Bileaflet mechanical valves have demonstrated excellent flow characteristics, low risk of late valve-related complications, including valve failure, and are currently the most commonly implanted type of mechanical valve prosthesis in the world.100

Although mechanical valves necessitate systemic anticoagulation, careful monitoring of the International Normalized Ratio (INR) reduces the risk of thromboembolic events and hemorrhagic complications and improves overall survival.101 Patients undergoing mechanical aortic valve replacement generally have a target INR of 2 to 3 times normal; however, after a randomized study, one of the mechanical aortic valves has been improved for an INR range of 1.5 to 2.0.102 Patients undergoing mechanical mitral valve replacement frequently have increased left atrial size, concomitant atrial fibrillation, and are at higher risk for thromboembolism that those undergoing aortic valve replacement and are thus recommended to have a target INR 2.5 to 3.5 times normal. When managed appropriately, the yearly risk of major bleeding is <1.4%. Patients with mechanical mitral valve prostheses have nearly twice the thromboembolic risk of those with mechanical aortic valve prostheses (1.3% vs. 0.8% per year).103

**Tissue Valves.** A xenograft valve is one implanted from another species, such as porcine xenograft valves, or manufactured from tissue such as bovine pericardium. A variety of xenograft tissue valves exist and are primarily differentiated by the presence or absence of a mounting stent. Stented valves are the most commonly implanted, and the most popular valve in the United States is a stented bovine pericardial valve.104

The more traditional stented valves are attached to a sewing ring, which decreases the technical complexity of valve replacement compared with stentless valves (Fig. 21-6). The chief disadvantage of stented tissue valves is a smaller effective orifice area, which increases the transvalvular gradient. This phenomenon is referred to as patient prosthetic mismatch. This effect is most pronounced in patients with small prosthetic valve areas, specifically <0.85 cm² valve area per square meter body surface area and may affect survival, symptomatic improvement, and the hemodynamic response to exercise following surgery.105

Stentless porcine xenograft valves were developed in order to minimize the limitations in flow characteristics seen in patients with small prosthetic valve areas and have demonstrated an increase in effective valve area of approximately 10% over stented xenografts of equivalent size.106 They can result in improved hemodynamics, both at rest and with exercise.106 The absence of a stent and sewing ring both increases the technical complexity of valve replacement and takes advantage of the biologic mobility of the aortic valve apparatus. Though results with stentless valves seem promising, some stentless valves have been shown to have poor durability,107 and stentless valves have not been widely adopted due to the technical complexity of implantation.

Recently, rapid deployment valves have been introduced to further decrease the complexity and time required for aortic valve replacement. Two of these valves are on the market in the United States. The Perceval valve by LivaNova is the only sutureless, stentless valve available.108 These rapid deployment valves have shown improved hemodynamics, particularly in patients with small annuli, and shorter implantation times. In some series, these benefits have been shown to reduce early morbidity.108

**Homografts.** Homograft valves from human cadavers, also known as allografts, have been used for aortic valve replacement since the technique was originally described over 50 years ago.109 Since that time, homografts have typically been used for aortic and pulmonary valve replacements and have been successfully harvested from brain dead organ donors and the
explanted hearts of heart transplant recipients. Following harvest, these valves are sterilized using an antibiotic solution and subsequently stored in fixative or cryopreservation solution.

Like other types of tissue valves, the risk of thromboembolic complications with homograft valves is low, and systemic anticoagulation therapy is not required. Additionally, the structure of homograft valves is naturally low-profile, allowing for larger effective valve orifices and lower postoperative transvalvular gradients compared with stented xenograft valves. Additionally, they have been shown to have some advantages in patients with endocarditis.  

The major shortcoming of homograft valves is their limited long-term durability due to tissue degeneration. Within one year of implantation, these valves undergo substantial loss of cellular components and subsequent structural compromise, which may ultimately lead to valve failure. Although enhanced preservation techniques has improved cellular viability, which approaches the 15-year viability of xenograft valves, the limited availability of these valves and techniques has limited the use of homograft tissue valves.

**Autografts.** In 1967, Donald Ross described a procedure in which the diseased aortic valve was replaced using the patient’s native pulmonary valve as an autograft, which was in turn replaced with a homograft in the pulmonic position. The procedure resulted in minimal transvalvular gradients and favorable left ventricular mechanics, both at rest and during exercise. Known as the Ross procedure, this operation has been shown to be particularly beneficial in children, as the pulmonary trunk grows with the child and long-term anticoagulation is not required.

The late results of the Ross procedure are discussed later in this chapter. In addition to potential concerns with durability, performance of the Ross procedure has also been limited by its technical complexity and the increased surgical risk associated with double valve replacement.

**Valve Repair.** Valve repair offers a number of advantages over valve replacement, due in large part to the preservation of the patient’s native valvular and subvalvular apparatus. In mitral valve (MV) surgery, preservation of the mitral apparatus has been shown to lead to better postoperative left ventricular function and survival. Additionally, as there is no implanted prosthesis, the patient avoids the risks of chronic anticoagulation, infection, thromboembolic complications, and prosthetic valve failure after surgery.

In the case of MV repair, freedom from reoperation and valve-related complications has been excellent in certain patient populations, even at 20-year follow-up. It has also been demonstrated that patients undergoing MV surgery with moderate functional tricuspid regurgitation (TR) do not experience increased perioperative complication rates when a concomitant tricuspid valve (TV) repair is performed. Midterm results in this group are encouraging, with greater than 98% freedom from reoperation reported by some groups at 5 years, suggesting further indications for valve repair.

Despite its advantages for the patient, valve repair is generally more technically demanding than valve replacement and may occasionally fail. Both the suitability of the patient for valve repair and the skill and expertise of the surgeon performing the operation are important when considering valve repair in the individual patient.

**MITRAL VALVE DISEASE**

**Mitral Stenosis**

**Etiology.** Acquired mitral stenosis (MS) is most often caused by rheumatic fever, with approximately 60% of patients with pure MS presenting with a clinical history of rheumatic heart disease. Rarely, other conditions can cause obstruction to filling of the left ventricle (LV), mimicking MS. Acquired causes of MV stenosis include left atrial myxoma, prosthetic valve thrombosis, mucopolysaccharidosis, previous chest radiation, and severe annular calcification.

**Pathology.** Although rheumatic heart disease is associated with a transmural pancarditis, pathological fibrosis of the valves.
SPECIFIC CONSIDERATIONS

Figure 21-7. Mitral stenosis. The thickened, fused leaflets of the diseased mitral valve are viewed through a left atriotomy. (Reproduced with permission from Centers for Disease Control and Prevention, Edwin P. Ewing, Jr.)

Results primarily from the endocarditic process. The damage caused by endocardial inflammation and fibrosis is progressive, causing commissural fusion, subvalvular shortening of the chordae tendineae, and calcification of the valvular and subvalvular apparatus. The resulting stenotic MV has a funnel shape, with doming of the leaflets, and a significantly narrowed orifice caused by interchordal and commissural fusion (Fig. 21-7). The degree of mitral stenosis should be determined preoperatively, as these pathological features may help determine the timing and type of intervention to perform.

**Pathophysiology.** As the normal MV area of 4.0 to 5.0 cm\(^2\) is reduced by the rheumatic process, blood can flow from the left atrium to the left ventricle only if it is propelled by an ever-increasing pressure gradient. This increased left atrial pressure causes left atrial enlargement and eventually pulmonary hypertension and decreased exercise tolerance. Patients with diastolic valve doming, usually accompanied by a history of rheumatic fever, are defined as having stage A MS, even with normal transmitral flow velocities. Stage B MS is defined by an increased transmural flow velocity with mitral valve area greater than 1.5 cm\(^2\). This condition is associated with mild to moderate left atrial enlargement but normal pulmonary arterial pressure. Stage C MS consists of a severely stenotic valve (mitral valve area ≤1.5 cm\(^2\)) without symptoms. This is frequently associated with commissural fusion and diastolic doming of the leaflets as well as a pulmonary arterial pressure of >30 mmHg. Stage D MS is defined as the aforementioned criteria with the onset of decreased exercise tolerance and/or dyspnea on exertion (Table 21-8).

The onset of symptoms is due to the evolution of pathophysiological processes, beginning with an elevation in left atrial pressure. The increased left atrial pressure is subsequently transmitted to the pulmonary venous system, causing pulmonary edema as the hydrostatic pressure in the vessels exceeds the plasma oncotic pressure. Decreased pulmonary venous compliance exacerbates the pulmonary venous hypertension, though a concomitant decrease in microvascular permeability may

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean gradient (mmHg)</td>
<td>&lt;5</td>
<td>5–10</td>
<td>&gt;10</td>
</tr>
<tr>
<td>Pulmonary artery systolic pressure</td>
<td>&lt;30</td>
<td>30–50</td>
<td>&gt;50</td>
</tr>
<tr>
<td>Valve area (cm(^2))</td>
<td>&gt;1.5</td>
<td>1.0–1.5</td>
<td>&lt;1.0</td>
</tr>
</tbody>
</table>

**Table 21-8**

Data from ACC/AHA guidelines for the classification of the severity of mitral valve disease in adults

**MITRAL STENOSIS**

**QUALITATIVE**

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angiographic grade</td>
<td>I+</td>
<td>2+</td>
<td>3+</td>
</tr>
<tr>
<td>Color Doppler jet area</td>
<td>Small, central jet (&lt;4 cm(^2) or &lt;20% left atrial area)</td>
<td>More than mild criteria, but no severe criteria present</td>
<td>Vena contracta width &gt;0.7 cm with large central jet (area &gt;40% of left atrial area) or with a wall-impinging jet of any size, swirling in left atrium</td>
</tr>
</tbody>
</table>

**MITRAL REGURGITATION**

**QUANTITATIVE (CATH OR ECHO)**

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regurgitant volume (ml per beat)</td>
<td>&lt;30</td>
<td>30–50</td>
<td>≥60</td>
</tr>
<tr>
<td>Regurgitant fraction (%)</td>
<td>&lt;30</td>
<td>30–49</td>
<td>≥50</td>
</tr>
<tr>
<td>Regurgitant orifice area (cm(^2))</td>
<td>0.2</td>
<td>0.2–0.39</td>
<td>≥0.4</td>
</tr>
</tbody>
</table>

**ADDITIONAL ESSENTIAL CRITERIA**

<table>
<thead>
<tr>
<th>Indicator</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Left atrial size</td>
<td></td>
<td>Enlarged</td>
</tr>
<tr>
<td>Left ventricular size</td>
<td></td>
<td>Enlarged</td>
</tr>
</tbody>
</table>

\(^{a}\)Valve gradients are flow dependent and when used as estimates of severity of valve stenosis should be assessed with knowledge of cardiac output or forward flow across the valve.
preclude pulmonary edema in the chronic setting.119 Patients may also develop pulmonary arterial hypertension, owing to vasoconstriction, intimal hyperplasia, and medial hypertrophy of the pulmonary arterioles in response to the increased pulmonary venous pressure. The secondary obstruction to flow caused by reactive pulmonary arterial hypertension may serve to protect against pulmonary edema, but it also exacerbates the intractable decrease in cardiac output that develops as stenosis worsens.120

Throughout the process, the left atrium becomes dilated and hypertrophied due to increased work in filling the ventricle against a fixed obstruction. Atrial fibrillation (AF) may develop, exacerbating the patient’s symptoms and increasing the risk of left atrial thrombus and subsequent embolization. Left ventricular structure and function are typically preserved owing to the protective effect of the stenotic valve.

Clinical Manifestations. The sudden opening of the thickened, nonpliable valve with left atrial contraction produces an opening snap, followed by a diastolic rumble caused by rapid entry of blood into the left ventricle. When diastole is complete, the MV subsequently closes very rapidly, causing an increased first heart sound. The murmur, classically known as the auscultatory triad, is best heard at the apex. Associated mitral and tricuspid insufficiencies are heard as a pansystolic murmur radiating to the axilla and a systolic murmur at the xiphoid process, respectively.

The first clinical signs of MS are those associated with pulmonary venous congestion, namely exertional dyspnea, decreased exercise capacity, orthopnea, and paroxysmal nocturnal dyspnea. Hemoptyisia and pulmonary edema may develop as the venous hypertension worsens. Advanced MS can also cause pulmonary arterial hypertension and subsequent right heart failure, manifested as jugular venous distention, hepatomegaly, ascites, and lower extremity edema.2

As mentioned previously, atrial fibrillation may develop as left atrial pathology worsens, causing atrial stasis and subsequent thromboembolism. Patients with MS may initially present with signs of arterial embolization, even rarely with angina from coronary occlusion.2

Diagnostic Studies. All patients with a clinical history and physical exam suggestive of MS should have an electrocardiogram (ECG) and chest X-ray. Abnormalities in the ECG may include atrial fibrillation, left atrial enlargement, or right axis deviation. Chest X-ray findings may include enlargement of the left atrium and pulmonary artery, creating a double contour behind the right atrial shadow and obliterating the normal concavity between the aorta and left ventricle. Findings consistent with pulmonary congestion may also be present.2

The diagnostic tool of choice is TTE, which not only confirms the diagnosis of MS, but also rules out other concomitant myocardial or valvular heart disease.121 Two-dimensional TTE can be used to calculate the MV area and to determine the morphology of the MV apparatus, including leaflet mobility and flexibility, leaflet thickness and calcification, subvalvular fusion, and the appearance of the commissures. Doppler TTE can also be used in combination with various equations to estimate the hemodynamic severity of MS in terms of the mean transmural pressure gradient, the MV area, and the pulmonary artery systolic pressure.

In most cases, further examinations are not necessary. A preoperative TEE may be utilized to rule out left atrial appendage thrombus, when the patient is being considered for percutaneous balloon mitral commissurotomy, or if the preoperative TTE is insufficient for diagnosis. Exercise TTE is indicated when resting TTE parameters are discordant with symptom severity.123 Routine coronary angiography is usually performed prior to valve surgery,98 except in young patients (≤30 years of age) with no risk factor for coronary artery disease.

Indications for Operation. Depending on the severity and the morphology of the diseased MV (see Table 21-8), balloon commissurotomy, surgical repair, or MV replacement may be indicated for the treatment of MS (Table 21-9).98

Mitrail Regurgiation

Etiology. The most important cause of MR in the United States is myxomatous degenerative disease of the MV, which occurs in 2% to 3% of the population.123 Other important causes of MR include rheumatic heart disease, infective endocarditis, ischemic heart disease, and dilated cardiomyopathies. Less frequently, MR can be caused by collagen vascular diseases, trauma, previous chest radiation, hypereosinophilic syndrome, carcinoid disease, and exposure to certain drugs.96

Pathology. The MV apparatus consists of the mitral leaflets, chordae tendineae, papillary muscles, and mitral annulus, and abnormalities in any one of these components has the potential to cause MR.124 The system for classifying MR proposed by Carpentier focuses on the functional and anatomic characteristics of the MV pathology and proposes three basic types of diseased valves based on the motion of the free edge of the leaflet relative to the plane of the mitral annulus.125

In Type I MR, valvular insufficiency occurs secondary to annular dilatation or leaflet perforation, and normal leaflet motion is maintained. Type II MR is seen in patients with mitral valve prolapse and is due to prolapse of often thickened excessive leaflet tissue that gives the valve a “billowing” appearance, in addition to ruptured or elongated chordae tendineae causing increased leaflet motion. Type III insufficiency, as seen in patients with rheumatic and ischemic heart disease, occurs from restricted leaflet motion, either during systole and diastole (Type IIIA) or during systole alone (Type IIIB).

Pathophysiology. The basic pathophysiologic abnormality of MR is the retrograde flow of a portion of the LV stroke volume into the left atrium during systole due to an incompetent MV or dilated MV annulus. Acute severe MR can result from ruptured chordae tendineae, a ruptured papillary muscle, or infective endocarditis and causes a sudden volume overload of both the left atrium and ventricle.96 Although an acute increase in preload provides a modest increase in overall stroke volume, the left atrium and ventricle are unable to fully accommodate the regurgitant volume or maintain forward stroke volume in the acute setting due to a lack of remodeling.

Chronic MR generally has a more indolent course, with increasing volume overload of the left atrium and ventricle as the effective regurgitant orifice size becomes larger. The resulting increase in left atrial and ventricular volume initially allows for an increase in the total stroke volume by Starling’s law and accommodation of the regurgitant volume, thus maintaining forward cardiac output and alleviating pulmonary congestion during the compensatory phase of chronic MR.126 However, as the left atrium becomes more dilated, the development of AF becomes more likely, disrupting atrioventricular synchrony and
<table>
<thead>
<tr>
<th>CLINICAL SETTING</th>
<th>CLASS OF RECOMMENDATION</th>
<th>LEVEL OF EVIDENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Balloon Valvotomy for Mitral Stenosis</strong></td>
<td>I</td>
<td>A</td>
</tr>
<tr>
<td>Symptomatic patients (NYHA II, III, IV) with moderate or severe MS and favorable valve morphology, without left atrial thrombus or moderate to severe MR</td>
<td>I</td>
<td>C</td>
</tr>
<tr>
<td>Asymptomatic patients with moderate or severe MS, favorable valve morphology, and pulmonary hypertension (PASP &gt; 50 mmHg at rest, &gt; 60 mmHg with exercise), without left atrial thrombus or moderate to severe MR</td>
<td>I</td>
<td>C</td>
</tr>
<tr>
<td>Symptomatic patients (NYHA III, IV) with moderate or severe MS and favorable valve morphology, who are high risk or not candidates for surgery</td>
<td>IIa</td>
<td>C</td>
</tr>
<tr>
<td>Asymptomatic patients with moderate or severe MS, favorable valve morphology, and new onset atrial fibrillation, without left atrial thrombus or moderate to severe MR</td>
<td>IIb</td>
<td>C</td>
</tr>
<tr>
<td>Symptomatic patients (NYHA II, III, IV) with MV area &gt; 1.5 cm² if there is evidence of hemodynamically significant MS (PASP &gt; 60 mmHg, PAWP ≥ 25 mmHg, mean MV gradient &gt; 15 mmHg during exercise)</td>
<td>IIb</td>
<td>C</td>
</tr>
<tr>
<td>Symptomatic patients (NYHA III, IV) with moderate or severe MS and favorable valve morphology, as an alternative to surgery</td>
<td>IIb</td>
<td>C</td>
</tr>
<tr>
<td>Patients with mild MS</td>
<td>III – Harm</td>
<td>C</td>
</tr>
<tr>
<td>Patients with moderate to severe MR or left atrial thrombus</td>
<td>III – Harm</td>
<td>C</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Surgery for Mitral Stenosis</strong></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptomatic patients (NYHA III, IV) with moderate or severe MS when:</td>
<td>I</td>
<td>B</td>
</tr>
<tr>
<td>Balloon valvotomy is unavailable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Balloon valvotomy is contraindicated due to thrombus or MR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valve morphology is not favorable for balloon valvotomy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symptomatic patients with moderate to severe MS who also have moderate to severe MR</td>
<td>I</td>
<td>C</td>
</tr>
<tr>
<td>Mildly symptomatic patients (NYHA I, II) with severe MS and severe pulmonary hypertension (PASP &gt; 60 mmHg)</td>
<td>IIa</td>
<td>C</td>
</tr>
<tr>
<td>Asymptomatic patients with moderate or severe MS and recurrent embolic events while receiving adequate anticoagulation, when the likelihood of successful MVr is high</td>
<td>IIb</td>
<td>C</td>
</tr>
<tr>
<td>MVr in the setting of mild MS</td>
<td>III – Harm</td>
<td>C</td>
</tr>
<tr>
<td>Closed commissurotomy in the setting of MVr; open commissurotomy should be performed</td>
<td>III – Harm</td>
<td>C</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Surgery for Mitral Regurgitation</strong></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptomatic patients with acute severe MR</td>
<td>I</td>
<td>B</td>
</tr>
<tr>
<td>Symptomatic patients (NYHA II, III, IV) with chronic severe MR without LV dysfunction (LVEF &lt; 0.30) and/or end-systolic dimension &gt; 55 mm</td>
<td>I</td>
<td>B</td>
</tr>
<tr>
<td>Asymptomatic patients with chronic severe MR and mild to moderate LV dysfunction (LVEF 0.30–0.60) and/or end-systolic dimension ≥ 40 mm</td>
<td>I</td>
<td>B</td>
</tr>
<tr>
<td>Asymptomatic patients with chronic severe MR and preserved LV function (LVEF &gt; 0.60, end-systolic dimension &lt; 40 mm), when the likelihood of successful MVr is &gt; 90%</td>
<td>IIa</td>
<td>B</td>
</tr>
<tr>
<td>Asymptomatic patients with chronic severe MR, preserved LV function, and 1) New onset atrial fibrillation, 2) Pulmonary hypertension (PASP &gt; 50 mmHg at rest, &gt; 60 mmHg with exercise)</td>
<td>IIa</td>
<td>C</td>
</tr>
<tr>
<td>Symptomatic patients (NYHA III, IV) with chronic severe MR due to a primary abnormality of the mitral apparatus and severe LV dysfunction (LVEF &lt; 0.30, end-systolic dimension &gt; 55 mm), when the likelihood of successful MVr is high</td>
<td>IIa</td>
<td>C</td>
</tr>
<tr>
<td>Symptomatic patients (NYHA III, IV) with chronic severe MR secondary to severe LV dysfunction (LVEF &lt; 0.30) who remain symptomatic despite optimal medical management for heart failure, including biventricular pacing</td>
<td>IIb</td>
<td>C</td>
</tr>
<tr>
<td>Asymptomatic patients with MR and preserved LV function (LVEF &gt; 0.60, end-systolic dimension &lt; 40 mm), when the likelihood of successful repair is low</td>
<td>III – Harm</td>
<td>C</td>
</tr>
<tr>
<td>Isolated MV surgery in the setting of mild or moderate MR</td>
<td>III – Harm</td>
<td>C</td>
</tr>
</tbody>
</table>

LV = left ventricular; LVEF = left ventricular ejection fraction; MR = mitral regurgitation; MS = mitral stenosis; MV = mitral valve; MVr = mitral valve repair; MVR = mitral valve replacement; NYHA = New York Heart Association; PASP = pulmonary artery systolic pressure; PAWP = pulmonary artery wedge pressure; * = mitral valve repair should be performed when possible in this population.
predisposing to thrombus formation. Additionally, chronic volume overload may lead to LV contractile dysfunction, resulting in impaired ejection and end-systolic volume increases. LV dilatation and elevated LV end-diastolic pressures may also worsen throughout the progression of MR, reducing cardiac output and causing congestion of the pulmonary vasculature. These changes herald LV decompensation and heart failure and often indicate irreversible myocardial injury.

Clinical Manifestations. In cases of acute severe MR, patients are often symptomatic and present with pulmonary congestion and reduced forward stroke volume. In severe cases, patients may present with cardiogenic shock. Because the LV has not remodeled in the acute setting, a hyperdynamic apical impulse may not be present in the precordium. The typical systolic murmur of MR may be holosystolic or absent, with a third heart sound and/or diastolic flow murmur being the only auscultatory findings.

In cases of chronic MR, patients may remain asymptomatic for long periods of time due to the compensatory mechanisms of the remodeled LV. However, once the LV begins to fail, patients become increasingly symptomatic from exertional dyspnea, decreased exercise capacity, orthopnea, and eventually pulmonary hypertension and right heart failure. Physical examination may demonstrate displacement of the LV apical impulse due to cardiac enlargement from chronic volume overload, a third heart sound, or an early diastolic flow rumble. The characteristic auscultatory findings also include an apical systolic murmur, which is variably transmitted to the axilla or the left sternal border, depending on the location of the pathology. As mentioned previously, patients may present with AF due to dilatation of the left atrium. Findings consistent with pulmonary hypertension frequently indicate late-stage disease.

Diagnostic Studies. In the setting of acute heart failure, TTE should be performed and may demonstrate the anatomical location and severity of the MV pathology. However, TTE may underestimate lesion severity due to inadequate views of the color flow jet. In this case, severe MR should be suspected if hyperdynamic systolic function of the LV is visualized, and TEE may be used to confirm the diagnosis and direct repair strategies.

In cases of chronic MR, ECG and chest X-ray are performed to assess rhythm status and the degree of pulmonary vascular congestion. An initial 2D and Doppler TTE should be performed for a baseline estimation of LV and left atrial size, LV systolic function, pulmonary artery pressure, MV morphology, and MR severity. A central color flow jet in the setting of a structurally normal MV on TTE suggests functional MR, which may be due to LV dilatation or tethering of the posterior leaflet in patients with ischemic heart disease. In the setting of organic MR (i.e., rheumatic and degenerative MR), which is suggested by the presence of an eccentric color flow jet and morphological abnormalities in the MV apparatus on TTE, the presence of calcium in the annulus or leaflets, the redundancy of the leaflets, and the anatomy of the MV pathology should be assessed. Follow-up TTE is indicated on an annual or semiannual basis in patients with asymptomatic moderate to severe MR in order to assess changes from baseline parameters and direct the timing of surgery. Any abrupt change in signs or symptoms in a patient with chronic MR is also an indication for TTE examination.

Additional preoperative studies are variably indicated in certain patient populations. Preoperative TEE is indicated in patients with indications for surgery or poor windows on TTE in order to determine the severity and anatomic basis of MR and to evaluate LV systolic function. Preoperative TEE is also indicated in cases when a discrepancy exists between a patient’s functional status and the severity of MR on TTE. It is helpful for preoperative planning when assessing the feasibility of repair in the individual patient. Exercise stress-echoangiography may also be useful to detect LV systolic dysfunction in well-compensated patients, who may not demonstrate a rise in the end-systolic dimension of the heart or a drop in ejection fraction on routine TTE. Coronary angiography should be performed prior to valve surgery in patients with evidence of ischemia, decreased LV systolic function, a history of coronary artery disease or coronary risk factors, including postmenopausal status and age ≥40 in men and premenopausal women.

Indications for Operation. Based on the etiology, morphology, and severity of MR (see Table 21-8), MV repair, MV replacement with preservation of part or all of the mitral apparatus may be variably performed for the treatment of MR. As the intraoperative findings may dictate MV replacement whenever a MV repair is planned, current recommendations are for MV surgery in general (see Table 21-9).

Mitral Valve Operative Techniques and Results

Mitral valve surgery is performed on the arrested heart with the assistance of cardiopulmonary bypass. Traditionally, a median sternotomy incision has been used; however, the left atrium can also be approached via minimally invasive incisions, such as a right thoracotomy, or a fully endoscopic approach. The MV is commonly exposed through a left atrial incision placed posterior and parallel to the intra-atrial groove or via a transseptal approach through the right atrium.

Commissurotomy. Upon opening the left atrium, the MV is visualized and the left atrium is examined for thrombus. A nerve hook or right-angle clamp is subsequently introduced beneath the commissures and used to evaluate the MV apparatus for leaflet mobility, commissural fusion, and subvalvular chordal abnormalities. The commissure is then carefully incised in a slightly anterior direction 2 to 3 mm at a time, making sure with each extension of the incision that the chordae tendineae remain attached to the commissural leaflets. The commissurotomy is generally stopped 1 to 2 mm from the annulus where the leaflet tissue thins, indicating the transition to normal commissural tissue. The papillary muscles are subsequently examined and incised as necessary in order to maximize the mobility of the leaflets.

After the commissurotomy is complete and the associated chordae tendineae and papillary muscles are mobilized, leaflet mobility is assessed. The anterior leaflet is grasped with forceps and brought through its complete range of motion. If subvalvular restriction or leaflet rigidity is identified, further division or excision of fused chordae and debridement of calcium may be necessary. Occasionally, the leaflets can be debrided carefully to increase mobility. In rheumatic patients, the thickened leaflets can be thinned by careful dissection. Valve replacement may be more appropriate if extensive secondary mobilization is required. At the end of the procedure, competence of the valve is assessed with injection of cold saline into the ventricle.

Open surgical commissurotomy has an operative risk of <1%, and has been shown to have good long-term results, with freedom from reoperation as high as 88.5%, 80.3%, and 78.7% at 10, 20, and 30 years, respectively. The incidence of
postoperative thromboembolic complications is generally <1% per patient-year, and the lack of required systemic anticoagulation precludes the development of hemorrhagic complications long term.\textsuperscript{132} In some institutions, balloon valvuloplasty has replaced commissurotomy.

**Mitrval Valve Replacement.** After exposing the valve, an incision is made in the anterior mitral leaflet, in the midline. The chordal attachments are preserved if possible, with leaflet tissue being excised as needed. Attempts are made to preserve the annular and subvalvular apparatus when possible. If it is necessary to excise the anterior leaflet and chordae, the papillary muscles can be reattached to the annulus with PTFE suture. If possible, the posterior leaflet along with its associated subvalvular structures are preserved. The annulus is subsequently sized, and an appropriate mitral prosthesis is implanted using pledged horizontal mattress sutures. The annular sutures may be placed from the atrial to the ventricular side, seating the valve intra-annularly, or from the ventricular to the atrial side, seating the valve in a supra-annular position. When placing the mattress sutures, care must be taken to stay within the annular tissue, as excessively deep bites may cause injury to critical structures such as the circumflex coronary artery posterolaterally, the atrioventricular node anteromedially, or the aortic valve anterolaterally. The sutures are subsequently placed through the sewing ring, and the valve prosthesis is lowered onto the annulus, where it is secured (Fig. 21-8).

The factors associated with increased operative risk for MV replacement include age, left ventricular systolic dysfunction, emergent procedure status, NYHA functional status, previous cardiac surgery, associated coronary artery disease, and concomitant disease in another valve. However, for most patients, MV replacement is associated with an operative mortality between 2% to 6%, and 65% to 70% 5-year survival.\textsuperscript{133,134} Although preservation of the mitral apparatus during MV replacement is important for subsequent left ventricular function, there was no difference between complete and partial preservation with respect to 30-day and 5-year mortality.\textsuperscript{133} Mechanical valves are associated with increased durability compared to bioprosthetic valves, and they have demonstrated a freedom from reoperation of 98% vs. 79% at 15 years, respectively.\textsuperscript{135} Despite these findings, the choice of prosthetic valve depends on many factors and should be decided on a patient-by-patient basis.

**Mitrval Valve Repair.** There are many techniques available for MV repair that are variably used depending on the preoperative and intraoperative assessment of valvular pathology. On opening the atrium, the endocardium is examined for a jet lesion, a roughened area caused by a regurgitant jet striking the wall, in order to better localize the area of valvular insufficiency. The commissures are examined for evidence of prolapse, fusion, and malformation. The subvalvular apparatus and individual leaflets are subsequently examined, and areas of prolapse, restriction, fibrosis, and calcification are identified. Leaflet perforations are generally repaired primarily, or with a pericardial patch. The degree of annular dilation is also noted. The basic components of MV repair based on this assessment may include resection of the posterior and/or anterior leaflet, chordal shortening, chordal transposition, artificial chordal replacement, and annuloplasty. Recent trends have been toward leaflet preservation.

One of the mainstays of MV repair is triangular resection of the posterior leaflet. Excision of the diseased leaflet tissue extends downwards towards but generally not to the mitral annulus. After repair has been completed, valvular competency is evaluated by injecting saline into the ventricle with a bulb syringe and assessing leaflet mobility and apposition. If focal insufficiency is identified in other areas, additional procedures are performed.

The anterior leaflet may be repaired via chordal shortening, chordal transposition, artificial chordal replacement, and triangular resection of the anterior leaflet. Chordal shortening has generally been abandoned in favor of chordal replacement. During chordal transposition, a resected portion of the posterior leaflet with attached chordae is transposed onto the prolapsed portion of the anterior leaflet to provide structural support, and this is followed by posterior leaflet repair, as described previously. The procedure of artificial chordal replacement uses polytetrafluoroethylene sutures to attach the papillary muscle to the free edge of the prolapsing anterior leaflet. Triangular resection with primary repair of the anterior leaflet removes the prolapsing segment of the anterior MV leaflet, while preserving adjacent chordal tissue, and may be especially helpful in patients with a ruptured chord or large amount of redundant anterior leaflet tissue.

Annular dilation is generally corrected using a MV annuloplasty device, such as a ring or partial band. Annuloplasty is known to improve the durability of all types of MV repair (Fig. 21-9).\textsuperscript{138} A number of devices are available and include rigid or semirigid rings that geometrically remodel the annulus, flexible rings or bands that restrict annular dilation while maintaining the physiologic sphincter motion of the annulus, and semirigid bands that provide a combination of annular remodeling and support of physiologic motion.

Another technique known as the “double-orifice” or “edge-to-edge” repair was introduced by Alfieri in 1995, and it involves tacking the free edge of the anterior leaflet to the opposing free edge of the posterior leaflet.\textsuperscript{137} This procedure effectively gives the valve a double-orifice “bow tie” configuration, and it has been used as both a primary repair technique and an adjunct to other repair techniques, usually in cases of anterior leaflet pathology or Barlow’s disease. While some groups report excellent late results, its use remains controversial, and it is used mainly as a bail-out procedure of last resort in many centers.

![Figure 21-8. Mitral valve replacement. A St. Jude bileaflet mechanical valve is viewed through a left atriotomy.](image-url)
ACQUIRED HEART DISEASE

CHAPTER 21

Figure 21-9. Mitral valve repair. The narrow arrow indicates the posterior leaflet repair, and the wide arrow indicates the ring annuloplasty as viewed through a left atriotomy.

Due to the variety in operations and etiologies of MV disease, there is heterogeneity in outcomes following MV repair. In general, the operative risk for elective, younger patients undergoing MV repair is <1%, and late results across a broad range of patients have demonstrated benefits in survival and valve-related complications, such as thromboembolic events, infective endocarditis, and anticoagulation-related hemorrhage, compared to MV replacement. Patients with MR due to degenerative disease have especially encouraging outcomes, demonstrating rates of survival and freedom from reoperation of >50% and >94% at 20 years, respectively. Historically, isolated anterior leaflet prolapse increased the risk of reoperation fivefold in this population. However, increasing experience and the expanded use of chordal replacement has greatly improved these results in recent series. Independent predictors of mortality have included higher NYHA class, lower left ventricular ejection fraction, renal dysfunction, and age. Older patients have demonstrated slightly worse outcomes overall, with an operative mortality of approximately 4%, and a 10-year survival of 54% in patients ≥65 years of age. However, the superiority of repair over replacement persists even for patients >80 years of age.

Patients with rheumatic disease have demonstrated slightly worse outcomes, with one study showing significantly better freedom from operation at 10 years in patients with non-rheumatic MV disease (88% vs. 73%, P < 0.005). Similarly, in patients with MR secondary to myocardial ischemia, there is growing recent evidence that mitral valve replacement may be significantly more durable than repair. Despite these differences in outcomes, MV repair remains the procedure of choice for the majority of patients with amenable MV disease.

Transcatheter Mitral Valve Repair and Replacement. Since the successful introduction of transcatheter aortic valve replacement, efforts have been made to translate the lessons learned to treatment of the mitral valve. The first device to receive FDA approval for repair of severe mitral regurgitation due to degenerative mitral disease was the MitraClip (Abbott, Abbott Park, IL), introduced in 2003. This device allows a surgeon or cardiologist to grasp the anterior and posterior leaflets of the mitral valve together, approximating the Alfieri double-orifice repair technique. Transcatheter mitral valve repair is now in clinical use in patients with chronic severe primary MR in whom surgery would be too great a risk, as judged by a heart team approach including a cardiologist skilled in structural heart intervention and an experienced mitral valve surgeon (see Table 21-9). In a recent randomized controlled trial comparing transcatheter mitral valve repair to open surgical mitral valve repair, 20.1% (n = 37) of patients who received transcatheter mitral valve repair (n = 184) underwent second intervention with surgical mitral valve repair or replacement within 12 months compared with 2% (n = 2) of patients who underwent surgical repair (n = 95) (P < 0.001). Surgical repair was associated with higher rates of blood transfusion and mechanical ventilation >48 hours but had otherwise equivalent safety. Other transcatheter techniques, including transcatheter mitral valve replacement, have begun clinical trials. This is an area of active research; however, at present, open surgical repair or replacement remains the standard of care for most patients.

AORTIC VALVE DISEASE

Aortic Stenosis

Etiology. The most common cause of adult aortic stenosis (AS) is calcification of a normal trileaflet or congenital bicuspid aortic valve, particularly in patients >70 years of age. Another important cause of AS is rheumatic heart disease, which is particularly common in developing countries (Fig. 21-10).

Pathology. Calcific aortic valve disease, also known as senile or degenerative disease, is an age-related disorder characterized by lipid accumulation, proliferative and inflammatory changes, upregulation of angiotensin-converting enzyme activity, oxidative stress, and infiltration of macrophages and T lymphocytes. This process, which closely resembles atherosclerotic vascular calcification, initially results in bone formation within the base of the cusps, reducing leaflet motion. Calcification progresses to involve the leaflets, and eventually results in obstructive

Figure 21-9. Mitral valve repair. The narrow arrow indicates the posterior leaflet repair, and the wide arrow indicates the ring annuloplasty as viewed through a left atriotomy.

Figure 21-10. Aortic stenosis. The aorta has been removed to demonstrate the thickened, fused aortic valve leaflets associated with rheumatic heart disease. (Reproduced with permission from Centers for Disease Control and Prevention, Edwin P. Ewing, Jr.)
Table 21-10
Data from ACC/AHA guidelines for the classification of the severity of aortic valve disease in adults

<table>
<thead>
<tr>
<th>INDICATOR</th>
<th>MILD (mmHg)</th>
<th>MODERATE (mmHg)</th>
<th>SEVERE (mmHg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jet velocity (m per s)</td>
<td>&lt;30</td>
<td>3.0–4.0</td>
<td>&gt;40</td>
</tr>
<tr>
<td>Mean gradient (mmHg)(^a)</td>
<td>&lt;25</td>
<td>25–40</td>
<td>&gt;4.0</td>
</tr>
<tr>
<td>Valve area (cm(^2))</td>
<td>&gt;1.5</td>
<td>1.0–1.5</td>
<td>&lt;1.0</td>
</tr>
<tr>
<td>Valve area index (cm(^2) per m(^2))</td>
<td></td>
<td></td>
<td>&lt;0.6</td>
</tr>
</tbody>
</table>

AORTIC STENOSIS

<table>
<thead>
<tr>
<th>QUALITATIVE</th>
<th>MILD</th>
<th>MODERATE</th>
<th>SEVERE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angiographic grade</td>
<td>1+</td>
<td>2+</td>
<td>3–4+</td>
</tr>
<tr>
<td>Color Doppler jet width</td>
<td>Central jet, width &lt;25% of left ventricular outflow tract</td>
<td>Greater than mild, but no signs of severe regurgitation</td>
<td>Central jet, width &gt;65% of left ventricular outflow tract</td>
</tr>
<tr>
<td>Doppler vena contracta width (cm)</td>
<td>&lt;0.3</td>
<td>0.3–0.6</td>
<td>&gt;0.6</td>
</tr>
</tbody>
</table>

AORTIC REGURGITATION

<table>
<thead>
<tr>
<th>QUANTITATIVE (CATH OR ECHO)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Regurgitant volume (ml per beat)</td>
<td>&lt;30</td>
<td>30–59</td>
<td>≥60</td>
</tr>
<tr>
<td>Regurgitant fraction (%)</td>
<td>&lt;30</td>
<td>30–49</td>
<td>≥50</td>
</tr>
<tr>
<td>Regurgitant orifice area (cm(^2))</td>
<td>&lt;0.1</td>
<td>0.1–0.29</td>
<td>≥0.3</td>
</tr>
</tbody>
</table>

ADDITIONAL ESSENTIAL CRITERIA

<table>
<thead>
<tr>
<th>Left ventricular size</th>
<th>Enlarged</th>
</tr>
</thead>
</table>

\(^a\)Valve gradients are flow dependent and when used as estimates of severity of valve stenosis should be assessed with knowledge of cardiac output or forward flow across the valve.

disease, with a reduced effective valve area without signs of leaflet fusion.

Pathophysiology. In general, once moderate AS is present, the average rate of progression includes an increase in jet velocity of 0.3 m/s per year, an increase in mean pressure gradient of 7 mmHg per year, and a decrease in valve area of 0.1 cm\(^2\) per year (Table 21-10).\(^{108}\) In most adult patients with AS, obstruction develops gradually and includes a long latent period free from symptoms. During this time, the LV typically hypertrophies in response to systolic pressure overload, and normal intracavitary volume is maintained.\(^{146}\) Afterload, which is defined as left ventricular systolic wall stress, and thus ejection fraction remain normal early in this process as the increase in myocardial thickness is usually enough to counter the increased intracavitary systolic pressures. Patients without a typical hypertrophic response to systolic pressure overload or with a depressed contractile state of the myocardium do not follow the common clinical course, but they experience an early decrease in ejection fraction due to excessively increased afterload without a compensatory response.\(^{147}\)

Concentric LV hypertrophy without chamber dilatation eventually leads to increased end-diastolic pressures and diastolic dysfunction. Forceful atrial contraction in the face of elevated end-diastolic pressures becomes an important component of ventricular filling, even as mean left atrial and pulmonary venous pressures remain in the normal range. Disorders such as atrial fibrillation that disrupt atrial contraction can lead to clinical deterioration. Although systolic function is generally preserved long into the natural history of the disease, left ventricular decompensation eventually occurs in the setting of longstanding increased afterload and is an indication for surgery even in the absence of other symptoms.

Although concentric hypertrophy is a compensatory mechanism to maintain ejection fraction in the face of high intracavitary pressures, the hypertrophied heart becomes increasingly vulnerable to ischemic injury. Coronary blood flow may become inadequate, despite the absence of epicardial coronary artery disease.\(^{148}\) Coronary vasodilation is mitigated by the hypertrophied myocardium, and the hemodynamic stress of exercise or tachyarrhythmias can lead to subendocardial ischemia and further systolic or diastolic dysfunction. When ischemic insults occur, patients with ventricular hypertrophy experience larger infarcts and higher mortality rates than those without hypertrophy.\(^{149}\) In some patients, ventricular hypertrophy occurs in excess of what is needed to compensate for increased intracavitary pressures, creating a high-output state that is also associated with increased perioperative morbidity and mortality.\(^{150}\)

Clinical Manifestations. The characteristic auscultatory findings of AS include a harsh, crescendo-decrescendo systolic murmur at the right second or third intercostal space, often with radiation to the carotid arteries.\(^2\) As the disease progresses, aortic valve closure may follow pulmonic valve closure, causing paradoxical splitting of the second heart sound. Other physical findings associated with AS include an apical impulse commonly described as a “prolonged heave,” and the presence of a narrow and sustained peripheral pulse, known as pulsus parvus et tardus.

The classic symptoms of AS are exertional dyspnea, angina, and syncope.\(^2\) Although many patients are diagnosed prior to the onset of symptoms, the most common clinical
presentation in patients with a known diagnosis of AS is worsening exertional dyspnea. Angina occurs in over half of patients with AS and is due to the increased oxygen demand of the hypertrophied myocardium in the setting of reduced oxygen supply secondary to coronary compression. Although some patients may have concomitant coronary disease, angina occurs without significant epicardial coronary artery disease in half of all patients with AS.\textsuperscript{153} Syncope is most common during exertion, as systemic vasodilation in the setting of a fixed cardiac output causes decreased cerebral perfusion. However, at times, it may occur at rest secondary to paroxysmal atrial fibrillation and subsequent loss of atrial booster pump function. Late findings of AS include atrial fibrillation, pulmonary hypertension, systemic venous hypertension, and, rarely, sudden death.

**Diagnostic Studies.** Evidence of LV hypertrophy is found in approximately 85\% of patients with AS on routine ECG, though the correlation between the absolute electrocardiographic voltages in precordial leads and the severity of AS is poor.\textsuperscript{2} ECG also may demonstrate signs of left atrial enlargement and various forms and degrees of atrioventricular and intraventricular block due to calcific infiltration of the conduction system. Routine chest X-ray usually demonstrates a normal heart size, with rounding of the left ventricular border and apex. A late finding on chest X-ray is cardiac enlargement, or cardiomegaly, a sign of LV failure.

Transesophageal echocardiography is indicated in all patients with a systolic murmur graded ≥2/6, a single second heart sound, or symptoms characteristic of AS.\textsuperscript{98} Initial TTE examinations are often diagnostic and provide an assessment of left ventricular size and function, the degree of left ventricular hypertrophy, the degree of valvular calcification, and the presence of other associated valvular disease. Doppler evaluation should be performed to define the maximum jet velocity, which is a useful measure for following disease severity and predicting clinical outcome.\textsuperscript{2} Additionally, color flow Doppler assesses the severity of the stenotic lesion by allowing calculation of the mean transvalvular pressure gradient and effective valve orifice area (see Table 21-10).\textsuperscript{98} Follow-up TTE is indicated depending on the severity of AS in order to assess changes from baseline parameters and direct the timing of surgery: yearly for severe AS; every 1 to 2 years for moderate AS; and every 3 to 5 years for mild AS. Any abrupt change in signs or symptoms in a patient with AS is an indication for TTE examination.

Additional preoperative studies may be necessary in some patients. Rarely, when TTE images are suboptimal, TEE or fluoroscopy may be indicated to assess the degree of valve calcification and effective valve orifice area. As in other patients with valvular heart disease, coronary angiography should be performed prior to aortic valve surgery in most patients.\textsuperscript{2} Since the symptoms of AS often mimic those of ischemic heart disease, left heart catheterization and coronary angiography may be necessary at the initial evaluation in patients with AS. Stress echocardiography may also be useful in the asymptomatic patient with AS in order to elicit exercise-induced symptoms or abnormal blood pressure responses during exertion. It is also useful in the evaluation of low-gradient AS in patients with depressed LV function.\textsuperscript{98} However, exercise stress-echocardiography is contraindicated in patients with ischemic heart disease.\textsuperscript{98} In patients with evidence of aortic root disease by TTE, chest computed tomography is useful in evaluating aortic dilatation at several anatomic levels and is necessary for clinical decision making and surgical planning.\textsuperscript{2}

**Indications for Operation.** Based on the severity of AS (see Table 21-10) and the predicted risk with surgical aortic valve replacement (SAVR) determined using the STS risk calculator, SAVR or transcatheter aortic valve replacement (TAVR) may be recommended for the treatment of AS (Table 21-11).\textsuperscript{98,152} As this field is rapidly evolving, attention to guideline updates and a multidisciplinary heart team approach to risk stratification and treatment selection are mandatory. In patients with severe calcific AS, AVR via either approach is the only effective treatment, though controversy exists as to the timing of intervention in asymptomatic patients. Balloon aortic valvuloplasty creates a modest hemodynamic effect and temporary symptom improvement in patients with calcific AS. However, the procedure has not been shown to affect long-term outcomes and is often used in hemodynamically unstable patients as a bridge to AVR.\textsuperscript{98}

**Aortic Insufficiency**

**Etiology.** The most common cause of isolated aortic insufficiency (AI) in patients undergoing AVR is aortic root disease, and it represents over 50\% of such patients in some studies.\textsuperscript{2} Other common causes of AI include congenital abnormalities of the aortic valve such as bicuspid aortic valve, calcific degeneration, rheumatic disease, infective endocarditis, systemic hypertension, myxomatous degeneration, dissection of the ascending aorta, and Marfan syndrome. Less common causes of AI include traumatic injuries to the aortic valve, ankylosing spondylitis, syphilitic aortitis, rheumatoid arthritis, osteogenesis imperfecta, giant cell aortitis, Ehlers-Danlos syndrome, Reiter’s syndrome, discrete subaortic stenosis, and ventricular septal defects with prolapse of an aortic cusp.\textsuperscript{96} Although most of these lesions produce chronic aortic insufficiency, rarely acute severe aortic regurgitation can result, often with devastating consequences.

**Pathology.** Regardless of its cause, AI produces volume overload with dilatation and hypertrophy of the left ventricle and subsequent dilatation of the MV annulus. Depending on the severity of AI, the left atrium may undergo dilatation and hypertrophy as well. Frequently, the regurgitant jet causes endocardial lesions at the site of impact on the left ventricular wall.

Diseases causing AI can be classified as primary disorders of the aortic valve leaflets and/or disorders involving the wall of the aortic root. Diseases causing dilatation of the ascending aorta are a more common indication for AVR due to isolated AI, and they include disorders such as age-related (degenerative) aortic dilatation, cystic medial necrosis of the aorta as is seen in Marfan syndrome, aortic dilatation secondary to bicuspid valves, and aortic dissection, to name a few.\textsuperscript{153} In these disorders, the aortic annulus becomes dilated, causing separation of the valve leaflets and subsequent AI. The diseased aortic wall may dissect secondarily and further escalate regurgitation across the valve, and secondary thickening and shortening of the valve cusps may occur due to undue tension placed on the valvular apparatus by the dilated aortic root. As the disease progresses, the valve leaflets become too small to close the aortic orifice, causing further aortic insufficiency and exacerbating dilatation of the ascending aorta.

There are also many primary valvular diseases that cause AI, generally in association with AS. One such disorder is age-related calcific AS, which causes some degree of AI in up to 75\% of patients.\textsuperscript{154} Infective endocarditis may involve the aortic valve apparatus and cause AI through direct destruction of the valve leaflets, perforation of a leaflet, or formation of vegetations that interfere with proper coaptation of the valve.
Pathophysiology. The basic pathophysiologic abnormality of AI is the retrograde flow of a portion of the LV stroke volume into the LV during diastole, producing volume overload. Acute severe AI results most commonly from infective endocarditis, acute aortic dissection, or trauma, and it causes a sudden volume overload of the left ventricle. Although an acute increase in preload provides a small increase in overall
stroke volume due to the Starling mechanism, the left ventricle often is unable to accommodate the large regurgitant volume and maintain forward stroke volume in the acute setting due to a lack of remodeling. Left ventricular end-diastolic and left atrial pressures increase dramatically as the LV is unable to develop compensatory chamber dilation. Although tachycardia develops as a compensatory mechanism to maintain forward flow, this attempt is often inadequate, and patients frequently present in heart failure and even cardiogenic shock. Moreover, subendocardial myocardial ischemia frequently develops as a result of decreased coronary diastolic perfusion pressures and increased LV end-diastolic pressure, as well as increased myocardial oxygen demand due to acute dilation. In the setting of a chronic ventricular hypertrophy and preexisting diastolic dysfunction, the pressure-volume relationship is even more extreme, exacerbating the hemodynamic derangements seen in acute AI.

Chronic AI generally has a more indolent course, with volume overload of the LV causing compensatory increases in left ventricular end-diastolic volume and chamber compliance as well as a combination of eccentric and concentric hypertrophy. Compensatory remodeling of the LV allows for accommodation of the regurgitant volume without a significant increase in filling pressures and maintains the preload reserve of the chamber. Eccentric left ventricular hypertrophy develops, permitting normal contractile performance across the enlarged chamber circumference and subsequent ejection of a larger total stroke volume in order to maintain forward flow, despite the regurgitant fraction. However, the enlarged chamber size results in an increase in systolic myocardial wall stress and causes further ventricular hypertrophy. As the disease progresses, recruitment of preload reserve and compensatory hypertrophy maintains ejection fraction within the normal range despite elevated afterload, causing many patients to remain asymptomatic throughout the compensatory phase.

Eventually, left ventricular compensatory mechanisms fail, and systolic dysfunction ensues. As the disease progresses, preload reserve may become exhausted, the hypertrophic response may become inadequate, and impaired myocardial contractility may develop so that ejection fraction begins to decline. Although left ventricular systolic dysfunction related to excessive afterload is reversible early in the course, irreversible damage occurs once chamber enlargement predominates as the primary cause of diminished myocardial contractility.

Clinical Manifestations. In cases of acute severe AI, patients are symptomatic and invariably present with compensatory tachycardia, often associated with acute pulmonary congestion and cardiogenic shock. Because the left ventricular and aortic pressures often equalize before the end of diastole, the diastolic murmur of AI may be short and/or soft. The reduced systolic pressure may attenuate the increase in peripheral pulse pressure seen in chronic AI, and early closing of the mitral valve due to elevated left ventricular end-diastolic pressures may diminish the intensity of the first heart sound in the acute setting.

In patients with chronic AI, symptoms of heart failure and myocardial ischemia develop after the compensatory phase. Patients gradually begin to complain of exertional dyspnea, fatigue, orthopnea, and paroxysmal nocturnal dyspnea, often after significant myocardial dysfunction has developed. Angina is a common complaint late in the course, especially during sleep when heart rate slows and arterial diastolic pressure falls. Patients also may experience exertional angina secondary to diminished coronary perfusion in the setting of myocardial hypertrophy. Occasionally, the compensatory tachycardia that develops with chronic AI will cause palpitations, and the increased pulse pressure will cause a sensation of pounding in the patient’s head. Peripherally, the widened pulse pressure causes a forceful, bounding, and quickly collapsing pulse known as Corrigan’s water-hammer pulse. Premature ventricular contractions have been reported to cause particularly troubling symptoms, owing to the heave of the volume-loaded left ventricle during the postextrasystolic beat. The classic auscultatory finding associated with AI is a high-pitched decrescendo diastolic murmur heard best in the left third intercostal space; an associated S gallop is often indicative of late disease. The Austin Flint murmur has also been described, and it is heard as a middiastolic rumble at the apex that simulates mitral stenosis and occurs in severe AI when the regurgitant jet impedes mitral valve leaflet opening.

Diagnostic Studies. In the acute setting, TTE should be performed to confirm the presence and severity of aortic regurgitation, the degree of pulmonary hypertension, and the cause of valvular dysfunction. When aortic dissection is suspected as the cause of acute AI, TEE or chest CT angiography may be substituted if more readily available. In these patients with confirmed aortic dissection, cardiac catheterization and coronary angiography are rarely indicated and can delay life-saving urgent surgical intervention.

In cases of chronic AI, the ECG frequently demonstrates left axis deviation and, late in the course, intraventricular conduction defects associated with left ventricular dysfunction. On chest X-ray, the left ventricle enlarges predominantly in an inferior and leftward direction, causing marked increase in the long axis diameter of the heart, frequently with little or no change in the transverse diameter. The chest X-ray should be examined for aneurysmal dilation of the aorta. An initial TTE should be performed to confirm the diagnosis and severity of AI, assess the cause of AI (including valve morphology and aortic root size and morphology), and assess the degree of left ventricular hypertrophy, volume, and systolic function. Follow-up TTE is indicated on an annual or semiannual basis in patients with asymptomatic moderate to severe AI in order to assess changes from baseline parameters and direct the timing of surgery. Any abrupt change in signs or symptoms in a patient with chronic AI is an indication for TTE examination.

Additional preoperative studies are indicated in certain patient populations. In patients with poor windows on TTE, TEE or MRI is indicated for initial and serial assessment of AI severity and left ventricular volume and function at rest. In symptomatic patients with chronic AI, it is reasonable to proceed directly to TEE or cardiac catheterization if TTE examinations are inadequate. Exercise stress testing may be helpful for an assessment of functional capacity and symptomatic responses in patients with equivocal symptoms. Coronary angiography should be performed prior to valve surgery in most patients.

Indications for Operation. Based on the morphology and severity of valve dysfunction (see Table 21-10), AV repair or replacement may be performed for the treatment of AI (see Table 21-11). Although the indications for AV repair and AV replacement do not differ, it is recommended that AV repair be performed only in those surgical centers that have developed the appropriate technical expertise, gained experience in patient selection, and which have demonstrated outcomes equivalent to those of valve replacement.
**Aortic Valve Operative Techniques and Results**

Aortic valve surgery has traditionally been performed through a median sternotomy with the assistance of cardiopulmonary bypass. However, minimally invasive incisions for aortic valve surgery have been introduced, including mini-sternotomy and mini-thoracotomy approaches. After the aorta is cross-clamped, cold blood cardiopulgia is delivered antegrade through the aortic root and/or retrograde through the coronary sinus.

**Aortic Valve Replacement.** During aortic valve replacement, an aortotomy is performed, extending medially from approximately 1 to 2 cm above the right coronary artery and inferiorly into the noncoronary sinus. The valve is completely excised. The annulus is thoroughly debrided of calcium deposits. At this point, the annulus is sized and an appropriate prostheses is selected. Pledged horizontal mattress sutures are then placed into the aortic valve annulus and subsequently through the sewing ring of the prosthetic valve, taking care to avoid damage to the coronary ostia, the conduction system, and the MV apparatus. The annular sutures may be placed from below the annulus, seating the valve supra-annularly, or from above the annulus for intra-annular placement (Fig. 21-11).

The major components to increased operative risk associated with surgical AVR include age, body surface area, diabetes, renal failure, hypertension, chronic lung disease, peripheral vascular disease, neurologic events, infectious endocarditis, previous cardiac surgery, myocardial infarction, cardiogenic shock, NYHA functional status, and pulmonary hypertension. For most patients, the risk of operative mortality associated with AVR is 1% to 5%, and 5-year survival has been reported to be >80%, even in patients >70 years of age. The choice of valve is dependent on many patient-related factors, and it is accompanied by the attendant postoperative risks of decreased durability and thromboembolic vs. hemorrhagic complications for biologic and mechanical valves, respectively.

**Aortic Valve Repair.** Although aortic valve replacement is performed more commonly, AV repair may be recommended at centers of excellence.

For patients with aortoannular ectasia, AI is due to annular dilatation and distortion of the sinotubular junction. For these patients, competence of the aortic valve can be achieved by functionally repairing the annulus in a method analogous to homograft implantation. The aneurysmal portion of the aortic root is excised, and the aortic valve is reimplemented inside a tubular Dacron graft, with concomitant reimplantation of the coronary arteries. Alternatively, the aneurysmal tissue and supravalvular tissue can be excised in their entirety, with subsequent implantation of the Dacron graft onto the superior aspect of the annulus and reimplantation of the coronary arteries.

Valve-sparing root replacement for root and annular stabilization in patients with AI due to aortoannular ectasia has led to a more durable outcome than is seen with subcommissural annuloplasty or leaflet-related procedures alone. One study demonstrated equivalent overall survival between patients undergoing subcommissural annuloplasty or aortic valve repair without annuloplasty and patients undergoing valve-sparing root replacement at 6 years. However, patients who underwent valve-sparing root replacement had higher freedom from reoperation and aortic insufficiency >2+ (100% vs. 90%, P = 0.03; and 100% vs. 77%, P = 0.002, respectively) at midterm follow-up.

For patients with AI associated with redundant leaflet tissue, aortic valve repair may be accomplished with free margin plication or resuspension of the valve cusps, with or without triangular resection of the redundant segment. Excision of the diseased portion of the involved valve cusp improves symmetry of the valve leaflets, and annular plication of one or both commissures helps to ensure adequate coaptation. Generally, the free margins of the excised leaflets are reaproximated primarily, but in the absence of adequate cusp tissue, a triangular autologous or bovine pericardial patch may be used for cusp restoration.

AV cusp repair with a free margin plication or resuspension technique has demonstrated encouraging results in expert centers, both in patients with tricuspid and bicuspid aortic valves. Freedom from AV reoperation in patients with a tricuspid AV has been reported to be 89% to 92% at 10 years, with a freedom from recurrent AI >2+ of 80% to 86% at the same time point. In patients with bicuspid aortic valves, who generally represent a younger cohort of patients, 10-year survival has been reported at 94% following AV repair, with a freedom from AV reoperation of 81% at the same time point.

**Ross Procedure.** As mentioned previously, the Ross procedure involves replacing the diseased AV with the patient’s native pulmonary valve as an autograft, which is in turn replaced with a homograft in the pulmonic position. The autograft may be implanted in the aortic position directly with resuspension of the valve commissures, or in association with a root replacement, which requires reimplantation of the coronary ostia.

The cylinder root replacement technique is most reproducible and involves transecting the native aorta approximately 5 mm above the sinotubular ridge, with subsequent excision of the aortic valve leaflets and supra-annular tissue. The main pulmonary artery is transected at the bifurcation, and the right ventricular outflow tract is incised, allowing the pulmonary valve and artery to be removed en bloc from the outflow tract. The annulus of the pulmonary autograft is sewn to the native aortic annulus with continuous or interrupted sutures, and the coronary ostia are reimplanted into the pulmonary artery graft. The pulmonary valve and right ventricular outflow tract are subsequently reconstructed using an aortic homograft.

![Figure 21-11. Aortic valve replacement. The stented porcine bioprosthesis as viewed through an aortotomy.](image-url)
The primary benefit of the Ross procedure compared to traditional AV surgery is a low risk of thromboembolism without the need for systemic anticoagulation. Although patients undergoing the Ross procedure are generally younger, perioperative mortality has been reported to be as low as 2.5% in this group, with an overall survival of 90% at 18-year follow-up. However, the long-term durability of the procedure is somewhat questionable. Although Ross reported a freedom from autograft replacement of 75% at 20 years, other groups have reported freedom from autograft reoperation and allograft reintervention of 51% and 82%, respectively, at 18-year follow-up. Progressive aortic insufficiency has been described as a cause of late failure in these patients, as well as calcification of the pulmonary homograft and pulmonary stenosis.

**Transcatheter Aortic Valve Replacement.** Transcatheter aortic valve replacement (TAVR) has proven beneficial for the treatment of AS in patients who are either moderate or high-risk candidates for conventional surgery. TAVR is now indicated for patients with severe AS with a STS score predicted risk of mortality of greater than or equal to 3%. Clinical trials in the low-risk population are currently underway.

There are two types of transcatheter valves that are approved for commercial use: a balloon-expandable valve (Edwards) and a self-expandable valve (CoreValve). A transcatheter valve may be inserted via the femoral artery, the left subclavian artery, the ascending thoracic aorta, or LV apex via a small left anterior thoracotomy. By far the most common route is transfemoral, making up the majority of TAVR in most centers. The principle of valves placed via these routes are to place the aortic prosthesis inside the patient’s native aortic valve. Rigorous preoperative planning is needed to ensure adequate sizing of the valve as well as placement to ensure there is no risk of coronary occlusion or malalignment of the valve.

A series of large, multicenter clinical trials have been performed investigating the role and safety of TAVR in patients with severe aortic stenosis requiring surgical treatment. The PARTNER I trial looked at mortality rate as the primary endpoint in patients with severe aortic stenosis who were not suitable candidates for surgery (high-risk patients). TAVR, as compared with standard surgical treatment (SAVR), significantly reduced the rates of death from any cause (30.7% vs. 50.7%, at 1 year, \( P < 0.001 \)), the composite endpoint of death from any cause or repeat hospitalization (42.5% vs. 71.6%, \( P < 0.001 \)), and cardiac symptoms (25.2% vs. 58.0%, \( P < 0.001 \)), despite the higher incidence of major strokes (5.0% vs. 1.1%, \( P = 0.06 \)) and major vascular events (16.2% vs. 1.1%, \( P < 0.001 \)). In the PARTNER II trial, 2032 intermediate-risk patients with severe aortic stenosis were randomly assigned to undergo either TAVR or SAVR. It was found that in intermediate-risk patients, TAVR was similar to SAVR with respect to the primary end point of death or disabling stroke (\( P = 0.001 \) for noninferiority). TAVR resulted in larger aortic valve areas than did surgery and also resulted in lower rates of acute kidney injury, severe bleeding, and new-onset atrial fibrillation; surgery resulted in fewer major vascular complications and less paravalvular aortic regurgitation.

Similarly, the Surgical or Transcatheter Aortic-Valve Replacement in Intermediate-Risk patients trial (SURTAVI) evaluated the clinical outcomes in intermediate-risk patients with severe, symptomatic aortic stenosis in a randomized trial comparing TAVR (performed with the use of a self-expanding prosthesis) with surgical aortic-valve replacement. TAVR was found to be a noninferior (i.e., similar composite endpoint of death from any cause or disabling stroke at 24 months) alternative to surgery in patients with severe aortic stenosis at intermediate surgical risk, with a different pattern of adverse events associated with each procedure. Surgery was associated with higher rates of acute kidney injury, atrial fibrillation, and transfusion requirements, whereas TAVR had higher rates of residual aortic regurgitation and need for pacemaker implantation. The PARTNER III trial, currently under investigation, is designed to establish the safety and effectiveness of TAVR in patients with severe, calcific aortic stenosis who are at low operative risk for standard surgical aortic valve replacement (SAVR). Ongoing trials will continue to define the roles for SAVR and TAVR in the future.

**TRICUSPID VALVE DISEASE**

**Tricuspid Stenosis and Insufficiency**

**Etiology.** Acquired tricuspid valve (TV) disease can be classified as either organic or functional and affects approximately 0.8% of the general population. Tricuspid stenosis (TS) is almost always a result of rheumatic heart disease or rarely endocarditis. In the case of rheumatic disease, tricuspid stenosis with or without associated insufficiency is invariably associated with mitral valve disease. Other less common causes of obstruction to right atrial emptying include congenital tricuspid atresia, right atrial tumors, and endomyocardial fibrosis.

Tricuspid insufficiency (TR), on the other hand, is most often a functional disease caused by secondary dilation of the tricuspid annulus due to pulmonary hypertension and/or right heart failure. This is most commonly caused by MV disease. Conditions such as right ventricular infarction and pulmonic stenosis can also lead to increased right ventricular pressures and functional TR. The less common causes of organic TR, with or without associated stenosis, include endocarditis, carcinoid syndrome, radiation therapy, trauma such as repeated endomyocardial biopsy, and Marfan syndrome.

**Pathology.** The changes associated with TS closely resemble those associated with MS, including fusion of the commissures. In the case of rheumatic disease, mixed TS and TR may result from fusion and shortening of the chordae tendineae, and fusion of the commissures, causing retraction of the valve leaflets. The right atrium is frequently dilated and thickened in chronic TS, and chronic obstruction to right ventricular filling often produces signs of systemic venous congestion such as hepatomegaly and splenomegaly.

In most cases of TR, dilation and deformation of the tricuspid annulus is the most prominent feature; the valve leaflets oftentimes appear stretched but are otherwise pliable and normal in appearance. When TR is caused by carcinoid syndrome, white fibrous carcinoid plaques are found on the ventricular surfaces of the TV, causing the cusps to adhere to the underlying right ventricular wall and stenting the valve open.

**Pathophysiology.** The basic pathophysiologic abnormality of both TS and severe TR is elevated right atrial pressure, producing systemic congestion and right heart failure. Severe TS is marked by a valve area <1.0 cm², and severe TR is defined as a vena contracta width of >0.7 cm in combination with systolic flow reversal in the hepatic veins. However, in patients with TS, a diastolic pressure gradient of only 5 mmHg, or a TV
orifice <1.5 cm², is frequently enough to cause jugular venous
distention, organomegaly, and peripheral edema. In severe
cases, cardiac output is compromised, especially during exercise
when the fixed obstruction prevents an increase in forward flow.
Patients with severe insufficiency and pulmonary hypertension
experience similar hemodynamic derangements.

Clinical Manifestations. Patients with TS and severe TR
develop symptoms of right heart failure associated with chroni-
cally elevated right atrial pressures. The classic clinical signs
and symptoms of TS and severe TR are jugular venous disten-
tion, hepatomegaly, splenomegaly, ascites, and lower extremity
dema. Uncomfortable fluttering in the neck has been reported
in patients with TV disease, and sensations of throbbing in the
eyeballs and pulsatile varicose veins have been reported to
occur, especially in patients with severe TR.

The low cardiac output syndrome occasionally associated
with TS and severe TR can cause fatigue, weakness, and exercise
intolerance in these patients. In the absence of pulmonary
hypertension, dyspnea is not a prominent feature of tricuspid
disease. The auscultatory findings associated with TS include
a presystolic and middiastolic murmur characterized by a tri-
cuspid opening snap that increases on inspiration. The lower
left parasternal murmur of TR may be holosystolic or less than
holosystolic, depending on the degree of regurgitation, may be
associated with a middiastolic murmur in severe cases, and may
increase on inspiration.

Diagnostic Studies. In patients with TV disease, chest
X-ray frequently demonstrates enlargement of the right atrium
and ventricle. Patients with TS demonstrate an exaggerated a
wave and a diminished rate of y descent in the jugular venous
pulse, while patients with TR have abnormal systolic c and v
waves. TTE examination should be performed in patients with
TV disease in order to characterize the structure and motion
of the TV, the size of the tricuspid anulus, and other cardiac
abnormalities that may affect TV function. In patients with a
pulmonary artery systolic pressure >55 mmHg, TI commonly
occurs in the setting of structurally normal valves; however,
structural derangement of the TV apparatus is frequently present
if TR is documented with a pulmonary artery systolic pressure
<40 mmHg. Doppler TTE allows estimations of the severity of
TR, the right ventricular systolic pressure, and the TV diastolic
gradient.

Indications for Operation. As an isolated lesion, mild or
moderate TV disease does not require surgical correction. How-
ever, patients with severe TV disease should be considered for
surgical intervention, especially in the setting of right ventricu-
lar enlargement and impaired systolic function, as this improves
life expectancy and the development of sequelae such as heart
failure and atrial fibrillation. Depending on the patient’s clini-
cal status and the cause of TV dysfunction, TV repair and TV
replacement be variably recommended for the treatment of TV
dysfunction (Table 21-12). In patients with TR, the valve can
usually be repaired with modern techniques.

Operative Techniques and Results. The TV can be
approached through a median sternotomy, a right thoracotomy,
or port-based techniques. Surgery is performed with the assist-
ance of cardiopulmonary bypass and, though TV surgery is
usually performed on the beating heart, a brief period of car-
dioplegic arrest may be rarely needed to allow for complete
inspection of the interatrial septum and to close any defects that
may be present. TV repair may include a suture or ring annuloplasty as well
as valvuloplasty, and multiple methods have been described.
Historically, bicuspidization of the TV was accomplished by a
figure-of-eight suture plication of the annulus of the posterior
leaflet; however, this technique has been essentially replaced
by suture or ring annuloplasty. Suture annuloplasty is gener-
ally performed by placing pledged suture along the base of
the anterior and posterior leaflets, partially encircling the annulus.
Ring annuloplasty can be accomplished by suturing the TV
annulus to a variety of rigid or semirigid annuloplasty rings,
which generally have an opening at the level of the anterospe-
tal commissure to avoid passing the anchoring sutures near to
the conduction system. Most surgeons favor ring over suture
annuloplasty. In severe annular dilatation, augmentation of the
anterior leaflet with autologous pericardium has been used with
some success. Tricuspid valvuloplasty is infrequently performed
and may include commissurotomy, triangular leaflet resection,

Table 21-12

Data from ACC/AHA guidelines for TV surgery in specific clinical contexts

<table>
<thead>
<tr>
<th>CLINICAL SETTING</th>
<th>CLASS OF RECOMMENDATION</th>
<th>LEVEL OF EVIDENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TVr for severe TI in patients with MV disease requiring MV surgery</td>
<td>I</td>
<td>B</td>
</tr>
<tr>
<td>TVR or annuloplasty for severe symptomatic primary TI</td>
<td>IIa</td>
<td>C</td>
</tr>
<tr>
<td>TVR for severe TI secondary to diseased/abnormal TV leaflets not amenable to annuloplasty or TVr</td>
<td>IIa</td>
<td>C</td>
</tr>
<tr>
<td>Annuloplasty for less than severe TI in patients undergoing MV surgery in the setting of 1) Pulmonary hypertension 2) Tricuspid annular dilatation</td>
<td>IIb</td>
<td>C</td>
</tr>
<tr>
<td>TVR or annuloplasty is not indicated in asymptomatic patients with TI, a normal MV, and a PASP &lt;60 mmHg</td>
<td>III – Harm</td>
<td>C</td>
</tr>
<tr>
<td>TVR or annuloplasty is not indicated in patients with mild primary TI</td>
<td>III – Harm</td>
<td>C</td>
</tr>
</tbody>
</table>

MV = mitral valve; PASP = pulmonary artery systolic pressure; TI = tricuspid insufficiency; TV = tricuspid valve; TVr = tricuspid valve repair; TVR = tricuspid valve replacement.
primary perforation repair, and traditional leaflet repair techniques such as chordal transfer, shortening, and replacement, papillary muscle plication, tricuspid leaflet augmentation, and the edge-to-edge repair technique used in MV prolapse.

For patients with functional TV disease, TV repair is generally preferred to replacement due to favorable results without the associated risks of thrombosis and the need for anticoagulation. In the setting of concomitant mitral valve surgery, TV repair has not been associated with additional perioperative complications, and 5-year freedom from reoperation has been impressive at 98%. However, a subgroup of patients report late failure following TV repair, and this may be worse following suture annuloplasty compared with ring annuloplasty.117

Prosthetic valve replacement may be necessary due to extensive leaflet destruction, as may be seen in patients with endocarditis, or marked annular dilatation not amenable to repair. In some cases, the valve prosthesis may be anchored directly to the leaflet tissue instead of the valve annulus, reducing the risk of injury to the conduction system.106 If this technique is used, it should be confirmed that the residual tissue does not interfere with the movement of the prosthetic leaflets after implantation. Pledgeted sutures should be used and may be placed on the ventricular or atrial side of the annulus.

Outcomes data following TV replacement are difficult to interpret, as most reports are in patients with previous TV surgery and/or signs of severe right heart failure. Operative mortality has been ≥20% in some studies.106 One study of 87 patients undergoing TV replacement between 1994 and 2007 showed an in-hospital mortality of only 1.4%. The choice of prosthetic valve is also somewhat controversial. Though bioprosthetic valves are more durable in the tricuspid than mitral or aortic positions, valve degeneration is an important cause of bioprosthetic valve dysfunction at reoperation. The ability to replace a degenerated tricuspid valve bioprosthesis with a transcatheter valve has led some surgeons to favor biological valves in this position. This is particularly true since the increased risk of valve thrombosis seen with mechanical valves mandates rigorous systemic anticoagulation. Even with these precautions, mechanical tricuspid valves are associated with an increased risk of hemorrhagic and thrombotic complications. The choice of valve is usually decided on a case-by-case basis, and late outcomes have been similar with biological and mechanical valves in this position. In general, TV replacement may be a reasonable choice in select patients, though more data are needed regarding long-term outcomes in the modern era.

### Multivalve Disease

Pathology involving multiple valves is relatively common and may result from diseases such as rheumatic fever, calcific disease, Marfan syndrome, and other connective tissue disorders. However, multivalve disease may also be caused by secondary valvular dysfunction due to a distal valvular lesion, as in the case of myxomatous degeneration of the mitral valve, resulting in pulmonary hypertension, dilation of the tricuspid annulus, and functional TR. If the primary pathology is corrected early in the disease course, these secondary functional changes may resolve without the need for intervention.

In patients with multivalve disease, the clinical manifestations may be dependent on the severity of each individual valve lesion, but this is not always the case.2 In patients with concomitant mitral and tricuspid dysfunction, the prominent symptoms of dyspnea, paroxysmal nocturnal dyspnea, and orthopnea commonly associated with MV dysfunction are sometimes diminished by associated TV dysfunction. Symptoms of multivalve disease are most commonly masked when valvular abnormalities are of approximately equal severity, highlighting the importance of careful examination of each valve both preoperatively and in the operating room.

Surgery for multivalve disease is associated with a higher perioperative mortality than single-valve procedures, and this risk is exacerbated by factors such as pulmonary artery hypertension, age, triple-valve procedures, concomitant coronary artery bypass grafting, previous heart surgery, renal insufficiency, and diabetes.170 Failing to recognize significant concomitant valvular dysfunction at the time of surgery is also associated with higher perioperative mortality. For this reason, patients suspected of having multivalve involvement should undergo full preoperative Doppler TTE or TEE evaluation and heart catheterization.18 In selected patients, procedures correcting multivalve disease demonstrate significant clinical improvement in symptoms and quality of life, as well as acceptable mortality and survival rates.170

### SURGICAL THERAPY FOR THE FAILING HEART

#### Epidemiology of Heart Failure

Heart failure affects approximately 5 million patients in the United States, with >550,000 new cases diagnosed annually.171 The disorder is the primary reason for 12 to 15 million office visits and >1 million hospitalizations each year. Overall 1-year mortality is estimated to be around 25%, but this can increase to as high as 75% for patients with more advanced heart failure (NYHA class IV).172 While heart transplantation remains the gold standard for the treatment of end stage disease, an increasing number of patients deteriorate while on the waiting list, and up to 30% die before transplantation.173 The total direct and indirect costs associated with the treatment of heart failure are estimated to be $32 billion, and this is projected to increase to $70 billion by 2030.174 Advances in the surgical management of heart failure over the two decades have pushed surgery for CHF into the mainstream. As a result, there is an increasing number of patients with late- or end-stage disease who are being considered for surgical therapies.

#### Etiology and Pathophysiology

Heart failure can be classified as acute or chronic, genetic or acquired, left-sided and/or right-sided, and systolic and/or diastolic dysfunction. The underlying causes and treatments for each of these vary considerably. In the Framingham Heart Study, coronary artery disease accounted for 67% of heart failure cases, valvular heart disease accounted for 10%, and 20% of cases were attributable to primary myocardial diseases, of which dilated cardiomyopathy predominated.175 In all cases, heart failure is a progressive disorder that through complex mechanisms of ventricular remodeling, altered hemodynamics, neurohumoral activation, cytokine overexpression, and vascular and endothelial dysfunction either disrupts the ability of the myocardium to generate force or results in a loss of functioning cardiac myocytes, thereby preventing normal myocardial contraction.

#### CABG for Ischemic Cardiomyopathy

Surgical coronary revascularization is among the most commonly performed procedures for CHF. CABG is beneficial as it protects from further myocardial infarction and/or malignant...
ventricular arrhythmias. It is most successful when treating hibernating as opposed to infarcted myocardium. While the majority of evidence supporting CABG for patients with ischemic cardiomyopathy comes from nonrandomized, retrospective studies, the prospective, randomized, multicenter international Surgical Treatment of Ischemic Heart Failure (STICH) trial compared CABG with medical therapy to medical therapy alone. Entry criteria included an EF ≤35% with CAD and anatomy suitable for CABG. No significant difference was seen in overall mortality by study completion, but patients who underwent CABG did have fewer deaths or hospitalizations from cardiovascular causes (58% vs. 68%, P < 0.001).

Myocardial viability testing has been shown by multiple studies to be pivotal in identifying patients that will have improved outcomes following CABG for ischemic cardiomyopathy. A meta-analysis performed by Allman et al demonstrated an 80% reduction in mortality in patients who underwent revascularization with viable myocardium compared to patients who received medical therapy alone (3.2% vs. 16%, P < 0.0001). Most importantly, in this analysis, CABG had no benefit over medical therapy for patients without viable myocardium. A more recent study by Gerber et al prospectively compared CABG and medical therapy to medical therapy alone in 114 patients with CAD and low EF (24% ± 8%) who underwent viability testing using delayed-enhancement cardiac MRI. This study demonstrated worse 3-year survival in medically treated patients with dysfunctional but viable myocardium than in medically treated patients with nonviable myocardium (48% vs. 77%, P = 0.02). This corresponded with a 4.56 times increased hazard of death when medical treatment was selected over full revascularization. In contrast, survival after CABG was not significantly different whether myocardium was viable or not (88% vs. 71%, P = NS). These studies underscore both the importance of viable myocardium as well as the adverse consequences of not offering a patient with viability surgical intervention.

Patients with ischemic cardiomyopathy are a heterogeneous group, and, as with any surgery, appropriate patient selection is central to success. In one retrospective study of 96 patients with ischemic cardiomyopathy (EF ≤25%), age and poor distal vessel quality were predictors of poor outcomes. Mortality in patients with poor vessel quality was 100%, compared with 90% when vessel quality was fair and 10% when it was good. Therefore, poor vessel quality should be considered a contraindication to surgical revascularization even in the presence of angina.

LV size and LV dyssynchrony are also risk factors for adverse short- and intermediate-term outcomes. A LV end-diastolic dimension of >100 mL/m² is associated with a significantly reduced 5-year survival following CABG (85% vs. 53%, P < 0.05), as well as worse 5-year freedom from recurrent CHF (85% vs. 31%). Moreover, LV dyssynchrony has been shown to have a significant impact on mortality in patients undergoing moderate- to high-risk revascularization and may compound risk in patients with nonviable myocardium. In patients with severe preoperative LV dyssynchrony, the 30-day mortality was 27% vs. 3% in patients without significant dyssynchrony (P < 0.001). Similar differences were seen with the presence of postoperative LV dyssynchrony, and outcomes were worse when patients also had fewer segments of viable myocardium.

Secondary Mitral Regurgitation
Secondary mitral regurgitation describes MR that results from damage to the left ventricle as a result of either ischemia or dilated cardiomyopathy rather than from a problem with the valve itself. Ischemic MR (IMR) typically results from systolic restriction of the mitral leaflets due to tethering of the subvalvular apparatus. This occurs mainly from regional wall motion abnormalities in areas of the LV adjacent to papillary muscle attachments. Alterations in the size and shape of the mitral annulus and posterior displacement of the posteroendocardial papillary muscle, which occurs primarily after an infarcted posterior MI, may also contribute. Additionally, functional MR (FMR) is caused by LV dilatation and increased sphericity, which displace the papillary muscles apically and radically, creating lateral forces on the valve that lead to increased retraction of the mitral leaflets by the chordae tendineae. LV dyssynchrony may also contribute to FMR through poor coordination of the contraction of the septum and lateral walls, producing MR that may vary in intensity during the cardiac cycle. Functional MR is usually referred to as a Carpentier class I/IIib lesion due to the presence of both annular dilatation (Carpentier type I) and systolic restriction of the mitral leaflets due to LV dysfunction (Carpentier type IIib). Ultimately, increased regurgitation leads to increased preload, LV wall tension, and LV work load, all of which contribute to progressive dysfunction of the LV and worsening heart failure.

Several observational and population-based studies have demonstrated a significant impact of secondary MR on long-term survival. Following MI, the 5-year survival rate dropped significantly from 61% in patients who did not have MR to 47% and 29% in patients with mild and moderate to severe MR, respectively. Similarly, in a series of 2057 patients with symptomatic heart failure and an LVEF <40%, the 5-year survival rate for patients without MR was 54%, and it decreased to 40% in patients with moderate to severe secondary MR. Moreover, medical therapy and PCI have not reduced the impact of IMR on late mortality.

Although specific recommendations to intervene for secondary MR are controversial and have not been rigorously defined, guidelines are available (Table 21-13). Some surgeons initially advocated performing only revascularization in cases of moderate ischemic MR with the idea that revascularizing viable myocardium would lead to improvements in LV function and effect reverse remodeling, ultimately contributing to a decrease in MR. While several studies, including a recent large, multicenter, randomized, controlled trial, have shown that MR often persists following revascularization alone, the addition of a mitral valve annuloplasty in those studies did not improve long-term functional status or survival in patients with ischemic MR. Nevertheless, other studies have shown that the persistence of MR after CABG is associated with a decreased survival rate and that CABG alone only has modest effects on reducing MR at 1 month follow-up. As a result, some centers continue to repair moderate MR in this patient population. Indications for surgery in ischemic MR patients in the absence of revascularization options are even less well defined.

For patients with functional MR, the goal of mitral valve surgery is to avoid or postpone transplantation in eligible patients. Mitral valve repair has been considered the procedure of choice when surgery is indicated for secondary MR. However, in patients with severe ischemic MR, a recent randomized, multicenter trial showed improved late freedom from moderate or severe recurrent MR with mitral valve replacement compared to repair (2.3% vs. 32.6%, P < 0.001). There was not a significantly higher mortality in the replacement group in this study.
In patients with poor LV function, dilated LV, and severe MR with significant leaflet tethering, we favor MV replacement with preservation of the subvalvular apparatus. Bioprosthetic valve is usually used due to the poor late survival in this group of patients. Currently, for patients undergoing repair recommendations are to use a semirigid or rigid annuloplasty ring to downsize the mitral annulus. There have also been various techniques proposed to correct the papillary muscle displacement, but most reports are single center, retrospective, and small. Mitral valve replacement with preservation of the subvalvular apparatus is indicated when repair is not feasible due to severe tethering of the leaflets or massive LV dilation.

Outcomes from surgery vary among centers and among patients in this heterogeneous group. Operative mortality ranges between 0% and 9% in most modern series. Generally speaking, mortality and recurrence rates are higher and long-term prognosis is worse compared to outcomes for primary MR. Recurrent MR is as high as 15% to 30% in some series, and 5-year mortality is between 44% and 48%. Some reductions in left atrial dimension and LV reverse remodeling may be achieved.

**Left Ventricular Aneurysmmorphy and Surgical Ventricular Restoration**

**Pathophysiology of Ventricular Aneurysms.** A transmural infarction of approximately 5% to 10% of the myocardium may result in formation of an LV aneurysm as necrotic myocardium is replaced by fibrous tissue. This usually occurs 4 to 8 weeks following the infarct. In the last decade, prompt revascularization of the culprit artery by either surgical or interventional techniques generally results in sparing of the subepicardial muscle while the subendocardial muscle remains necrotic. Therefore, it is not uncommon for the LV wall to show both living myocardium during thallium testing and an akinetic zone on echocardiogram or angiogram. It has been demonstrated that once more than 20% of the myocardium is necrosed there is irreversible progression to ventricular dilation and failure. Once heart failure develops after postinfarction remodeling, the 1-year mortality reaches 32% despite current therapies. The classic aneurysm is a 4 to 6 mm thick scar, which bulges outward in paradoxical motion as the LV contracts during systole. More than 80% develop in the anteroseptal and apical portions of the left ventricle as a result of left anterior descending artery occlusion. The rest are inferior in location and the result of circumflex or right coronary occlusion.

This patient population typically suffers from associated ventricular arrhythmias for several reasons. First, electrical dysynchrony results from postinfarction remodeling, and triggers for ventricular arrhythmias typically occur in the scar border zone in patients with ischemic cardiomyopathy. Second, increased ventricular volume causes high wall stress and stretch, and stretch has been shown to be arrhythmogenic. Third, LV aneurysms represent an independent risk factor for SCD after MI. Surgical ventricular restoration (SVR) addresses each of these issues by removing the anatomic substrate during resection of the postinfarct scar and/or aneurysm, accomplishing volume reduction and mechanical resynchronization and relieving ischemia through complete revascularization and reduction in myocardial wall tension and oxygen demand.

**Clinical Presentation and Diagnosis.** Symptoms of LV aneurysms include angina, CHF, ventricular arrhythmias, and rarely, embolic phenomenon. Rupture is extremely uncommon. Patients generally present for coronary artery bypass or during evaluation of CHF or arrhythmias. While transthoracic echocardiography gives pertinent information regarding LV function, size, mitral valve function, and the presence of thrombus, it is generally accepted that cardiac MRI is the best diagnostic modality to accurately identify areas of scar and viable tissue and to best define ventricular geometry.

**Surgical Treatment and Results.** In 1985, Vincent Dor described a surgical technique called the endoventricular circular patch plasty that was intended to improve geometric reconstruction compared with the standard linear repair in LV aneurysm surgery. SVR is a somewhat broader term that arose from surgical repair of ventricular aneurysms and has now come to be applied to a group of surgical procedures designed to correct the effects of postinfarction ventricular remodeling. It is also sometimes referred to as surgical ventricular remodeling or reconstruction, surgical anterior ventricular endocardial reconstruction (SAVER), or the Dor procedure. SVR is specifically intended to reduce the size and sphericity of the LV by...
excluding akinetic and dyskinetic areas, most often by using a circular patch inserted inside the ventricle on contractile myocardium (Fig. 21-12A,B).

Candidates for SVR are typically patients who have had a remote anterior or anteroseptal myocardial infarction, significant ventricular enlargement with a significant area of akinetic or dyskinetic myocardium, a discrete aneurysm, a clinical picture consistent with heart failure (LVEF <40%), retained function of the basilar and lateral portions of the heart, and good right ventricular function. These patients should also be candidates for repair of any other concomitant cardiac disease. Dor currently emphasizes the importance of complete revascularization and repair of any mitral pathology at the time of operative SVR. In patients with spontaneous (13%) or inducible (25%) ventricular tachycardia (VT), it is additionally necessary to perform nonguided endocardial resection and cryoablation encircling the resected area.197

Results with this approach have been good in treating both heart failure and its sequelae, such as VT. In Dor’s series of 1150 patients, the operative mortality varied based on the LVEF, ranging from 1% (patients with EF >40%) to 13% (patients with EF <30%), and the 5-year survival approached 85%.197 Overall, more than 80% of survivors either stabilized or improved, and the quality of life was shown to improve significantly by 6 months after the Dor procedure.203 This is likely due in part to the fact that the Dor procedure restores LV geometry, resulting in a mean ejection fraction increase between 10% and 15%, with significant alleviation of symptoms.197,204-206 These data are reinforced by the international RESTORE group, which examined SVR in a registry of 1198 postinfarction patients between 1998 and 2003.207 They found that 5-year overall freedom from hospital readmission for CHF was 78%. Moreover, 67% of patients had preoperative NYHA class III or IV symptoms, whereas 85% of patients were NYHA functional class I or II postoperatively.

With respect to VT, Dor et al reported on 106 patients with ischemic ventricular arrhythmias that underwent reconstruction for postinfarction LV aneurysm and visually directed endocardectomy plus or minus cryoablation and coronary revascularization.208 At a mean follow-up of 21.3 months, only 10.8% of patients had inducible VT, and no spontaneous VT was documented. Results from similar series have also been excellent,197,206,209 but the efficacy of left ventricular restoration alone has been controversial.209,211 Inferior results seen in some series have been attributed to failure to perform endocardial resection and/or cryoablation at the border of the transitional zone, as well as differences in stimulation protocols and possible inadequate volume reduction of the ventricle.

A large, randomized, multicenter study, the STICH trial, concluded that adding SVR to reduce ventricular volume to CABG does not improve symptoms or exercise tolerance and fails to lower death rate or cardiac rehospitalization compared to CABG alone.199 While this trial has some shortcomings, it has resulted in a marked decrease in referrals for this procedure. The main problem is that the LV volume was reduced by only 19% in the STICH trial, reflecting an inadequate repair as determined by the Surgery Therapy Committee, whose “acceptable STICH procedure” guideline required a 30% reduction at the 4-month postoperative cardiac MRI.212 Previous studies have reported an average reduction of end-systolic volume index (ESVI) of 40% with a range between 30% and 58%, suggesting that the STICH SVR procedure may have involved an inadequately small LV plication or limited intracavitary reconstruction.212 Moreover, this trial enrolled 13% of patients who had never had an MI and changed criteria such that enrollment required documented LV anterior wall dysfunction rather than demonstration of scar. This could have captured patients with hibernating myocardium that would recover following CABG alone. Dor subsequently published the results of 117 patients who would have been eligible for the STICH trial and demonstrated durable improvement in left ventricular function.213 However, this was a single-center, retrospective experience. Caution should be exercised so as not to broadly extrapolate the results of the STICH trial and inappropriately deny appropriate patients effective treatment. This remains an area of controversy.

With the recent advances in percutaneous interventions, the Parachute device has been trialed in human subjects. It is composed of a self-expanding nitinol frame covered with an impermeable fluoropolymer that is deployed into the LV apex walling off akinetic or dyskinetic segments of the LV. The PARACHUTE trial reported 3-year echocardiographic and clinical outcomes of patients with ischemic heart failure who underwent placement of the Parachute device as a feasibility and safety study. They demonstrated that of the 31 patients who received the device, there was improvement or maintenance of NYHA functional class in 85% as well as significant reduction of the LV EDV index.198
Mechanical Circulatory Support

Intra-Aortic Balloon Pump. The intra-aortic balloon pump (IABP) is a commonly used device for mechanical circulatory support and has been in use since 1968. The device is inserted percutaneously through a peripheral artery into the thoracic aorta. The balloon is synchronized so that it inflates during diastole and deflates during systole, resulting in augmentation of diastolic perfusion of the coronary arteries and decreased afterload. Typically, this improves cardiac index and decreases both preload and myocardial oxygen consumption.

Common indications for use of an IABP are cardiogenic shock during or following cardiac catheterization or cardiac surgery. It is also utilized for preoperative stabilization of high-risk patients with either severe coronary artery disease, LV dysfunction, or refractory, unstable angina. Kang et al have reported that risk-adjusted mortality was significantly lower for selected high-risk patients undergoing open heart surgery when a preoperative IABP was used. In 2012, Thiele et al reported their data following a randomized, prospective, multicenter clinical trial looking at the outcome of using intra-aortic balloon pump in the treatment of cardiogenic shock complicating acute myocardial infarction in patients who underwent early revascularization (by means of percutaneous coronary intervention or bypass surgery) (IABP-SHOCK II trial). IABP did not significantly reduce 30-day mortality in this group of patients. Additionally, they reported no significant differences in secondary end points, including the time to hemodynamic stabilization, the length of stay in the intensive care unit, serum lactate levels, and major complications including major bleeding and peripheral ischemia.

Generally, an IABP is used for a few days and the amount of support is weaned as the patient’s condition improves. Mortality associated with device use is typically minimal; however, in one series of 911 patients undergoing CABG who received an IABP, there was a 12% incidence of minor or major vascular complications, including an approximately 3% incidence of limb ischemia requiring thromboembolectomy. This is the most serious complication of IABP placement. To prevent this problem, frequent lower extremity neurovascular checks are necessary while an IABP is in place.

Ventricular Assist Device Indications and Cannulation. Patients in need of ventricular assist devices (VADs) may have preexisting chronic heart failure, refractory ventricular arrhythmias, or acute heart failure following an MI, cardiopulmonary arrest, viral illness, pregnancy, or cardiomyopathy. Device therapy is intended to preserve end-organ perfusion and function and may be categorized as short- or long-term support for the left heart, the right heart, or both. In general, VADs may be used rarely for support while the heart recovers (bridge to recovery, BTR), while the patient waits for a heart transplant (bridge to transplant, BTT) or increasingly more commonly to treat a chronic heart failure patient who is not a transplant candidate (destination therapy, DT). The Interagency Registry for Mechanically Assisted Circulatory Support (INTERMACS) database, a joint effort by the NHLBI, FDA, CMS, academia, and industry to prospectively track patient outcomes, reported that in 2015 and 2016 indications for device implantation were BTR (0.2%), BTT listed (26.4%), BTT likely (15.5%), BTT moderate (7.7%), BTT unlikely (2.5%), DT (49.2%), and rescue therapy (0.5%). The percentage of patients receiving a VAD as destination therapy has markedly increased over the last decade as results and devices improved.

Left ventricular assist devices (LVADs) provide support for the failing heart by unloading blood from the left ventricle and pumping it into the aorta. Cannulas may be inserted into the LV apex or the left atrium for inflow into the pump, and return is through an arterial cannula or graft sewn to either the ascending or descending aorta. For right-sided devices, inflow drainage is most often from a cannula in the right atrium, and blood is returned through a graft sewn to the pulmonary artery or right ventricular outflow tract.

Left Ventricular Assist Devices. The first generation LVADs were pulsatile devices. They provided adequate support for the heart but were limited by their large size and durability. More recently, continuous-flow LVADs based on rotary pump technology have been introduced. These devices are smaller, quieter, and durable enough for long term support. The two most commonly used devices today are the HeartMate II (Thoratec, Pleasanton, CA) and the HeartWare HVAD (HeartWare, Inc., Framingham, MA) (Figs. 21-13A,B and 21-14A,B). These devices differ in that the HeartMate II is implanted subdiaphragmatically, whereas the smaller HeartWare HVAD is implanted within the pericardium. Frequently used short-term support devices include the Abiomed BVS 5000 (Abiomed, Inc., Danvers, MA) and the CentriMag (Thoratec), which are both extracorporeal pumps, as well as the Impella (Abiomed), which may be inserted percutaneously. These devices are commonly used in either post-MI or postcardiotomy heart failure. They have the benefit of faster and easier insertion, making them ideal rescue devices and allowing time for patient transfer to a tertiary referral center, device weaning, transplantation, or transition to a permanent VAD as DT or BTT.

Bridge to Recovery. The ideal clinical situation would be for all LVADs to be temporary with the goal of myocardial recovery. However, as noted previously, this is rare with only 0.2% of devices in the most recent INTERMACS data placed with intent for bridge to recovery. The LVAD Working Group Recovery Study, a prospective multicenter trial investigating myocardial recovery in BTT patients, has shown significant improvements in left ventricular ejection fraction and significant reductions in left ventricular end-diastolic diameter following support with continuous flow pumps, but myocardial recovery resulting in device explantation was still only seen in six patients (9%). Current data suggest that significant reverse remodeling is more likely to occur in the young and those with myocarditis.

Nevertheless, some encouraging results have been reported using a combination of treatment modalities. In a few small studies of patients with LVADs inserted for nonischemic cardiomyopathy, deliberate and aggressive medical therapy, including the β2-agonist clenbuterol, resulted in successful LVAD explantation in 69% to 73% of patients but these results have been difficult to replicate. Moreover, early results from clinical trials using stem cell therapy to treat patients with ischemic cardiomyopathy suggest that stem cells may be another adjuvant treatment with potential to aid in myocardial recovery.

Bridge to Transplant. LVADs are used as a bridge to transplant in patients who are candidates for heart transplantation but are not predicted to survive the waiting list period due to sequelae of cardiac failure, including end-organ dysfunction, rising pulmonary artery pressures, escalating inotrope
Figure 21-13. The HeartMate II LVAD viewed from the (A) outside and (B) inside. The device is an axial flow, rotary pump that produces no pulsatile action. The pump contains a magnet, and the rotor assembly functions by the electromotive force generated by the motor. The result is that blood is propelled from the inflow cannula to systemic circulation at flows up to 10 L/min. (HeartMate, HeartMate II and St. Jude Medical are trademarks of St. Jude Medical, LLC or its related companies. Reproduced with permission of St. Jude Medical, ©2017. All rights reserved.)

Figure 21-14. The HeartWare HVAD system. A. Both the device controller and batteries are held in a wearable carrying case and connected to the ventricular assist device through the driveline. B. The main component is a centrifugal blood pump, called the HVAD, which is implanted within the pericardium. The only moving part in the device, the impeller, is suspended within the pump using magnets and thrust bearings. Similar to the HeartMate II, it can deliver a flow rate of up to 10 L/min. (Reproduced with permission of Medtronic, Inc.)
Most recently, in 2017, Mehra et al reported their multicenter outcome (Mechanical Circulatory Support Therapy with HeartMate 3 trial—MOMENTUM 3) comparing the safety and effectiveness of centrifugal-flow pump (HeartMate 3) to the axial-flow pump (HeartMate II). Implantation of a fully magnetically levitated centrifugal-flow pump (HeartMate 3) was associated with better outcomes at 6 months than was implantation of an axial-flow pump (HeartMate II). The improved outcomes were primarily due to the lower rate of reoperation for pump malfunction. Additionally, no patients in the HeartMate 3 group were suspected or found to have pump thrombosis.

**Destination Therapy.** The Randomized Evaluation of Mechanical Assistance for Treatment of Congestive Heart Failure (REMATCH) trial was conducted to compare the efficacy of LVAD insertion against optimal medical management in patients with NYHA class IV heart failure. While the pulsatile devices used in this trial had high failure rates, poor durability, and high associated mortality, there was still a clear survival benefit in patients treated with LVADs. This led to the FDA approval of the first LVADs for destination therapy in 2002.

Subsequent trials have proven the increased efficacy of second-generation devices for DT. In one such landmark trial, patients with advanced heart failure who were ineligible for transplantation were randomized in a 2:1 ratio to either a HeartMate II or HeartMate XVE. While both groups showed significant improvements in functional capacity and quality of life, actuarial survival at 2 years was superior for HeartMate II patients (58% vs. 24%, $P = 0.008$) and adverse event rates were significantly lower. These data established the benefit of continuous flow LVADs over optimal medical management for end-stage heart failure, and led to FDA approval of the HeartMate II for DT in 2010. In certain populations, 2-year survival with the HeartMate II is now 80%. Several smaller third-generation devices are in various stages of development or clinical trials. Some of these devices eliminate the drive line by using alternative energy sources, thereby removing a significant nidus for device infections. Long-term outcomes with these devices are expected to continue to improve, approaching that of cardiac transplantation and providing a viable solution to organ shortage for many patients.

In 2013, Slaughter et al reported their multicenter clinical trial outcome to evaluate the HeartWare Ventricular Assist Device (HVAD) system for BTT in patients with advanced heart failure in the United States (ADVANCE BTT trial). The use of the HeartWare pump as a BTT continued to come into the light, and the use of the HeartMate XVE was approved for Destination Therapy. However, the HVAD system was limited due to longer survival and lower transplant rates. Current eligibility criteria for mechanical support as destination therapy include (a) NYHA class III or IV heart failure despite guideline-directed medical therapy including cardiac resynchronization therapy if indicated; (b) peak oxygen consumption <12 mL/kg per min or failure to wean from continuous IV inotropes; (c) left ventricular ejection fraction <25%; and (d) presence of a contraindication for heart transplantation (i.e., age >65 years, irreversible pulmonary hypertension, chronic renal failure, insulin-dependent diabetes with end-organ damage, or other clinically significant comorbidities).

Once a patient has an LVAD inserted as DT, close and intensive follow-up by a multidisciplinary heart failure team is required in order to optimize medical therapy, reduce device-related morbidity, and improve survival.

It is also important to keep in mind that while some contraindications to transplantation are irreversible, others can be modified. As such, approximately 10% of patients implanted with an initial strategy of destination therapy become BTT patients, and in some patients, the LVAD itself facilitates this transition. For example, an improvement in mean pulmonary vascular resistance was reported following implantation of the HeartMate II in patients with end-stage heart failure (2.1 vs. 3.6 Woods units, $P < 0.001$). These data are also relevant to patients that receive LVADs as a bridge to decision.

**Right Ventricular Assist Devices and Biventricular Assist Devices**

Most patients who present with advanced heart failure and a failing left ventricle also have some degree of right ventricular dysfunction, but the majority of these patients do well with only an LVAD. However, implantation of an LVAD may cause acute worsening of tricuspid regurgitation and exacerbations of right heart failure through leftward deviation of the intraventricular septum and as a result of the significant volume-loading and transfusion requirement that is often necessary to achieve adequate flows postoperatively. Overall, approximately 20% of HeartMate II BTT patients had persistent right ventricular failure (RVF) requiring either a subsequent RVAD (6%) or intravenous inotropic support for >14 days (14%), and these patients had significantly worse 6-month survival compared to those without RVF (71% vs. 89%, $P < 0.001$). Typically, mechanical right-ventricular support is temporary with intent to wean the device, and isolated right-ventricular assist devices are unusual.

Biventricular support is most commonly indicated for acute cardiogenic shock after an MI or postcardiotomy heart failure. Biventricular support is temporary, although some patients may be successfully bridged to transplant or permanent left-sided assist devices. There is currently no destination therapy device for biventricular failure.

**Total Artificial Heart**

The total artificial heart (TAH, SynCardia Systems, Tucson, AZ) is currently indicated as a bridge to transplant for patients in biventricular failure, particularly for those who are critically ill and too large for extracorporeal BiVAD support. Unlike ventricular assist devices, the TAH replaces the entire heart. The ventricles of the TAH are implanted orthotopically to the atrial cuffs on the ventricular side of the AV groove, and the outflow conduits are attached to the great vessels. This approach has the benefit of obviating the hemodynamic influence of pulmonary hypertension, right heart failure, myocardial or valvular problems, cardiac arrhythmias, and inotropic agents. While this device has failed to reach its potential as a replacement for cardiac transplantation, the TAH has achieved favorable results as a BTT with a >70% survival in selected centers. However, at most centers results with the TAH have been suboptimal, and it is not frequently used. A total of 226 TAH implants between 2013 and 2016 were reported to INTERMACS database. One-year and 2-year survival was reported at 52% and 37%, respectively. Compare to the previously reported survival (2006 through 2012), no significant improvement in survival was noted with time.

**Surgery for Arrhythmias**

The success of catheter-based ablation and implantable cardioverter defibrillators (ICDs) has significantly diminished referrals for the surgical treatment of arrhythmias such as ventricular...
tachycardia, Wolff-Parkinson-White syndrome, atrial flutter, and atrioventricular nodal reentry. On the other hand, the introduction of surgical ablation modalities such as radiofrequency and cryothermal energy, has simplified the surgical treatment of atrial fibrillation and has led to an increase in the number of surgical procedures performed annually for AF, although this has plateaued in recent years.241,242

**Atrial Fibrillation**

**Epidemiology of Atrial Fibrillation.** AF remains the most common arrhythmia in the world with an overall prevalence of 0.4% to 1% that increases to 8% in those older than 80 years old.243 The most serious complication of AF is thromboembolism with resultant stroke,244 but serious morbidity and mortality may also result from hemodynamic compromise due to loss of atrial contraction and exacerbations of CHF from atrioventricular asynchrony and tachycardia-induced cardiomyopathy.

**Medical Management.** Most patients are treated medically, but the shortcomings of pharmacological management have left an important role for interventional therapies. Antiarrhythmic medications have been limited by modest efficacy and significant proarhythmic and systemic toxicities.245 Conversely, rate control strategies leave the patient in AF, do not address the impaired hemodynamics or symptoms associated with this arrhythmia, and may render subsequent attempts at rhythm control therapies less effective for younger patients who may suffer irreversible cardiac remodeling due to the prolonged period of time in AF. Additionally, AF is associated with a fivefold greater risk of ischemic stroke or systemic embolism compared with normal sinus rhythm. Annual risk of major bleeding in those on anticoagulation is estimated at 1.2%.246

Restoration of normal sinus rhythm has several potential benefits over other strategies.247-249 These include improvement in atrial systolic function, which improves cardiac output and often symptoms of CHF; lowered risks of stroke; potential freedom from anticoagulation; and likely reversal of atrial structural and/or electrical remodeling.

**Indications for Surgical Management.** Consensus guidelines published by the Heart Rhythm Society state that surgical ablation for atrial fibrillation is indicated for (a) all symptomatic AF patients undergoing other cardiac surgery; (b) selected asymptomatic AF patients undergoing cardiac surgery in which the ablation can be performed with minimal additional risk; and (c) symptomatic patients with lone AF who have failed medical therapy and prefer a surgical approach, have failed one or more attempts at catheter ablation, or are poor candidates for catheter ablation.243 At our institution, relative indications for surgical ablation in patients with permanent AF that were not included in the consensus statement are (a) a contraindication to long term anticoagulation for patients at high risk for stroke (CHADS2 score ≥2) and (b) a history of stroke while on therapeutic anticoagulation. Since the consensus statement was released, a multicenter, randomized, controlled trial of surgical ablation in patients undergoing mitral valve surgery showed a significant improvement in freedom from atrial fibrillation in patients receiving surgical ablation (63% vs. 29%, *P* <0.001).250 Controversially, this trial did not show a difference between left atrial and bilateral lesion; this may have been due to technical issues with the operations.251 The STS has recently released guidelines for surgical ablation that give a Class I Level A recommendation for concomitant surgical ablation at the time of mitral valve surgery and a Class I Level B-NR recommendation for concomitant surgical ablation at the time of AVR, CABG, or AVR-CABG.252

**The Cox-Maze IV Procedure.** The first successful operation for atrial fibrillation, the Cox-Maze procedure, was introduced clinically in 1987 by James Cox. The procedure involved the completion of a maze-like pattern of surgical incisions across both the right and left atrial that were designed to interrupt the multiple macroreentrant circuits thought to be responsible for AF, while still allowing propagation of the sinus impulse, restoring atrioventricular synchrony, and preserving atrial transport function. While effective at eliminating AF and reducing the risk of thromboembolism, it was not widely performed because it was technically difficult and significantly prolonged time on cardiopulmonary bypass. In 2002, the Cox-Maze IV, was introduced. The Cox-Maze IV uses a combination of bipolar radiofrequency (RF) ablation and cryoablation to effectively replace the majority of incisions that comprise the Cox-Maze III while significantly shortening cross-clamp time and reducing operative complexity.

The Cox-Maze IV is performed on cardiopulmonary bypass through either a median sternotomy, often in combination with other cardiac surgery or a right minithoracotomy.253,254 In most cases, the right atrial lesion set performed on the beating heart, whereas the left atrial lesions are performed during cardioplegic arrest (Fig. 21-15).

**Results from the Cox-Maze IV procedure have been excellent.** The Washington University group reported a series of 576 consecutive patients in 2015, demonstrating freedom from atrial tachyarrhythmias in 92% of patients at 1 year and 73% of patients at 5 years postoperatively.255 Additionally, freedom from atrial tachyarrhythmias and antiarrhythmic drugs was 81% at 1 year and 61% at 5 years. A recent propensity-matched analysis showed that the addition of the Cox-Maze IV procedure to a routine cardiac surgery did not significantly increase postoperative morbidity or mortality and was associated with improved late survival compared with patients with untreated AF and a similar survival to patients without a history of AF.256 A propensity analysis has shown that results are similar between the traditional “cut-and-sew” maze (Cox-Maze III) and the Cox-Maze IV.257 This procedure is often successful in patients who are poor candidates for catheter-based ablation, such as those with large left atria and patients with long-standing persistent AF.

The combination of surgical management of the left atrial appendage (LAA) and restoration of normal sinus rhythm after the Cox-Maze procedure significantly reduces stroke risk. It is our practice to stop warfarin at 3 months postoperatively in patients who are in normal sinus rhythm and without another indication for anticoagulation, regardless of CHA2DS2-VASc score. With this approach, the stroke rate following the Cox-Maze procedure off anticoagulation has been remarkably low (annual risk = 0.2%).258 In contrast, in one report the annual rate of intracranial hemorrhage in anticoagulated patients with AF was 0.9% per year, and the overall rate of major bleeding complications was 2.3% per year.259

**Left Atrial Lesion Sets.** Some surgeons perform more limited ablation procedures, such as isolated pulmonary vein isolation or lesion sets that are limited to the left side of the heart. This is done in order to further reduce the complexity of the procedure and takes advantage of the fact that in most patients AF
Pulmonary Vein Isolation. Pulmonary vein isolation (PVI) is an attractive therapeutic option because it can be performed off of cardiopulmonary bypass (CPB) through small or thoracoscopic incisions. The results of PVI have been variable and highly dependent on patient selection since outcomes are consistently worse in patients with longstanding persistent AF. In a study from Edgerton et al, only 56% of patients were free from AF at 6 months (35% off antiarrhythmic drugs), and with concomitant procedures, the success rate of PVI has been even lower.267 Several devices are available to close the LAA at the time of PVI. These include staplers and epicardial clips that can be placed without the need for CPB.263

While surgical PVI has had poorer results than a Cox-Maze procedure, it has had superior results to catheter-based PVI. The Atrial Fibrillation Catheter Ablation Versus Surgical Ablation Treatment (FAST) Trial, which was a two-center, randomized clinical trial, compared catheter-based ablation to thoracoscopic PVI in patients with antiarrhythmic drug-refractory AF and either left atrial dilatation and hypertension or failed prior catheter-ablation.264 This study demonstrated that the 12-month freedom from AF and antiarrhythmic drugs was 37% for the catheter ablation group and 66% for the PVI group (P = 0.002).264

**Surgery for Pericardial Disease**

**Acute Pericarditis**

Pericarditis is characterized by infiltration of the cellular and fibrous pericardium by inflammatory cells. The exact incidence and prevalence of pericarditis is unknown, but it is estimated that pericarditis is found in approximately 1% of autopsies and accounts for up to 5% of presentations of nonischemic chest pain.265,266 The etiologies of acute pericarditis are diverse and may result from primary pericardial disorders or occur secondary to a systemic illness.267 In developed countries, 80% to 90% of cases are now considered idiopathic or related to a viral pathogen, but nonviral infection, autoimmune diseases, myocardial infarction, radiation, malignancy, endocrinopathy, myocarditis, aortic dissection, uremia, trauma, pharmacologic side effects, and previous cardiothoracic surgery must be included in the differential diagnosis. The relative incidences of peri-infarction pericarditis, which was once common, and postcardiac injury syndrome have been dramatically reduced with the advent of thrombolytics and coronary angioplasty.267

**Clinical Presentation and Diagnosis**

Diagnosis of acute pericarditis typically requires the identification of at least two of four cardinal features (Table 21-14). The presentation may be confused with several more common cardiopulmonary conditions, particularly myocardial infarction, making a careful history and physical critical. Patients with pericarditis classically complain of sudden onset, retrosternal pain that may be pleuritic in nature. The pain may also be positional, with alleviation of pain when the patient is upright and leaning forward. Pain from pericarditis is typically sharp or stabbing, as opposed to the dull pain or pressure that is common with angina, and it typically does not crescendo. While both conditions cause pain that often radiates to the neck, arms, and shoulders, pericarditis pain may
uniquely radiate to the trapezius ridge due to innervation from the phrenic nerve.\textsuperscript{268-269}

The presence of a pericardial friction rub is pathognomonic for pericarditis, but it tends to vary in intensity over time and may be absent in 15\% to 65\% of patients.\textsuperscript{265,268} As such, the sensitivity of this physical finding is dependent on the frequency and quality of auscultation. A pericardial friction rub is best heard at the left lower sternal border and is typically described as a high-pitched scratchy or squeaky sound with a triphasic cadence corresponding to the movement of the heart during atrial systole, ventricular systole, and early ventricular diastole. However, it may be monophonic or biphasic in up to 50\% of patients.

Electrocardiogram changes typically progress through four stages representing global subepicardial myocarditis and subsequent recovery. Pericarditis patients may have concave ST deflections with diffuse changes, spanning the leads of multiple coronary artery distributions, but ST segment abnormalities are absent in 20\% to 40\% of patients.\textsuperscript{270,271} Acute pericarditis should not result in the development of infarct patterns, such as Q waves or loss of R waves, and T-wave inversions from pericarditis tend to result later in the disease process after the ST segment has returned to baseline.

Echocardiography is routinely performed in the evaluation of acute pericarditis. Its role is primarily to assess for a pericardial effusion. However, in a patient who can be demonstrated to have previously had normal cardiac function, it may be used to exclude segmental wall motion abnormalities that may suggest ischemia.

The remaining workup should attempt to determine the underlying cause of the pericarditis and should be directed by the history and physical. Most inflammatory markers and laboratory tests are nonspecific, but C-reactive protein may be useful in predicting recurrence risks and in guiding the duration of anti-inflammatory medications.\textsuperscript{272} Rarely, other imaging modalities, such as CT scanning, pericardial biopsies, or pericardiocentesis may aid in diagnosis.

Treatment. The preferred treatment depends on the underlying cause of the pericarditis. The disease usually follows a self-limited and benign course and can be successfully treated with a short course of nonsteroidal anti-inflammatory agents (NSAIDs). The addition of colchicine may be beneficial.\textsuperscript{273} Some patients may require judicious use of steroids or IV antibiotics. In cases of purulent pyogenic pericarditis, surgical exploration and drainage are occasionally necessary. Rarely, accumulation of fluid in the pericardium may lead to tamponade, requiring prompt evacuation of the pericardial space. While pericardiocentesis will typically suffice, surgical drainage may be required for thick, viscous, or clotted fluid or in patients with significant scarring from previous operations. More commonly, surgical intervention is required to manage recurrent disease.

Relapsing Pericarditis

As many as one-third of patients with acute pericarditis will develop at least one episode of relapse.\textsuperscript{267} While many of these patients can be treated medically during their initial relapse and do not experience further episodes, a subset of patients experience chronic relapsing pericarditis that can significantly impact their quality of life. Recurrence may develop either from the original etiology or from an autoimmune process that occurs as a consequence of damage from the initial episode. Relapsing pericarditis normally responds to a longer course of NSAIDS ± colchicine. While steroids may induce rapid symptomatic response, their use should be limited to patients who have multiple relapses and are unresponsive to first-line agents, as several studies have suggested that steroid administration may favor relapse.\textsuperscript{273,274}

Pericardiectomy may be considered a last resort treatment in patients with relapsing pericarditis who are severely symptomatic despite optimal medical management, are unable to tolerate steroids, or have recurrence with tamponade. Evidence for this approach is lacking, as few studies have described pericardiectomy in this population.\textsuperscript{275-277} The largest study and the only one to compare surgical treatment with medical management for patients with persistent relapsing pericarditis was a report of 184 patients from the Mayo Clinic.\textsuperscript{276} About 58 patients were identified as having undergone a pericardiectomy after failed medical treatment, whereas the remainder were treated with medical management only. Compared to medical treatment only, pericardiectomy resulted in significantly fewer relapses (8.6\% vs. 28.6\%, $P = 0.009$) at long term follow-up, as well as a nonsignificant trend towards less medication and corticosteroid usage. Of note, 80\% of patients in the pericardiectomy group who had relapses reported significant improvements in their symptoms and had fewer relapses than before pericardiectomy. No perioperative deaths were observed, and only two patients (3\%) had major complications. Hence, at experienced centers pericardiectomy may be a safe and viable option in select patients with relapsing pericarditis.

Chronic Constrictive Pericarditis

**Etiology, Pathology, and Pathophysiology.** Constrictive pericarditis can occur after any pericardial disease process but remains a rare outcome of recurrent pericarditis. It results when chronic pericardial scarring and fibrosis cause adhesion of the visceral and parietal layers and resultant obliteration of the pericardial space. While the pericardium is often grossly thickened with either focal or diffuse calcification in chronic disease, constriction may occur with normal pericardial thickness in approximately 20\% of cases.\textsuperscript{267,278} In developed nations, idiopathic causes and cardiac surgery (accounting for almost 40\% of cases in some series) are the predominant underlying etiologies, followed by mediastinal radiation, pyogenic infections (i.e., Staphylococcus), and other miscellaneous causes. Tuberculosis is an additional common cause in immunosuppressed patients and in developing or underdeveloped countries.

Clinically, pericardial constriction limits diastolic filling of the ventricles and mimics right heart failure since the right-sided chambers are more affected by a rise in filling pressures. Subsequent increases in central venous pressure result in the progressive development of hepatomegaly, ascites, peripheral edema, abdominal pain, dyspnea on exertion, anorexia, and nausea (in part due to hepatic and bowel congestion). In many patients, these symptoms develop insidiously over a course
of years. Since many of these symptoms are similar to those seen in patients with restrictive cardiomyopathy, the distinction between the two entities is difficult, but it remains critical because the treatment is completely different for restriction. The primary difference is that restrictive cardiomyopathy is defined by a nondilated ventricle with a rigid myocardium that causes a significant decrease in myocardial compliance, which is not seen in constrictive pericarditis.

**Clinical and Diagnostic Findings.** Classic physical exam findings include jugular venous distention with Kussmaul’s sign, diminished cardiac apical impulses, peripheral edema, ascites, pulsatile liver, a pericardial knock, and, in advanced disease, signs of liver dysfunction, such as jaundice or cachexia. The “pericardial knock” is an early diastolic sound that reflects a sudden impediment to ventricular filling, similar to an S3 but of higher pitch.

Several findings are characteristic of noninvasive and invasive testing. CVP is often elevated 15 to 20 mmHg or higher. ECG commonly demonstrates nonspecific low voltage QRS complexes and isolated repolarization abnormalities. Chest X-ray may demonstrate calcification of the pericardium, which is highly suggestive of constrictive pericarditis in patients with heart failure, but this is present in only 25% of cases. Cardiac CT or MRI (cMRI) typically demonstrate increased pericardial thickness (>4 mm) and calcification, dilation of the inferior vena cava, deformed ventricular contours, and flattening or leftward shift of the ventricular septum. Pericardial adhesions may also be seen on tagged cine MRI studies.

As discussed, it is most important to distinguish pericardial constriction from restrictive cardiomyopathy, which is best done with either echocardiography or right heart catheterization. Findings favoring constriction on echocardiography include respiratory variation of ventricular septal motion and mitral inflow velocity, preserved or increased mitral annulus early diastolic filling velocity, and increased hepatic vein flow reversal with expiration. Cardiac catheterization will show increased atrial pressures, equalization of end-diastolic pressure and early ventricular diastolic filling with a subsequent plateau, called the “square-root sign.” Additional findings upon catheterization that would favor constriction include respiratory variation in ventricular filling and increased ventricular interdependence, manifest as a discordant change in the total area of the LV and RV systolic pressure curve with respiration.

**Surgical Treatment.** Transient constrictive pericarditis may occur weeks to months after an initial injury and follows a self-limiting course of weeks to a few months. These patients are best treated with medical therapy alone. They often lack calcification of their pericardium, and the degree of late gadolinium enhancement of the pericardium on cardiac MRI has shown promise in predicting which patients may have resolution of the process. Still, there is no ideal way to distinguish these patients from those who will develop chronic constrictive pericarditis, which is permanent. Therefore, if a newly diagnosed patient is hemodynamically stable, it is recommended that conservative management is attempted for 2 to 3 months prior to performing a pericardectomy. Surgical therapy should not be delayed indefinitely, however, as results are improved when the operation is performed earlier in the course of the disease. A series of 938 patients undergoing pericardiectomy reported by the Mayo Clinic, 355 of whom underwent pericardiectomy for constrictive pericarditis, showed significantly lower survival in patients with constrictive pericarditis compared with patients with effusive/relapsing pericarditis. Patients with left ventricular systolic dysfunction or right ventricular dilatation are at increased risk of early mortality. Additional factors that predict adverse long-term outcomes include older age and prior ionizing radiation, as well as cardiopulmonary and renal dysfunction. Surgery should therefore be approached cautiously in patients with advanced, “end-stage” constrictive pericarditis, mixed constrictive-restrictive disease (often from radiation), and significant myocardial or renal dysfunction, as those patients are at increased risk from surgery and may not experience improvement of symptoms.

In order to minimize recurrence following pericardiectomy, complete pericardial resection is desirable. This is typically performed through either a median sternotomy or left anterolateral thoracotomy while on cardiopulmonary bypass. Radical pericardiectomy involves wide resection of the constricting pericardium from the anterior surface of the heart between the phrenic nerves and the diaphragmatic surface. Decortication of the right atrium and vena cavae is not universally performed, but doing so improves the risk of persistent disease or relapse.

The extent of myocardial involvement may also affect long-term outcomes, and, thus, the depth of decortication is an important consideration. Even when an adequate pericardiectomy is performed, epicardial sclerosis can cause persistent hemodynamic instability or a delayed response to surgery. Sclerotic epicardium is often thin and nearly transparent, but in cases of severe chronic constrictive pericarditis it can be difficult to remove it without injury to the heart.

**Surgical Results.** While most patients experience significant improvement in their symptoms following pericardiectomy, symptomatic relief may take several months. Since there is a significant perioperative morbidity and mortality, pericardiectomy is best performed by experienced surgeons at high-volume centers. Between 1970 and 1985, the operative mortality was reported to be 12%, but a lower mortality of approximately 4% to 8% was noted between 1977 and 2006 at several experienced centers.

Long-term survival is in part determined by etiology of the disease. In a report from the Cleveland Clinic, 7-year survival rates following pericardiectomy for idiopathic, postsurgical, and radiation-induced constrictive pericarditis were 88%, 66%, and 27%, respectively. Results are worst for radiation-induced disease because ionizing radiation is often associated with myocardial injury as well as pericardial disease.

Despite the risks, many patients experience significant benefits from surgical treatment. In one large series, 83% of patients were reported to be free of symptoms at last follow-up. This is in agreement with other studies that have shown a significant improvement in NYHA functional status from class III/IV preoperatively to class I/II following pericardiectomy in >95% of patients.

**CARDIAC NEOPLASMS**

**Overview and General Clinical Features**

Cardiac neoplasms are rare, with an incidence ranging from 0.001% to 0.3% in autopsy studies and a 0.15% incidence in major echocardiographic series. In one large autopsy series, 99.2% of cardiac tumors were metastatic in origin; however,
these patients almost never present for surgical management as they usually have fatal diffuse metastatic disease. As a result, a majority of surgical series describe management of primary cardiac neoplasms. Benign cardiac tumors are most common and account for 75% of primary neoplasms. Approximately 50% of benign cardiac tumors are myxomas, with the remainder being papillary fibroelastomas, lipomas, rhabdomyomas, fibromas, hemangiomas, teratomas, lymphangiomas, and others, in order of decreasing frequency. Most malignant primary cardiac tumors are sarcomas (angiosarcoma, rhabdomyosarcoma, fibrosarcoma, leiomyosarcoma, and liposarcoma), with a small incidence of malignant lymphomas.

**Clinical Presentation.** The clinical presentation of cardiac neoplasms varies greatly depending on the location of the tumor, as well as its size, rate of growth, invasiveness, and friability. While as many as 10% of patients are asymptomatic, most manifest some combination of symptoms from the classic triad resulting from blood flow obstruction, tumor embolization, and constitutional symptoms. Systemic manifestations of disease include fever, myalgias, chills, night sweats, weight loss, and fatigue and occur in up to one-third of patients.

Obstruction of cardiac blood flow accounts for the majority of presenting symptoms. When the tumor is located in the left atrium, symptoms tend to mimic mitral valve disease with dyspnea and pulmonary edema; although more severe presentations with syncopal episodes, hypotension, and sudden cardiac death have been reported from temporary valve orifice occlusion. When the tumor is located in the right atrium, symptoms may mimic right heart failure and include hepatomegaly, ascites, and peripheral edema. Outflow tract obstruction is rare but may be caused by large ventricular tumors.

Tumor lysis and embolization may also lead to neurologic presentations such as stroke, retinal artery occlusion, or cerebral aneurysms, particularly in the case of pedunculated tumors and those with frond-like projections. Embolic tumor cells are able to lodge and penetrate distant vessel walls via subintimal growth, which leads to weakening of the arterial wall and subsequent aneurysm formation. This has been documented as late as 5 years after successful primary myxoma resection. Alternatively, embolic implants may metastasize and create space occupying lesions. While rare, myxomatous tumor emboli have also been identified in the coronary arteries, common iliac and femoral arteries, kidney, spleen, pancreas, and liver.

Certain clinical features may be helpful in distinguishing benign from malignant primary cardiac tumors. Malignant tumors, primarily sarcomas, do not demonstrate a gender preference and tend to present after the fourth decade of life. They are often multifocal within the right atrium, and intramyocardial invasion can lead to refractory congestive heart failure, arrhythmias, hemopericardium, and ischemia. Conversely, benign tumors, primarily myxomas, are typically unifocal in the left atrium, have a 3:1 female preference, and occur in younger patients. Arrhythmias and pericardial effusions are very rare in this population.

**Diagnosis and Characterization of Cardiac Masses.** Trans-thoracic echocardiography is the mainstay imaging technique for the detection of cardiac tumors. However, echocardiography is limited by dependence on an acoustic window, suboptimal visualization of extracardiac extension, and poor soft-tissue visualization. TEE is generally only beneficial for small localized tumors due to its limited field of view. cMRI is therefore the current standard for delineating the anatomical extent of the tumor and assessing the paracardiac space and great vessels. Advantages of cMRI over CT scans include better soft-tissue evaluation without the need for iodinated contrast and no exposure to ionizing radiation.

It is important in the initial workup to distinguish a cardiac tumor from an intracardiac thrombus, which may be common in the atria of patients with AF and can mimic echocardiographic features of atrial myxomas. This determination is critical, as an atrial thrombus may be expected to resolve with anticoagulation, whereas a tumor requires surgical intervention. Moreover, anticoagulation can potentially increase the risk of peripheral embolization in patients with cardiac tumors. Delayed enhancement cMRI is the best modality to separate these two entities. cMRI may show vascularization, areas of necrosis, hemorrhage, or calcification in cardiac tumors that are not present in thrombi.

**Myxoma**

**Pathology and Genetics.** Cardiac myxomas are the most common primary cardiac tumor and are characterized by several distinguishing features. About 75% of the time, they arise from the interatrial septum near the fossa ovalis in the left atrium. Most others will develop in the right atrium, but, less commonly, they can arise from valvular surfaces and the walls of other cardiac chambers. Macroscopically, these tumors are pedunculated with a gelatinous consistency, and the surface may be smooth (65%), villous, or friable. Size varies greatly with these tumors and ranges from 1 to 15 cm in diameter. Internally, myxomas are heterogeneous and often contain hemorrhage, cysts, necrosis, or calcification. Histologically, these tumors contain cells that arise from a multipotent mesenchyme and are contained within a mucopolysaccharide stroma.

While the majority of myxomas occur spontaneously with the highest incidence in women aged 40 to 60 years old, approximately 7% of cases are familial as part of Carney complex. Carney complex is an autosomal dominant disorder characterized by two or more of the following conditions: atrial and extracardiac myxomas, schwannomas, cutaneous lentiginosis, spotty pigmentation, myxoid fibroadenomas of the breast, endocrine overactivity (pituitary adenomas or primary adrenal hyperplasia with Cushing’s syndrome), and testicular tumors. Compared to sporadic myxomas, those that occur as part of Carney complex are more commonly found in the right atrium (37% vs. 18%) or one of the ventricles (25% vs. 0%), more often multicentric (33% vs. 6%) and more likely to recur (20% vs. 3%). They also present earlier at an average age of 24 years old (range 4–48 years).

**Pathophysiology.** Larger tumors are more likely to be associated with cardiovascular symptoms from obstruction, and embolic symptoms tend to occur from organized thrombi present on friable or villous tumors (Fig. 21-16). The relative frequencies of symptoms was illustrated by a series of 112 patients who reported cardiovascular symptoms (67%), most commonly resembling mitral valve obstruction; systemic embolization (29%); neurologic deficits (20%); and constitutional symptoms (34%). Similar incidences of symptoms have been reported in other large studies.

**Treatment.** Cardiac myxomas should be promptly excised after diagnosis due to the significant risk of embolization and cardiovascular complications, including sudden death. Resection may be performed through either a median sternotomy or...
a minimally invasive right thoracotomy while on cardiopulmonary bypass. Care is taken not to manipulate the tumor before cross clamping of the aorta in order to avoid embolization. Left atrial tumors may be approached through a standard left atriotomy. Exposure of large tumors attached to the interatrial septum may be facilitated by an additional parallel incision in the right atrium, but this is rarely necessary. An ideal resection encompasses both the tumor and a portion of the cardiac wall or interatrial septum to which it is attached. In order to prevent recurrence, a full thickness excision of the attachment site is preferred, but partial thickness excisions and cryoablation of the base have been performed with good late results. The defect created in the atrial septum can either be repaired primarily or with a small patch. Finally, patients with valvular involvement may require additional valvular reconstruction or replacement, and rare cases of cardiac autotransplantation (with atrial reconstruction) or transplantation have been reported as strategies for complex cases of recurrent atrial myxoma.

Short- and long-term results following excision are excellent for benign cardiac myxomas. Operative mortality is low, and the probability of disease-free survival at 20 years has been reported to be as high as 92% for benign, sporadic myxomas. Risk of recurrence is significantly higher for familial cases. Other risk factors for recurrence include younger age, smaller tumor mass, and ventricular tumor location.

Other Benign Cardiac Tumors
There are several benign cardiac tumors apart from myxomas that are infrequent but have distinct pathophysiologic features. Papillary fibroelastomas are the second most common primary cardiac tumor, representing approximately 8% of all cases. These tumors typically occur in more elderly patients; are small (<1 cm in diameter) sessile, pedunculated masses that arise from the mitral or aortic valves; and frequently result in embolization. Fibroelastomas can almost always be resected with preservation of the native valve leaflets, and cryoablation of the valve leaflet after resection can help prevent recurrence. Lipomas are encapsulated tumors that usually arise from the epicardium and remain asymptomatic in most patients. Hemangiomas, which may arise from any cardiac structure, including the pericardium, account for 2% of benign cardiac tumors, and atrioventricular node tumors, which often lead to sudden cardiac death from heart block and ventricular fibrillation, are exceedingly rare.

In children, rhabdomyomas are the most common primary cardiac tumor, whereas fibromas are the most commonly resected cardiac tumor. Rhabdomyomas are myocardial hamartomas that are often multicentric in the ventricles. About 50% of cases are associated with tuberous sclerosis, and while resection is occasionally necessary, most disappear spontaneously. Fibromas are congenital lesions that one-third of the time are found in children younger than 1-year old. These tumors, conversely, are ordinarily solitary lesions found in the inner interventricular septum, and they may present with heart failure, cyanosis, arrhythmias, syncopal episodes, chest pain, or sudden cardiac death.

Malignant Cardiac Tumors
Primary cardiac malignancies are very rare, but when they occur they tend to have a right-sided predominance and frequently demonstrate extracardiac extension and involvement. Malignant cardiac tumors include intimal sarcoma, angiosarcoma, osteosarcoma, leiomysarcoma, rhabdomyosarcoma, liposarcoma, and primary cardiac lymphomas. Intimal sarcoma is the most common subtype. Angiosarcomas are aggressive, rapidly invading adjacent structures, and 47% to 89% of patients present with lung, liver, or brain metastases by the time of diagnosis. Leiomyosarcomas are sessile masses with a mucous appearance that are typically found in the posterior wall of the left atrium. Rhabdomyosarcomas are bulky (>10 cm in diameter) tumors that usually occur in children and do not have a predilection for any particular chamber. They frequently invade nearby cardiac structures and are multicentric in 60% of cases. Finally, while not as frequent as secondary cardiac lymphomas, primary cardiac lymphomas are increasing in frequency due to lymphoproliferative disorders caused by Epstein-Barr virus in immunosuppressed patients. The absence of necrotic foci in lymphomas can be used to differentiate these tumors from cardiac sarcomas.

Metastatic Cardiac Tumors
Cardiac metastases have been found in approximately 10% of autopsies performed for malignant disease. Secondary cardiac tumors, unlike primary tumors, are therefore relatively common. They may arise from direct extension of mediastinal tumors, hematological spread, intracavitary extension from the inferior vena cava or lymphatic extension, although the latter is the most common mechanism.

While they can occur with most any primary tumor, they are generally observed late in the course of disease. Malignant melanomas have a high potential for cardiac involvement, but...
other soft tissue tumors such as lung cancer, breast cancer, sarcomas, renal carcinoma, esophageal cancer, hepatocellular carcinoma, and thyroid cancer may all progress to cardiac involvement. Cardiac metastases may also develop from leukemia and lymphoma in 25% to 40% of cases.304

Metastatic cardiac tumors are typically found in random locations, excluding the valvular tissue where lymphatics are absent, and they may be multifocal or diffusely extend along the epicardial surface. Signs of malignant cardiac involvement in cancer patients include pericardial effusion or tamponade, tachyarrhythmias, and heart failure symptoms. Workup is similar to other cardiac tumors. Treatment is generally with combined chemotherapy and radiation and is rarely effective.

REFERENCES

Entries highlighted in bright blue are key references.

1. Task Force for the Diagnosis and Management of Syncope; European Society of Cardiology (ESC); European Heart Rhythm Association (EHRA); et al. Guidelines for the diagnosis and management of syncope (version 2009). Eur Heart J. 2009;30(21):2631-2671.


98. Nishimura RA, et al. 2014 AHA/ACC guideline for the management of patients with valvular heart disease: a report


ACQUIRED HEART DISEASE

CHAPTER 21


283. Chowdhury UK, Subramaniam GK, Kumar AS, et al. Pericardiecomy for constrictive pericarditis: a clinical, echo-

ANATOMY OF THE AORTA

The aorta consists of two major segments—the proximal aorta and the distal aorta—whose anatomic characteristics affect both the clinical manifestations of disease in these segments and the selection of treatment strategies for such disease (Fig. 22-1). The proximal aortic segment includes the ascending aorta and the transverse aortic arch. The ascending aorta begins at the aortic valve and ends at the origin of the innominate artery. The first portion of the ascending aorta is the aortic root, which includes the aortic valve annulus and the three sinuses of Valsalva; the coronary arteries originate from two of these sinuses. The aortic root joins the tubular portion of the ascending aorta at the sinotubular ridge. The transverse aortic arch is the area from which the brachiocephalic branches arise. The distal aortic segment includes the descending thoracic aorta and the abdominal aorta. The descending thoracic aorta begins distal to the origin of the left subclavian artery and extends to the diaphragmatic hiatus, where it joins the abdominal aorta. The descending thoracic aorta gives rise to multiple bronchial and esophageal branches, as well as to the segmental intercostal arteries, which provide circulation to the spinal cord.

The volume of blood that flows through the thoracic aorta at high pressure is far greater than that found in any other vascular structure. For this reason, any condition that disrupts the integrity of the thoracic aorta, such as aortic dissection, aneurysm rupture, or traumatic injury, can have catastrophic consequences.

Historically, open surgical repair of such conditions has been an intimidating undertaking associated with significant morbidity and mortality. Strategies for protecting the brain and spinal cord during such repairs have become critical in preventing devastating complications. Endovascular therapy for such conditions in selected patients has become accepted practice, producing fewer adverse outcomes than traditional approaches.

THORACIC AORTIC ANEURYSMS

Aortic aneurysm is defined as a permanent, localized dilatation of the aorta to a diameter that is at least 50% greater than is normal at that anatomic level. The annual incidence of thoracic aortic aneurysms is estimated to be 5.9 per 100,000 persons. The clinical manifestations, methods of treatment, and treatment results in patients with aortic aneurysms vary according to the cause and the aortic segment involved. Causes of thoracic aortic aneurysms include degenerative disease of the aortic wall, aortic dissection, aortitis, infection, and trauma. Aneurysms can be localized to a single aortic segment, or they can involve multiple segments. Thoracoabdominal aortic aneurysms, for example, involve both the descending thoracic aorta and the abdominal aorta. In the most extreme cases, the entire aorta is aneurysmal; this condition is often called mega-aorta.

Aortic aneurysms can be either “true” or “false.” True aneurysms can take two forms: fusiform and saccular. Fusiform aneurysms are more common and can be described as symmetrical dilatations of the aorta. Saccular aneurysms are localized outpouchings of the aorta. False aneurysms, also called pseudoaneurysms, are leaks in the aortic wall that are contained by the outer layer of the aorta and/or the periaortic tissue; they are caused by disruption of the aortic wall and lead blood to collect in pouches of fibrotic tissue.

Aneurysms of the thoracic aorta consistently increase in size and eventually progress to cause serious complications. These include rupture, which usually is a fatal event. Therefore, aggressive treatment is indicated in all but the poorest surgical candidates. Small, asymptomatic thoracic aortic aneurysms can be followed, especially in high-surgical-risk patients, and can be treated surgically later if symptoms or complications develop, or if progressive enlargement occurs. Meticulous control of hypertension is the primary medical treatment for patients with small, asymptomatic aneurysms.

Elective resection with graft replacement is indicated in asymptomatic patients with an aortic diameter of at least twice...
**Key Points**

1. Assessing urgency of repair is essential to developing the appropriate management plan. Although emergent repair carries greater operative risk than does elective repair, any inappropriate delay of repair risks death.

2. The clinical progression of an aortic aneurysm is continued expansion and eventual dissection or rupture. Hence, regular noninvasive imaging studies, as part of a lifelong surveillance plan, are necessary to ensure long-term patient health. Even small asymptomatic aneurysms should be routinely imaged to assess overall size and yearly rate of expansion.

3. Endovascular repair devices are approved for the treatment of descending thoracic aortic aneurysms, descending thoracic aortic dissections, aortic trauma, and penetrating aortic ulcer.

4. Practice guidelines have been published to help standardize the decision-making process and select an appropriate surgical intervention, as well as to standardize the use of imaging studies for patients with thoracic aortic disease.

5. Ascending aortic aneurysms that are symptomatic or \(\geq 5.5\) cm in diameter should be repaired regardless of whether normal in the involved segment (5 to 6 cm in most thoracic segments). Elective repair is contraindicated by extreme operative risk due to severe coexisting cardiac or pulmonary disease and by other conditions that limit life expectancy, such as malignancy. An emergency operation is performed for any patient in whom a ruptured aneurysm is suspected.

Patients with thoracic aortic aneurysm often have coexisting aneurysms of other aortic segments. A common cause of death after repair of a thoracic aortic aneurysm is rupture of a different aortic aneurysm. Therefore, staged repair of multiple aortic segments often is necessary. As with any major operation, careful preoperative evaluation for coexisting disease and subsequent medical optimization are important for successful surgical treatment.

An alternative to traditional open repair of a descending thoracic aortic aneurysm is endovascular stent grafting. Certain anatomic criteria for use—such as a landing zone that includes at least 2 cm of landing zone of healthy aortic tissue proximal and distal to the targeted aneurysm—are preferable, but not absolutely necessary. Although few data on long-term outcomes have recently been published, endovascular repair of descending thoracic aortic aneurysm has become an accepted practice that produces excellent midterm results.

3. Ascending aortic aneurysms that are symptomatic or \(\geq 5.5\) cm in diameter should be repaired regardless of whether normal in the involved segment (5 to 6 cm in most thoracic segments). Elective repair is contraindicated by extreme operative risk due to severe coexisting cardiac or pulmonary disease and by other conditions that limit life expectancy, such as malignancy. An emergency operation is performed for any patient in whom a ruptured aneurysm is suspected.

Patients with thoracic aortic aneurysm often have coexisting aneurysms of other aortic segments. A common cause of death after repair of a thoracic aortic aneurysm is rupture of a different aortic aneurysm. Therefore, staged repair of multiple aortic segments often is necessary. As with any major operation, careful preoperative evaluation for coexisting disease and subsequent medical optimization are important for successful surgical treatment.

An alternative to traditional open repair of a descending thoracic aortic aneurysm is endovascular stent grafting. Certain anatomic criteria for use—such as a landing zone that includes at least 2 cm of landing zone of healthy aortic tissue proximal and distal to the targeted aneurysm—are preferable, but not absolutely necessary. Although few data on long-term outcomes have recently been published, endovascular repair of descending thoracic aortic aneurysm has become an accepted practice that produces excellent midterm results.

Causes and Pathogenesis

**General Considerations.** The normal aorta derives its elasticity and tensile strength from the medial layer, which contains approximately 45 to 55 lamellae of elastin, collagen, smooth muscle cells, and ground substance. Elastin content is highest within the ascending aorta, as would be expected because of its compliant nature, and decreases distally into the descending and abdominal aorta. Maintenance of the aortic matrix involves complex interactions among smooth muscle cells, macrophages, proteases, and protease inhibitors. Any alteration in this delicate balance can lead to aortic disease.

Thoracic aortic aneurysms have a variety of causes (Table 22-1). Although these disparate pathologic processes differ in biochemical and histologic terms, they share the final common pathway of progressive aortic expansion and eventual rupture.

Hemodynamic factors clearly contribute to the process of aortic dilatation. The vicious cycle of increasing diameter and increasing wall tension, as characterized by Laplace’s law (tension = pressure × radius), is well established. Turbulent
blood flow also is recognized as a factor. Poststenotic aortic dilatation, for example, occurs in some patients with aortic valve stenosis or coarctation of the descending thoracic aorta. Hemodynamic derangements, however, are only one piece of a complex puzzle.

Atherosclerosis is commonly cited as a cause of thoracic aortic aneurysms. However, although atherosclerotic disease often is found in conjunction with aortic aneurysms, the notion that atherosclerosis is a distinct cause of aneurysm formation has been challenged. In most thoracic aortic aneurysms, atherosclerosis appears to be a coexisting process, rather than the underlying cause.

Research into the pathogenesis of abdominal aortic aneurysms has focused on the molecular mechanisms of aortic wall degeneration and dilatation. For example, imbalances between proteolytic enzymes (e.g., matrix metalloproteinases) and their inhibitors contribute to abdominal aortic aneurysm formation. Building on these advances, current investigations are attempting to determine whether similar inflammatory and proteolytic mechanisms are involved in thoracic aortic disease, in hope of identifying potential molecular targets for pharmacologic therapy.

**Nonspecific Medial Degeneration.** Nonspecific medial degeneration is the most common cause of thoracic aortic disease. Histologic findings of mild medial degeneration, including fragmentation of elastic fibers and loss of smooth muscle cells, are expected in the aging aorta. However, an advanced, accelerated form of medial degeneration leads to progressive weakening of the aortic wall, aneurysm formation, and eventual dissection, rupture, or both. The underlying causes of medial degenerative disease remain poorly understood.

**Aortic Dissection.** An aortic dissection usually begins as a tear in the inner aortic wall, which initiates a progressive separation of the medial layers and creates two channels within the aorta. This event profoundly weakens the outer wall. As the most common catastrophe involving the aorta, dissection represents a major, distinct cause of thoracic aortic aneurysms and is discussed in detail in the second half of this chapter.

**Heritable Conditions.** Several heritable conditions cause thoracic aortic aneurysms. To better characterize these disorders, the National Institutes of Health (NIH) sponsored a longitudinal registry for individuals affected by genetically triggered thoracic aneurysms and cardiovascular conditions (GenTAC) more than a decade ago. The registry enrollment includes adults and children in 13 clinical categories, including Marfan syndrome, Ehlers-Danlos syndrome, Loeys-Dietz syndrome, familial thoracic aortic aneurysms and dissections, aneurysms-osteoarthritis syndrome, and congenital bicuspid aortic valve.

**Marfan Syndrome** Marfan syndrome is an autosomal dominant genetic disorder characterized by a specific connective tissue defect that leads to aneurysm formation. The phenotype of patients with Marfan syndrome typically includes a tall stature, high palate, joint hypermobility, eye lens disorders, mitral valve prolapse, and aortic aneurysms. The aortic wall is weakened by fragmentation of elastic fibers and deposition of extensive amounts of mucopolysaccharides (a process previously called cystic medial degeneration or cystic medial necrosis). Patients with Marfan syndrome have a mutation in the fibrillin gene located on the long arm of chromosome 15. The traditional view has been that abnormal fibrillin in the extracellular matrix decreases connective tissue strength in the aortic wall and produces abnormal elasticity, which predisposes the aorta to dilatation from wall tension caused by left ventricular ejection impulses. More recent evidence, however, shows that the abnormal fibrillin causes degeneration of the aortic wall matrix by increasing the activity of transforming growth factor beta (TGF-β). Between 75% and 85% of patients with Marfan syndrome have dilatation of the ascending aorta and annuloaortic ectasia (dilatation of the aortic sinuses and annulus). Marfan syndrome also is frequently associated with aortic dissection, and aortic complications are the most common cause of death among patients with Marfan syndrome.

**Loeys-Dietz Syndrome** Loeys-Dietz syndrome is phenotypically distinct from Marfan syndrome. It is characterized as an aneurysmal syndrome with widespread systemic involvement. Loeys-Dietz syndrome is an aggressive, autosomal dominant condition that is distinguished by the triad of arterial tortuosity and aneurysms, hypertelorism (widely spaced eyes), and bifid uvula or cleft palate. It is caused by heterozygous mutations in the genes encoding TGF-β receptors. Patients with Loeys-Dietz syndrome—including young children—are at increased risk of aortic rupture and aortic dissection; diameter-based thresholds of repair tend to be lower for patients with this syndrome than for patients with other heritable disorders.

**Ehlers-Danlos Syndrome** Ehlers-Danlos syndrome includes a spectrum of inherited disorders of collagen synthesis. The subtypes represent differing defective steps of collagen production. Vascular type Ehlers-Danlos syndrome is characterized by an autosomal dominant defect in type III collagen synthesis, which can have life-threatening cardiovascular manifestations. Spontaneous arterial rupture, usually involving the mesenteric vessels, is the most common cause of death in these patients. Thoracic aortic aneurysms and dissections are less commonly associated with Ehlers-Danlos syndrome, but when they do occur, they pose a particularly challenging surgical problem because of the reduced integrity of the aortic tissue. An Ehlers-Danlos variant of periventricular heterotopia associated with joint and skin hyperextensibility and aortic dilation has been described as being caused by mutations in the gene encoding filamin A.
(FLNA), an actin-binding protein that links the smooth muscle cell contractile unit to the cell surface.\textsuperscript{12}

**Familial Thoracic Aortic Aneurysm and Dissection** Families without the heritable syndromes described earlier also can be affected by genetic conditions that cause thoracic aortic aneurysm. In fact, it is estimated that at least 20% of patients with thoracic aortic aneurysms and dissections have a genetic predisposition to them. The involved mutations are characterized by autosomal dominant inheritance with decreased penetrance and variable expression. The number of genes for which mutations have been identified as causes of familial thoracic aortic aneurysm and dissection is expanding rapidly; involved genes include those related to TGF-β receptors (TGFBR1 and TGFBR2), TGF-β ligands (TGFβ2 and TGFβ3), myosin (MYH11 and MYLK), elastin (ELN), elastin microfibril interfacial 1 (EMLI1), microfibril-associated glycoprotein 2 (MFAP5), fibrillin-2 (FBN2), fibulin-4 (FBLN4), lysyl oxidase (LOX), and α-smooth muscle cell actin (ACTA2).\textsuperscript{3,13-16} ACTA2 mutations are present in approximately 14% of families with familial thoracic aortic aneurysms and dissections.

**Aneurysms-Osteoarthritis Syndrome** Aneurysms-osteoarthritis syndrome is an autosomal dominant disorder characterized by aortic and arterial aneurysms, arterial tortuosity, aortic dissection, mild craniofacial abnormalities, and early-onset osteoarthritis. Aneurysms-osteoarthritis syndrome is caused by mutations in the gene encoding SMAD3, a transcription factor for TGF-β. Affected patients have a high incidence of aortic dissection, which often occurs in a mildly dilated aorta and causes sudden death.\textsuperscript{17}

**Congenital Bicuspid Aortic Valve** Bicuspid aortic valve is the most common congenital malformation of the heart or great vessels, affecting up to 2% of Americans.\textsuperscript{18} Compared to patients with a normal, trileaflet aortic valve, patients with bicuspid aortic valve have an increased incidence of ascending aortic aneurysm formation and, often, a more rapid rate of aortic enlargement.\textsuperscript{19} The location of the fused leaflet, or raphe, may be predictive of aortic dilation and other abnormalities.\textsuperscript{20} Fifty to 70% of adults with bicuspid aortic valve, but without significant valve dysfunction, have echocardiographically detectable aortic dilatation.\textsuperscript{21,22} This dilatation usually is limited to the ascending aorta and root.\textsuperscript{23} Dilation occasionally is found in the arch and only rarely in the descending or abdominal aorta. In addition, aortic dissection occurs 10 times more often in patients with bicuspid valves than in the general population.\textsuperscript{24} Recent findings suggest that aneurysms associated with bicuspid aortic valve have a fundamentally different pathobiologic cause than aneurysms that occur in patients with trileaflet valves.\textsuperscript{25}

Although the exact mechanism responsible for aneurysm formation in patients with bicuspid aortic valve remains unclear, evidence suggests that these patients have a congenital connective tissue abnormality that predisposes the aorta to medial degeneration.\textsuperscript{25-31} For example, fibrillin-1 content is significantly lower and matrix metalloproteinase activity is significantly higher in the aortic media in patients with bicuspid aortic valve than in persons with a normal, tricuspid aortic valve.\textsuperscript{26-27} Further, the process of medial degeneration in patients with bicuspid aortic valve may be exacerbated by the presence of chronic turbulent flow through the deformed valve.

**Bovine Aortic Arch** Bovine aortic arch—a common origin of the innominate and left common carotid arteries—has been considered a normal anatomic variant. Studies from Yale University have identified a higher prevalence of bovine aortic arch in patients with thoracic aortic disease; an association was found between this anomaly and a generalized increase in aortic aneurysmal disease (without any predisposition to a particular aortic region). However, bovine aortic arch was not associated distinctly with bicuspid aortic valve or aortic dissection, but with a higher mean aortic growth rate: 0.29 cm per year in patients with bovine aortic arch, compared with 0.09 cm per year in controls. Therefore, bovine aortic arch may be better characterized as a precursor of aortic aneurysm than as a simple normal anatomic variant.\textsuperscript{32} Further studies are needed to delineate the underlying mechanism for this association.

**Infection.** Primary infection of the aortic wall resulting in aneurysm formation is rare. Although these lesions are termed mycotic aneurysms, the responsible pathogens usually are bacteria rather than fungi. Bacterial invasion of the aortic wall may result from bacterial endocarditis, endothelial trauma caused by an aortic jet lesion, or extension from an infected laminar clot within a preexisting aneurysm. The most common causative organisms are *Staphylococcus aureus, Staphylococcus epidermidis, Salmonella,* and *Streptococcus.*\textsuperscript{33,34} Unlike most other causes of thoracic aortic aneurysms, which generally produce fusiform aneurysms, infection often produces saccular aneurysms located in areas of aortic tissue destroyed by the infectious process.

Although syphilis was once the most common cause of ascending aortic aneurysms, the advent of effective antibiotic therapy has made syphilitic aneurysms a rarity in developed nations. In other parts of the world, however, syphilitic aneurysms remain a major cause of morbidity and mortality. The spirochete *Treponema pallidum* causes an obliteratorative endarteritis of the vasa vasorum that results in medial ischemia and loss of the elastic and muscular elements of the aortic wall. The ascending aorta and arch are the most commonly involved areas. The emergence of HIV infection in the 1980s was associated with a substantial increase in the incidence of syphilis in both HIV-positive and HIV-negative patients. Because syphilitic aortitis often presents 10 to 30 years after the primary infection, the incidence of associated aneurysms may increase in the near future.

**Aortitis.** In patients with preexisting degenerative thoracic aortic aneurysms, localized transmural inflammation and subsequent fibrosis can develop. The dense aortic infiltrate responsible for the fibrosis consists of lymphocytes, plasma cells, and giant cells. The cause of the intense inflammatory reaction is unknown. Although the severe inflammation is a superimposed problem rather than a primary cause, its onset within an aneurysm can further weaken the aortic wall and precipitate expansion.

Systemic autoimmune disorders also cause thoracic aortitis. Aortic Takayasu arteritis generally produces obstructive lesions related to severe intimal thickening, but associated medial necrosis can lead to aneurysm formation. In patients with giant cell arteritis (temporal arteritis), granulomatous inflammation may develop that involves the entire thickness of the aortic wall, causing intimal thickening and medial destruction. Rheumatoid aortitis is an uncommon systemic disease that is associated with rheumatoid arthritis and ankylosing spondylitis. The resulting medial inflammation and fibrosis can affect the aortic root, causing annular dilatation, aortic valve regurgitation, and ascending aortic aneurysm formation.

**Pseudoaneurysms.** Pseudoaneurysms of the thoracic aorta usually represent chronic leaks that are contained by surrounding...
tissue and fibrosis. By definition, the wall of a pseudoaneurysm is not formed by intact aortic tissue; rather, the wall develops from organized thrombus and associated fibrosis. Pseudoaneurysms can arise from primary defects in the aortic wall (e.g., after trauma or contained aneurysm rupture) or from anastomotic or cannulation site leaks that occur after cardiovascular surgery. Anastomotic pseudoaneurysms can be caused by technical problems or by deterioration of the native aortic tissue, graft material, or suture. Commonly, they occur in patients with Marfan syndrome, Loeys-Dietz syndrome, or other heritable conditions that markedly weaken the vessel wall. Tissue deterioration usually is related to either progressive degenerative disease or infection. Improvements in sutures, graft materials, and surgical techniques have decreased the incidence of thoracic aortic pseudoaneurysm. Should thoracic aortic pseudoaneurysms occur, they typically require expeditious open surgical or catheter-based repair because they are associated with a high incidence of morbidity and rupture.

Clinical History
Treatment decisions in cases of thoracic aortic aneurysm are guided by our current understanding of the clinical history of these aneurysms, which classically is characterized as progressive aortic dilatation and eventual dissection, rupture, or both.

An analysis by Elefteriades of data from 1600 patients with thoracic aortic disease has helped quantify these well-recognized risks. Average expansion rates were 0.07 cm per year in ascending aortic aneurysms and 0.19 cm per year in descending thoracic aortic aneurysms. As expected, aortic diameter was a strong predictor of rupture, dissection, and mortality. For thoracic aortic aneurysms >6 cm in diameter, annual rates of catastrophic complications were 3.6% for rupture, 3.7% for dissection, and 10.8% for death. Critical “hinge-point” diameters, at which the incidence of expected complications significantly increased, were 6.0 cm for aneurysms of the ascending aorta and 7.0 cm for aneurysms of the descending thoracic aorta; the corresponding risks of rupture after reaching these diameters were 31% and 43%, respectively.

Certain types of aneurysms have an elevated propensity for expansion and rupture. For example, aneurysms in patients with Marfan or Loeys-Dietz syndrome tend to dilate at an accelerated rate and rupture or dissect at smaller diameters than sporadic, nonheritable aneurysms. Before the era of surgical treatment for aortic aneurysms, the aggressive form of aortic disease in Marfan patients resulted in an average life expectancy of 32 years, with aortic root complications causing the majority of deaths. Saccular aneurysms, which commonly are associated with aortic infection and typically affect only a discrete small section of the aorta, tend to grow more rapidly than fusiform aneurysms, which are associated with more widespread degenerative changes and generally affect a larger section of the aorta.

One common clinical scenario deserves special attention. A moderately dilated ascending aorta (i.e., 4 to 5 cm) often is encountered during aortic valve replacement or coronary artery bypass operations. The clinical history of these ectatic ascending aortas has been defined by several studies. Michel and colleagues studied patients whose ascending aortic diameters were >4 cm at the time of aortic valve replacement; 25% of these patients required reoperation for ascending aortic replacement. Prenger and colleagues reported that aortic dissection occurred in 27% of patients who had aortic diameters of >5 cm at the time of aortic valve replacement. Attention has been directed toward whether or not a mildly dilated aortic root should be replaced in patients with bicuspid aortic valve who are undergoing isolated valve replacement, and at what threshold to intervene. Although this is a controversial issue, many surgeons believe that the tendency toward late aortic dilatation in these patients warrants aggressive treatment. According to a recent guidelines clarification, in patients with bicuspid aortic valve who are undergoing aortic valve replacement or repair, replacing the ascending aorta is reasonable when the diameter of the ascending aorta is greater than 4.5 cm (Class Ila, Level C recommendation).

Clinical Manifestations
In many patients with thoracic aortic aneurysms, the aneurysm is discovered incidentally when imaging studies are performed for unrelated reasons. Therefore, patients often are asymptomatic at the time of diagnosis. However, thoracic aortic aneurysms that initially go undetected eventually create symptoms and signs that correspond with the segment of aorta that is involved. These aneurysms have a wide variety of manifestations, including compression or erosion of adjacent structures, aortic valve regurgitation, distal embolism, and rupture.

Local Compression and Erosion. Initially, aneurysmal expansion and impingement on adjacent structures causes mild, chronic pain. The most common symptom in patients with ascending aortic aneurysms is anterior chest discomfort; the pain is frequently precordial in location but may radiate to the neck and jaw, mimicking angina. Aneurysms of the ascending aorta and transverse aortic arch cause symptoms related to compression of the superior vena cava, the pulmonary artery, the airway, or the sternum. Rarely, these aneurysms erode into the superior vena cava or right atrium, causing acute high-output failure. Expansion of the distal aortic arch can stretch the recurrent laryngeal nerve, which results in left vocal cord paralysis and hoarseness. Descending thoracic and thoracoabdominal aneurysms frequently cause back pain localized between the scapulae. When the aneurysm is largest in the region of the aortic hiatus, it may cause middle back and epigastric pain. Thoracic or lumbar vertebral body erosion typically causes severe, chronic back pain; extreme cases can present with spinal instability and neurologic deficits from spinal cord compression. Although mycotic aneurysms have a peculiar propensity to destroy vertebral bodies, spinal erosion also occurs with degenerative aneurysms. Descending thoracic aortic aneurysms may cause varying degrees of airway obstruction, manifesting as cough, wheezing, stridor, or pneumonitis. Pulmonary or airway erosion presents as hemoptysis. Compression and erosion of the esophagus cause dysphagia and hematemesis, respectively. Thoracoabdominal aortic aneurysms can cause duodenal obstruction or, if they erode through the bowel wall, gastrointestinal bleeding. Jaundice due to compression of the liver or porta hepatitis is uncommon. Erosion into the inferior vena cava or iliac vein presents with an abdominal bruit, widened pulse pressure, edema, and heart failure.

Aortic Valve Regurgitation. Ascending aortic aneurysms can cause displacement of the aortic valve commissures and annular dilatation. The resulting deformation of the aortic valve leads to progressively worsening aortic valve regurgitation. In response to the volume overload, the heart remodels and becomes increasingly dilated. Patients with this condition may present with progressive heart failure, a widened pulse pressure, and a diastolic murmur.

Distal Embolization. Thoracic aortic aneurysms—particularly those involving the descending and thoracoabdominal aorta—are commonly lined with friable, atheromatous plaque and
mural thrombus. This debris may embolize distally, causing occlusion and thrombosis of the visceral, renal, or lower-extremity branches.

**Rupture.** Patients with ruptured thoracic aortic aneurysms often experience sudden, severe pain in the anterior chest (ascending aorta), upper back or left chest (descending thoracic aorta), or left flank or abdomen (thoracoabdominal aorta). When ascending aortic aneurysms rupture, they usually bleed into the pericardial space, producing acute cardiac tamponade and death. Descending thoracic aortic aneurysms rupture into the pleural cavity, producing a combination of severe hemorrhagic shock and respiratory compromise. External rupture is extremely rare; saccular syphilitic aneurysms have been observed to rupture externally after eroding through the sternum.

**Diagnostic Evaluation**

Diagnosis and characterization of thoracic aneurysms require imaging studies, which also provide critical information that guides the selection of treatment options. Although the best choice of imaging technique for the thoracic and thoracoabdominal aorta is somewhat institution-specific, varying with the availability of imaging equipment and expertise, efforts have been made to standardize key elements of image acquisition and reporting. Recent practice guidelines recommend that aortic imaging reports plainly state the location of aortic abnormalities (including calcification and the extent to which abnormalities extend into branch vessels), the maximum external aortic diameters (rather than internal, lumen-based diameters), internal filling defects, and any evidence of rupture. Whenever possible, all results should be compared with those of prior imaging studies.

**Plain Radiography.** Plain radiographs of the chest, abdomen, or spine often provide enough information to support the initial diagnosis of thoracic aortic aneurysm. Ascending aortic aneurysms produce a convex shadow to the right of the cardiac silhouette. The anterior projection of an ascending aneurysm results in the loss of the retrosternal space in the lateral view. An aneurysm may be indistinguishable from elongation and tortuosity. Importantly, chest radiographs (CXRs) may appear normal in patients with thoracic aortic disease and thus cannot exclude the diagnosis of aortic aneurysm. Aortic root aneurysms, for example, often are hidden within the cardiac silhouette. Plain CXRs may reveal convexity in the right superior mediastinum, loss of the retrosternal space, or widening of the descending thoracic aortic shadow, which may be highlighted by a rim of calcification outlining the dilated aneurysmal aortic wall. Aortic calcification also may be seen in the upper abdomen on a standard radiograph made in the anteroposterior or lateral projection (Fig. 22-2). Once a thoracic aortic aneurysm is detected on plain radiographs, additional studies are required to define the extent of aortic involvement.

**Echocardiography and Abdominal Ultrasonography.** Ascending aortic aneurysms are commonly discovered during echocardiography in patients presenting with symptoms or signs of aortic valve regurgitation. Both transthoracic and transesophageal echocardiography provide excellent visualization of the ascending aorta, including the aortic root. Transesophageal echocardiography also allows visualization of the descending thoracic aorta but is not ideal for evaluating the transverse aortic arch (which is obscured by air in the tracheobronchial tree) or the upper abdominal aorta. Effective echocardiography requires considerable technical skill, both in obtaining adequate images and in interpreting them. This imaging modality has the added advantage of being less expensive than CT or MRI and providing real-time imaging during intervention. Advantages of abdominal ultrasonography include the ability to identify mural thrombus, intraluminal thrombus, and atheroemboli, as well as the location of an aneurysm.

**Figure 22-2.** Chest radiographs showing a calcified rim (arrows) in the aortic wall of a thoracoabdominal aortic aneurysm. A. Anteroposterior view. B. Lateral view.
benefit of assessing cardiac function and revealing any other abnormalities that may be present. During ultrasound evaluation of a suspected infrarenal abdominal aortic aneurysm, if a definitive neck cannot be identified at the level of the renal arteries, the possibility of thoracoabdominal aortic involvement should be suspected and investigated by using other imaging modalities. Caution should be exercised while interpreting aneurysm dimensions from ultrasound imaging because intraluminal measurements are often reported, whereas external measurements are usually used in other imaging modalities.

**Computed Tomography.** Computed tomographic (CT) scanning is widely available, provides visualization of the entire thoracic and abdominal aorta, and permits multplanar and 3-dimensional aortic reconstructions. Consequently, CT is the most common—and arguably the most useful—imaging modality for evaluating thoracic aortic aneurysms. In addition to establishing the diagnosis, CT provides information about an aneurysm’s location, extent, anatomic anomalies, and relationship to major branch vessels. CT is particularly useful in determining the absolute diameter of the aorta, especially in the presence of laminated clot, and also detects aortic calcification. Contrast-enhanced CT provides information about the aortic lumen and can detect mural thrombus, aortic dissection, inflammatory periaortic fibrosis, and mediastinal or retroperitoneal hematoma due to contained aortic rupture. To increase consistency and ensure uniform reporting, current practice guidelines suggest that measurements be taken perpendicular to blood flow and at standard anatomic locations, (Fig. 22-3); this should reduce the likelihood of erroneous measurements, especially during serial imaging surveillance.

The major disadvantage of contrast-enhanced CT scanning is the possibility of contrast-induced acute renal failure in patients who are at risk (e.g., patients with preexisting renal disease or diabetes) even though the risk is smaller than was assumed in the past. If possible, surgery is performed at least 1 day after contrast administration to allow time to observe renal function and to permit diuresis. If renal insufficiency occurs or is worsened, elective surgery is postponed until renal function returns to normal or stabilizes.

**Magnetic Resonance Angiography.** Magnetic resonance angiography (MRA) is becoming widely available and can facilitate visualization of the entire aorta. This modality produces aortic images comparable to those produced by contrast-enhanced CT but does not necessitate exposure to ionizing radiation. In addition, MRA offers excellent visualization of branch-vessel details, and it is useful in detecting branch-vessel stenosis. However, MRA is limited by high expense and a susceptibility to artifacts created by ferromagnetic materials, and gadolinium—the contrast agent for MRA—may be linked to nephrogenic systemic fibrosis and acute renal failure in patients with advanced renal insufficiency. Furthermore, the MRA environment is not appropriate for many critically ill patients, and unlike CT imaging, MRA imaging is suboptimal in patients with extensive aortic calcification.

**Invasive Aortography and Cardiac Catheterization.** Although catheter-based contrast aortography was previously considered the gold standard for evaluating thoracic aortic disease, cross-sectional imaging (i.e., CT and MRA) has largely replaced this modality. Technologic improvements have enabled CT and MRA to provide excellent aortic imaging while causing less morbidity than catheter-based studies do, so CT and MRA are now the primary modes for evaluating thoracic aortic disease. Today, the use of invasive aortography in patients with thoracic aortic disease is generally limited to those undergoing endovascular therapies or when other types of studies are contraindicated or have not provided satisfactory results.

Unlike standard aortography, cardiac catheterization continues to play an important role in diagnosis and preoperative planning, especially in patients with ascending aortic involvement. Proximal aortography can reveal not only the status of the coronary arteries and left ventricular function but also the degree of aortic valve regurgitation, the extent of aortic root involvement, coronary ostial displacement, and the relationship of the aneurysm to the arch vessels.

The value of the information one can obtain from catheter-based diagnostic studies should be weighed against
the established limitations and potential complications of such studies. A key limitation of aortography is that it images only the lumen and may therefore underrepresent the size of large aneurysms that contain laminated thrombus. Manipulation of intraluminal catheters can result in embolization of laminated thrombus or atheromatous debris. Proximal aortography carries a 0.6% to 1.2% risk of stroke. Other risks include allergic reaction to the contrast agent, iatrogenic aortic dissection, and bleeding at the arterial access site. In addition, the volumes of contrast agent required to adequately fill large aneurysms can cause significant renal toxicity. To minimize the risk of contrast nephropathy, patients receive periprocedural intravenous (IV) fluids for hydration, mannitol for diuresis, and acetylcysteine. As with contrast-enhanced CT, surgery is performed ≥1 day after angiography whenever possible to ensure that renal function has stabilized or returned to baseline.

**Treatment**

**Selecting the Appropriate Treatment.** Once a thoracic aortic aneurysm is detected, management begins with patient education, particularly if the patient is asymptomatic, because aortic disease may progress rapidly and unexpectedly in some patients. A detailed medical history is collected, a physical examination is performed, and a systematic review of medical records is carried out to clearly assess the presence or absence of pertinent symptoms and signs, despite any initial denial of symptoms by the patient. Signs of heritable conditions such as Marfan syndrome or Loey-Dietz syndrome are thoroughly reviewed. If clinical criteria are met for a heritable condition, confirmatory laboratory tests are conducted. Patients with heritable disorders are best treated in a dedicated aortic clinic where they can be appropriately followed up. Surveillance imaging and aggressive blood pressure control are the mainstays of initial management for asymptomatic patients. When patients become symptomatic or their aneurysms grow to meet certain size criteria, the patients become surgical candidates.

Endovascular therapy has become an accepted treatment for descending thoracic aortic aneurysm. Its role in treating proximal aortic disease and thoracoabdominal aortic aneurysm remains experimental; nonetheless, endoluminal stenting is approved by the U.S. Food and Drug Administration for the treatment of isolated descending thoracic aortic aneurysm, and several different devices have been approved for the treatment of blunt aortic injury and penetrating aortic ulcer. In practice, however, the off-label application of aortic stent grafts is widespread and accounts for well over half their use; endovascular approaches may be helpful in emergent aneurysm repair, such as for patients with aortic rupture. Endovascular therapy has evolved to include hybrid repairs, which combine open “debranching” techniques (to reroute branching vessels) with endovascular aortic repair. Despite these advances, for the repair of aneurysms with proximal aortic involvement and of thoracoabdominal aortic aneurysms, open procedures remain the gold standard and preferred approach.

**Determination of the Extent and Severity of Disease.** Cross-sectional imaging with reconstruction is critical when one is evaluating a thoracic aneurysm, determining treatment strategy, and planning necessary procedures. Note that patients with a thoracic aortic aneurysm may also have a second, remote aneurysm. In such cases, the more threatening lesion is addressed first. In many patients, staged operative procedures are necessary for complete repair of extensive aneurysms involving the ascending aorta, transverse arch, and descending thoracic or thoracoabdominal aorta. When the descending segment is not disproportionately large (compared with the proximal aorta) and is not causing symptoms, the proximal aortic repair is carried out first. An important benefit of this approach is that it allows treatment of valvular and coronary artery occlusive disease at the first operation.

Proximal aneurysms (proximal to the left subclavian artery) usually are addressed via a sternotomy approach. Aneurysms involving the descending thoracic aorta are evaluated in terms of criteria (described in the following section) for potential endovascular repair; those unsuitable for an endovascular approach are repaired with open techniques through a left thoracotomy. A CT scan can reveal detailed information about aortic calcification and luminal thrombus. These details are important in preventing embolization during surgical manipulation.

**Indications for Operation**

Thoracic aortic aneurysms are repaired to prevent fatal rupture. Therefore, on the basis of clinical history studies and other data, practice guidelines for thoracic aortic disease recommend elective operation in asymptomatic patients when the diameter of an ascending aortic aneurysm is ≥5.5 cm, when the diameter of a descending thoracic aortic aneurysm is ≥6.0 cm, or when the rate of dilatation is ≥0.5 cm per year. In patients with heritable disorders such as Marfan and Loey-Dietz syndromes, the threshold for operation is based on a smaller aortic diameter (5.0 cm for the ascending aorta in patients with Marfan syndrome, 4.4 to 4.6 cm for the ascending aorta in patients with Loey-Dietz syndrome, and <6.0 cm for the descending thoracic aorta in patients with either disorder). For women with heritable disorders who are considering pregnancy, prophylactic aortic root replacement is considered because the risk of aortic dissection or rupture increases at an aortic diameter of 4.0 cm and greater. For patients with ascending aortic aneurysm and bicuspid aortic valve, repair is recommended if aortic diameter is 5.0 cm or greater and additional risk factors are present (e.g., family history of dissection, expansion rate exceeding 0.5 cm per year), if aortic diameter is 5.5 cm or larger and no additional risk factors are present, or if aortic diameter exceeds 4.5 cm and the patient is undergoing aortic valve replacement or repair. For low-risk patients with chronic aortic dissection, descending thoracic repair is recommended at an aortic diameter of 5.5 cm or greater.

The acuity of presentation is a major factor in decisions about the timing of surgical intervention. Many patients are asymptomatic at the time of presentation, so there is time for thorough preoperative evaluation and improvement of their current health status, such as through smoking cessation and other optimization programs. In contrast, patients who present with symptoms may need urgent operation. Symptomatic patients are at increased risk of rupture and warrant expeditious evaluation. The onset of new pain in patients with known aneurysms is especially concerning because it may herald significant expansion, leakage, or impending rupture. Emergent intervention is reserved for patients who present with aneurysm rupture or superimposed acute dissection.

**Open Repair vs. Endovascular Repair** As noted earlier, endovascular repair has become the standard approach for patients with isolated degenerative descending thoracic aortic aneurysm; in fact, practice guidelines recommend that endovascular repair be strongly considered for patients with descending thoracic aneurysm at an aortic diameter of 5.5 cm (which is slightly below the 6.0-cm threshold for open repair). For endovascular
reparis to produce optimal outcomes, several anatomic criteria must be met. For one, the proximal and distal neck diameters should fall within a range that will allow proper sealing. Also, the proximal and distal landing zones should ideally be at least 20 mm long so that an appropriate seal can be made. Note that the limiting structures proximally and distally are the brachiocephalic vessels and celiac axis, respectively. Vascular access continues to be one of the most important determinants of successful deployment of the current endovascular devices. The femoral and iliac arteries have to be wide enough to accommodate the sheaths used to deploy the stent grafts. As endovascular technology evolves, newer devices are being used for smaller sheaths (or are “sheathless” self-deployed stent grafts) to accommodate smaller arteries. Tortuosity of the iliac vessels and abdominal aorta can make these procedures technically challenging. Occasionally, an 8- or 10-mm polyester “side graft” is anastomosed to the iliac artery through a retroperitoneal incision if the femoral vessels are too small to access easily.

Of note, attempts have been made to extend the use of endovascular therapy to aortic arch aneurysms and thoracoabdominal aortic aneurysms. Although reports of purely endovascular repair of the aortic arch remain limited, Greenberg and colleagues have reported their experience with a large series of purely endovascular thoracoabdominal aortic repairs. Additionally, there have been numerous reports of small series of off-label, experimental hybrid procedures that involve debranching of the aortic arch or the visceral vessels of the abdominal aorta, followed by endovascular exclusion of the aneurysm. The majority of hybrid approaches involve repairing the aortic arch. In its simplest form, hybrid arch repair involves an open bypass from the left subclavian to the left common carotid artery, which is followed by deliberate coverage of the origins of the left subclavian artery by the stent graft. In its most complex form, hybrid arch repair involves rerouting all of the brachiocephalic vessels, followed by proximal placement of the stent graft in the ascending aorta and extending repair distally into the aortic arch and descending thoracic aorta.

The patients who theoretically benefit the most from an endovascular approach are those who are of advanced age or have significant comorbidities, as many of these patients face substantial risks when undergoing traditional open repair. For example, with regard to open repair of a descending thoracic aortic aneurysm, significant pulmonary morbidity can occur postoperatively; therefore, patients with borderline pulmonary reserve may better tolerate an endovascular procedure than a standard open repair. Patients with heritable syndromic conditions generally are not considered candidates for elective endovascular repair except in specific circumstances. Endovascular repair in patients with heritable syndromic conditions have produced poor results, which are mainly due to progressive dilatation, stent graft migration, and endoleak.

Preoperative Assessment and Preparation. Given the impact of comorbid conditions on perioperative complications, a careful preoperative assessment of physiologic reserve is critical in assessing operative risk. Therefore, most patients undergo a thorough evaluation—with emphasis on cardiac, pulmonary, and renal function—before undergoing elective surgery.

Cardiac Evaluation Coronary artery disease is common in patients with thoracic aortic aneurysm and is responsible for a substantial proportion of early and late postoperative deaths in such patients. Similarly, valvular disease and myocardial dysfunction have important implications when one is planning anesthetic management and surgical approaches for aortic repair. Transthoracic echocardiography is a satisfactory noninvasive method for evaluating both valvular and biventricular function. Dipyridamole-thallium myocardial scanning identifies regions of myocardium that have reversible ischemia, and this test is more practical than exercise testing in older patients with concomitant lower-extremity peripheral vascular disease. Cardiac catheterization and coronary arteriography are performed in patients who have evidence of coronary disease—as indicated by either the patient’s history or the results of noninvasive studies—or who have a left ventricular ejection fraction of ≤30%. If significant valvular or coronary artery disease is identified before a proximal aortic operation, the disease can be addressed directly during the procedure. Patients who have asymptomatic distal aortic aneurysms and severe coronary occlusive disease undergo percutaneous transluminal angioplasty or surgical revascularization before the aneurysmal aortic segment is replaced.

Pulmonary Evaluation Pulmonary function screening with arterial blood gas measurement and spirometry is routinely performed before thoracic aortic operations. Patients with a forced expiratory volume in 1 second of >1.0 L and a partial pressure of carbon dioxide of <45 mmHg are considered appropriate candidates for open surgical repair. In suitable patients, borderline pulmonary function can be improved by implementing a regimen that includes smoking cessation, weight loss, exercise, and treatment of bronchitis for a period of 1 to 3 months before surgery. Although surgery is not withheld from patients with symptomatic aortic aneurysms and poor pulmonary function, adjustments in operative technique should be made to maximize these patients’ chances of recovery. In such patients, preserving the left recurrent laryngeal nerve, the phrenic nerves, and diaphragmatic function is particularly important.

Renal Evaluation Renal function is assessed preoperatively by measuring serum electrolyte, blood urea nitrogen, and creatinine levels. Information about kidney size and perfusion can be obtained from the imaging studies used to evaluate the aorta.

Obtaining accurate information about baseline renal function has important therapeutic and prognostic implications. For example, perfusion strategies and perioperative medications are adjusted according to renal function. Patients with severely impaired renal function frequently require at least temporary hemodialysis after surgery. These patients also have a mortality rate that is significantly higher than normal. Patients with thoracoabdominal aortic aneurysms and poor renal function secondary to severe proximal renal occlusive disease undergo renal artery endarterectomy, stenting, or bypass grafting during the aortic repair.

Operative Repair Proximal Thoracic Aortic Aneurysms

Open Repair Traditional open operations to repair proximal aortic aneurysms—which involve the ascending aorta, transverse aortic arch, or both—are performed through a midsternal incision and require cardiopulmonary bypass. The best choice of aortic replacement technique depends on the extent of the aneurysm and the condition of the aortic valve. The spectrum of operations (Fig. 22-4) ranges from simple graft replacement of the tubular portion of the ascending aorta only (Fig. 22-4A) to replacement of the ascending aorta and the proximal aortic arch (Fig. 22-4B) to graft replacement of the entire proximal aorta, including the aortic root, and reattachment of the coronary
Figure 22-4. Illustrations of proximal aortic repairs in which the native aortic root is left intact. A. Graft replacement of the tubular portion of the ascending aorta with the aortic arch left intact. B. Hemiarach beveled graft replacement, in which the ascending aorta and a portion of the lesser curvature of the aortic arch are replaced. C. A modified arch with additional graft replacement of the innominate artery. D. Patch repair of the aortic arch. E. Traditional total arch replacement using an island approach to reattach the brachiocephalic vessels. F. The branched graft approach, which replaces the brachiocephalic vessels by following their original anatomic location. G. The elephant trunk approach with a concomitant island brachiocephalic artery reattachment. Contemporary Y-graft arch repairs include (H) the single Y-graft approach, (I) the double Y-graft approach, (J) the elephant trunk approach with a single Y-graft, and (K) the elephant trunk approach with a double Y-graft.
ar teries and brachi o cephalic branches. The options for treating aortic valve disease, repairing aortic aneurysms, and maintaining perfusion during repair procedures each deserve detailed consideration (Table 22-2).

### Options for graft repair of the aortic aneurysm
- Patch aortoplasty
- Ascending replacement only
- Beveled hemiarch replacement
- Total arch replacement with reattachment of brachi o cephalic branches (island technique)
- Elephant trunk technique with island reattachment
- Total arch repair with bypass grafts to the brachi o cephalic branches (including Y-graft approaches)
- Elephant trunk technique with Y-graft approach
- Hybrid aortic arch repairs (including “frozen elephant trunk technique”)

### Perfusion options
- Standard cardiopulmonary bypass
- Profound hypothermic circulatory arrest without adjuncts
- Hypothermic circulatory arrest with adjuncts
- Retrograde cerebral perfusion
- Antegrade cerebral perfusion
- Balloon perfusion catheters
- Right axillary artery cannulation
- Innominate artery cannulation

### Options for treating aortic valve disease
- Aortic valve annuloplasty (annular plication)
- Aortic valve replacement (with mechanical or biologic prosthesis)
- Aortic root replacement
  - Composite valve graft (with mechanical or biologic valve)
  - Aortic homograft
  - Stentless porcine root
  - Pulmonary autograft (Ross procedure)
- Valve-sparing techniques

### Options for open surgical repair of proximal aortic aneurysms

#### Table 22-2

<table>
<thead>
<tr>
<th>Options for graft repair of the aortic aneurysm</th>
<th>Perfusion options</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patch aortoplasty</td>
<td>Standard cardiopulmonary bypass</td>
</tr>
<tr>
<td>Ascending replacement only</td>
<td>Profound hypothermic circulatory arrest without adjuncts</td>
</tr>
<tr>
<td>Beveled hemiarch replacement</td>
<td>Hypothermic circulatory arrest with adjuncts</td>
</tr>
<tr>
<td>Total arch replacement with reattachment of brachi o cephalic branches (island technique)</td>
<td>Retrograde cerebral perfusion</td>
</tr>
<tr>
<td>Elephant trunk technique with island reattachment</td>
<td>Antegrade cerebral perfusion</td>
</tr>
<tr>
<td>Total arch replacement with bypass grafts to the brachi o cephalic branches (including Y-graft approaches)</td>
<td>Balloon perfusion catheters</td>
</tr>
<tr>
<td>Elephant trunk technique with Y-graft approach</td>
<td>Right axillary artery cannulation</td>
</tr>
<tr>
<td>Hybrid aortic arch repairs (including “frozen elephant trunk technique”)</td>
<td>Innominate artery cannulation</td>
</tr>
</tbody>
</table>

Aortic Valve Disease and Root Aneurysms

Many patients undergoing proximal aortic operations have aortic valve disease that requires concomitant surgical correction. When such disease is present and the sinus segment is normal, separate repair or replacement of the aortic valve and graft replacement of the tubular segment of the ascending aorta are carried out. In such cases, mild to moderate valve regurgitation with annular dilatation can be addressed by plicating the annulus with mattress sutures placed below each commissure, thereby preserving the native valve. In patients with more severe valvar regurgitation or with valvular stenosis, the valve is replaced with a stented biologic or mechanical prosthesis; mechanical prostheses necessitate following a lifelong anticoagulation regimen.

Separate replacement of the aortic valve and ascending aorta is not performed in patients with Marfan syndrome or Loeys-Dietz syndrome, because progressive dilatation of the remaining sinus segment eventually leads to complications that necessitate reoperation. Therefore, patients with Marfan syndrome, Loeys-Dietz syndrome, or annuloaortic ectasia require some form of aortic root replacement.

In many cases, the aortic root is replaced with a mechanical or biologic prosthesis that has both a valve and an aortic conduit. Currently, the following graft options are commercially available: composite valve grafts with a mechanical valve, which consist of a bileaflet mechanical valve attached to a polyester tube graft; composite valve grafts with a biological valve (available in Europe only at this point); aortic root homografts, which are harvested from cadavers; and stentless porcine root grafts. In the United States, because no biologic composite valve grafts are commercially available, another option for surgeons is to construct a bioprosthetic composite valve graft during the operation by suturing a stented tissue valve to a polyester tube graft.

Although select patients may be offered the Ross procedure—in which the patient’s pulmonary artery root is excised and placed in the aortic position and then the right ventricular outflow tract is reconstructed by using a cryopreserved pulmonary homograft—this option is rarely used. This is largely because it is a technically demanding procedure, and there are concerns about the potential for autograft dilatation in patients with heritable conditions.

An additional option is valve-sparing aortic root replacement, which has evolved substantially during the past decade. The valve-sparing technique that is currently favored is called aortic root reimplantation and involves excising the aortic sinuses, attaching a prosthetic graft to the patient’s annulus (Fig. 22-5), and resuspended the native aortic valve inside the graft. The superior hemodynamics of the native valve and the avoidance of anticoagulation are major advantages of the valve-sparing approach. Long-term results in carefully selected patients have been excellent. Although the durability of this procedure in patients with Marfan syndrome has been satisfactory in some centers, it remains uncertain whether long-term durability can be reliably achieved with this approach. Further, acceptable mid-term outcomes have been reported for patients with bicuspid aortic valve. Patients who have structural leaflet deterioration or excessive annular dilatation are typically deemed unsuitable for valve-sparing repair.

Regardless of the type of conduit used, aortic root replacement requires reattaching the coronary arteries to openings in the graft. In the original procedure described by Bentall and De Bono, this was accomplished by suturing the intact aortic wall surrounding each coronary artery to the openings in the graft. The aortic wall was then wrapped around the graft to establish hemostasis. However, this technique frequently produced leaks at the coronary reattachment sites that eventually led to pseudoaneurysm formation. Cabrol’s modification, in which a separate, small tube graft is sutured to the coronary ostia and the main aortic graft, achieves tension-free coronary anastomoses, and reduces the risk of pseudoaneurysm formation. Kouchoukos’ button modification of the Bentall procedure is currently the most widely used technique. The aneurysmal aorta is excised, and buttons of aortic wall are left surrounding both coronary arteries, which are then mobilized and sutured to the aortic graft (Fig. 22-6). The coronary suture lines may be reinforced with polytetrafluoroethylene felt or pericardium...
Figure 22-5. Illustration of our current valve-sparing procedure for replacing the aortic root and ascending aorta for treatment of (A) aortic root aneurysm. B. The ascending aorta is opened after cardiopulmonary bypass and cardioplegic arrest are established and the distal ascending aorta is clamped. The diseased aortic tissue (including the sinuses of Valsalva) is excised. Buttons of surrounding tissue are used to mobilize the origins of the coronary arteries. C. A synthetic graft is sewn to the distal ascending aorta with continuous suture. D. After the distal anastomosis is completed, six sutures reinforced with Teflon pledgets are placed in the plane immediately below the aortic valve annulus. E. The subannular sutures are placed through the base of a synthetic aortic root graft, which is then is parachuted down around the valve. F. After the root graft is cut to an appropriate length, the valve commissures and leaflets are positioned within the graft. The annular sutures are then tied. G. Each of the three commissures is then secured near the top of the graft. H. The supra-annular aortic tissue is sewn to the graft in continuous fashion. I. The button surrounding the origin of the left main coronary artery is sewn to an opening cut in the root graft. J. The two aortic grafts are sewn together with continuous suture. K. The button surrounding the origin of the right coronary artery is sewn to an opening cut in the root graft. L. The completed valve-sparing aortic root replacement and graft repair of the ascending aorta are shown. (Used with permission of Baylor College of Medicine.)

to enhance hemostasis. When the coronary arteries cannot be mobilized adequately because of extremely large aneurysms or scarring from previous surgery, the Cabrol technique or a related modification can be used. Another option, originally described by Zubiate and Kay, is the construction of bypass grafts by using interposition saphenous vein or synthetic grafts.

Aortic Arch Aneurysms Several options are also available for handling aneurysms that extend into the transverse aortic arch.
with arch debranching followed by exclusion of the aneurysm with an endovascular graft. For fusiform aneurysms, when the distal portion of the arch is a reasonable size, a single, beveled replacement of the lower curvature (hemiarch) is performed. More extensive arch aneurysms require total replacement involving a distal anastomosis to the proximal descending thoracic aorta and separate reattachment of the brachiocephalic branches. The brachiocephalic vessels can be reattached to one or more openings made in the graft, or if these vessels are aneurysmal, they can be replaced with separate, smaller grafts. Alternatively, a Y-graft approach can be used to debranch the brachiocephalic vessels and move them forward, thereby permitting the distal anastomosis to be brought forward, which aids in hemostasis. When the aneurysm involves the entire arch and extends into the descending thoracic aorta, it is approached by using Borst’s elephant trunk technique of staged total arch replacement. The distal anastomosis is performed such that a portion of the graft is left suspended within the proximal descending thoracic aorta (Fig. 22-7). A collared graft can be used to accommodate any discrepancy in aortic diameter. During a subsequent operation, the suspended “trunk” is used to facilitate repair of the remaining descending thoracic or thoracoabdominal aortic aneurysm by an endovascular technique or by open repair through a thoracotomy incision, depending on the extent of disease and other factors. The elephant trunk technique permits access to the distal portion of the graft during the second operation without the need for dissection around the distal transverse aortic arch; this reduces the risk of injuring the left recurrent laryngeal nerve, esophagus, and pulmonary artery if an open approach is used at the second stage. As described in the section on hybrid repair of arch aneurysms (see later), the elephant trunk can be completed by using a hybrid endovascular approach (Fig. 22-8) in certain settings. A newer technique that is currently under investigation involves using a graft comprising a conventional polyester proximal portion and a stent graft into the distal transverse arch. This “frozen elephant trunk” technique can enable treatment of the entire aortic pathology during a single procedure or can facilitate a subsequent endovascular procedure (Fig. 22-9).

**Cardiopulmonary Bypass Perfusion Strategies** Like the operations themselves, perfusion strategies used during proximal aortic surgery depend on the extent of the repair. Aneurysms that are isolated to the ascending segment can be replaced by using standard cardiopulmonary bypass and distal ascending aortic clamping. This provides constant perfusion of the brain and other vital organs during the repair. Aneurysms involving the transverse aortic arch, however, cannot be clamped during the repair, which necessitates the temporary withdrawal of cardiopulmonary bypass support; this is called *circulatory arrest*. To protect the brain and other vital organs during the circulatory arrest period, hypothermia must be initiated before pump flow is stopped. However, the deep levels of hypothermia (below 20°C) that have been traditionally used in open arch repair are not without risk, and pure hypothermic circulatory arrest continues to have substantial limitations. Importantly, although brief periods of total circulatory arrest generally are well tolerated at cold temperatures, as the duration of circulatory arrest increases, the well-recognized risks of brain injury and death increase dramatically. Additionally, deep levels of hypothermia are associated with coagulopathy.

Because of the inherent complexity of aortic arch repairs and their general tendency to require longer periods
of hypothermic circulatory arrest, two cerebral perfusion strategies—retrograde cerebral perfusion (RCP) and antegrade cerebral perfusion (ACP)—were developed to supplement this process by delivering cold, oxygenated blood to the brain and further reduce the risks associated with repair. Retrograde cerebral perfusion involves directing blood from the cardiopulmonary bypass circuit into the brain through the superior vena cava. However, RCP is thought to be less beneficial than ACP, and although it may be helpful in the retrograde flushing of air and debris from the arch, many centers have stopped using RCP.

In contrast, ACP delivers blood directly into the brachiocephalic arteries to maintain cerebral flow. Although its use was cumbersome in the past, contemporary ACP techniques have been simplified and commonly involve cannulating either the right axillary artery or the innominate artery and subsequent connection to the cardiopulmonary bypass circuit. Often, a small section of graft is used as a conduit to ease cannulation, but there remains a small procedure-related risk of brachial plexus or vascular injury. Upon initiation, cold blood is delivered into the brain via the right common carotid artery and, if bilateral ACP is desired, the left common carotid artery. Note that, with the unilateral ACP technique, blood flow to the left side of the brain requires collateral circulation, ideally through an intact circle of Willis.

Methods to help determine the adequacy of unilateral ACP to deliver cerebral cross-circulation include preoperative imaging and intraoperative monitoring. A commonly used method (Fig. 22-10) have been simplified and commonly involve cannulating either the right axillary artery or the innominate artery and subsequent connection to the cardiopulmonary bypass circuit. Often, a small section of graft is used as a conduit to ease cannulation, but there remains a small procedure-related risk of brachial plexus or vascular injury. Upon initiation, cold blood is delivered into the brain via the right common carotid artery and, if bilateral ACP is desired, the left common carotid artery. Note that, with the unilateral ACP technique, blood flow to the left side of the brain requires collateral circulation, ideally through an intact circle of Willis. Methods to help determine the adequacy of unilateral ACP to deliver cerebral cross-circulation include preoperative imaging and intraoperative monitoring. A commonly used method

Figure 22-7. Illustration of a contemporary Y-graft approach to total arch replacement for aortic arch aneurysm. A. The proximal portions of the brachiocephalic arteries are exposed. B. The first two branches of the graft are sewn end-to-end to the transected left subclavian and left common carotid arteries. The proximal ends of the transected brachiocephalic arteries are ligated. C. A balloon-tipped perfusion cannula is placed inside the double Y-graft and used to deliver antegrade cerebral perfusion. After systemic circulatory arrest is initiated, the innominate artery is clamped, transected, and sewn to the distal end of the main graft. D. The proximal aspect of the Y-graft is clamped. This directs flow from the axillary artery to all three brachiocephalic arteries. The arch is then replaced with a collared elephant trunk graft. E. The distal anastomosis between the elephant trunk graft and the aorta is created between the innominate and left common carotid arteries. The collared graft accommodates any discrepancy in aortic diameter. F. The aortic graft is clamped, and a second limb from the arterial inflow tubing of the cardiopulmonary bypass circuit is used to deliver systemic perfusion through a side-branch of the arch graft while the proximal portion of the ascending aorta is replaced. Once the proximal aortic anastomosis is completed, the main trunk of the double Y-graft is cut to an appropriate length, and the beveled end is then sewn to an oval opening created in the right anterolateral aspect of the ascending aortic graft, which completes the repair (G). (Modified with permission from LeMaire SA, Price MD, Parenti JL, et al. Early outcomes after aortic arch replacement by using the Y-graft technique, Ann Thorac Surg. 2011 Mar;91(3):700-707.)
Figure 22-8. Illustration of Borst’s elephant trunk technique using a contemporary Y-graft approach. A. Stage 1: The proximal repair includes replacing the ascending aorta and entire arch, with Y-graft reattachment of the brachiocephalic vessels. The distal anastomosis is facilitated by using a collared elephant trunk graft to accommodate the larger diameter of the distal aorta. A section of the graft is left suspended within the proximal descending thoracic aorta. B. Stage 2: The distal repair uses the floating “trunk” for the proximal anastomosis. C. An alternate “hybrid” approach may be used in patients with less extensive distal aortic disease. Endovascular stent grafts are placed within the elephant trunk to complete the repair. (Used with permission of Baylor College of Medicine.)

Figure 22-9. Illustration of a frozen elephant trunk repair, which is a hybrid approach to repair that combines open aortic replacement with endovascular aortic repair. A. Extensive aortic disease affects the proximal and distal aorta. B. Aortic repair is extended into the proximal portion of the descending thoracic aorta after the transverse aortic arch is fully replaced. (Used with permission of Baylor College of Medicine.)
of intraoperative monitoring is brain near-infrared spectroscopy (NIRS), which measures cerebral oxygenation. If NIRS monitoring indicates inadequate perfusion, an additional perfusion catheter can be inserted into the left common carotid artery to provide blood flow to the left side of the brain.

Because ACP provides excellent brain protection, many surgeons now use more moderate levels of hypothermia (often between 22°C and 24°C) to decrease the risks associated with deep hypothermia.67 However, some authors have raised the concern that reducing the degree of hypothermia increases the risk of ischemic complications involving the spinal cord, kidneys, and other organs that are ischemic (without the benefit of deep hypothermia) during the systemic circulatory arrest period.98 Consequently, some groups supplement ACP with additional perfusion strategies that provide flow to the descending aorta during arch repair.99,100

**Endovascular Repair** Experience with purely endovascular treatment of proximal aortic disease remains limited and only investigational.101 The unique anatomy of the aortic arch and the need for uninterrupted cerebral perfusion pose difficult challenges. There are reports of the use of “homemade” grafts to exclude arch aneurysms; however, these grafts are experimental at this time. For example, in 1999, Inoue and colleagues102 reported placing a triple-branched stent graft in a patient with an aneurysm of the aortic arch. The three brachiocephalic branches were positioned by placing percutaneous wires in the right brachial, left carotid, and left brachial arteries. The patient underwent two subsequent procedures: surgical repair of a right brachial pseudoaneurysm and placement of a distal stent graft extension to control a major perigraft leak. Since then, efforts to employ endovascular techniques in the treatment of the proximal aorta have been essentially limited to the use of approved devices for off-label indications, such as the exclusion of pseudoaneurysms in the ascending aorta.

**Hybrid Repair** In June 1991, the Ukrainian surgeon Nikolay Volodos and his colleagues performed the first hybrid aortic arch repair103,104; 22 years later, Volodos reported that the patient was still alive.105 Unlike purely endovascular approaches, hybrid repairs of the aortic arch have entered the mainstream clinical arena, although they remain controversial. Hybrid arch repairs involve some form of “debranching” of the brachiocephalic vessels (which are not unlike Y-graft approaches), followed by endovascular exclusion of some or all of the aortic arch (Fig. 22-11). Although this technique has many variants, they often involve sewing a branched graft to the proximal ascending aorta with the use of a partial aortic clamp. The branches of the graft are then sewn to the arch vessels. Once the arch is “debranched,” the arch aneurysm can be excluded with an endograft. Commonly, a zone 0 approach (Fig. 22-12) is undertaken in which the proximal end of the endograft is secured within the ascending aorta. Other hybrid approaches aim to extend repair into the distal arch and descending thoracic aorta (see the following section). The arguments for using a hybrid approach to treat aortic arch aneurysm include the potential to avoid using cardiopulmonary bypass, circulatory arrest, and cardiac ischemia.59,60

It is not yet clear whether hybrid repairs are as durable as traditional ones because little mid- or long-term data have been published, and there are very few comparative studies.65 Procedure-related risks include the risk of embolization and stroke due to wire and device manipulation within the aortic arch (this risk appears to be greatest in zone 0 repairs106), retrograde acute aortic dissection,107 type I endoleak,108 and paraplegia.27 Because iatrogenic retrograde dissection of the ascending aorta is a particularly lethal complication, special considerations, including careful blood pressure management and wire manipulation, are recommended to avoid this problem in patients who are undergoing hybrid arch zone 0 stent deployment.109 Notably, patients with an ascending aortic diameter greater than 4.2 cm may be more susceptible to retrograde dissection. In an effort to
reduce the risk of iatrogenic dissection, some centers have begun to replace a small section of the ascending aorta with a standard polyester graft such that the endograft’s proximal landing zone comprises prosthetic material rather than native aortic tissue.107

**Distal Thoracic Aortic Aneurysms**

**Open Repair** In patients with descending thoracic or thoracoabdominal aortic aneurysms, several aspects of treatment—including preoperative risk assessment, anesthetic management, choice of incision, and use of protective adjuncts—are dictated by the overall extent of aortic involvement. By definition, descending thoracic aortic aneurysms involve the portion of the aorta between the left subclavian artery and the diaphragm. Thoracoabdominal aneurysms can involve the entire thoracoabdominal aorta, from the origin of the left subclavian artery to the aortic bifurcation. Surgical repair of thoracoabdominal aortic aneurysms is categorized by the extent of aortic replacement according to the Crawford classification scheme (Fig. 22-13). Extent I thoracoabdominal aortic aneurysm repairs involve most of the descending thoracic aorta, usually beginning near the left subclavian artery, and extend down into the suprarenal abdominal aorta. Extent II repairs also begin near the left subclavian artery but extend distally into the infrarenal abdominal aorta, and they often reach the aortic bifurcation. Extent III repairs extend from the lower descending thoracic aorta (below the sixth rib) and into the abdomen. Extent IV repairs begin at the diaphragmatic hiatus and often involve the entire abdominal aorta.

Descending thoracic aortic aneurysms not amenable to endovascular therapy are currently repaired through a left thoracotomy. In patients with thoracoabdominal aortic aneurysm, the thoracotomy is extended across the costal margin and into the abdomen.110 Using a double-lumen endobronchial tube allows selective ventilation of the right lung and deflation of the left lung. Transperitoneal exposure of the thoracoabdominal aorta is achieved by performing medial visceral rotation and circumferential division of the diaphragm. During a period of aortic clamping, the diseased segment is replaced with a polyester tube graft. Important branch arteries—including intercostal arteries and the celiac, superior mesenteric, and renal arteries—are reattached to openings made in the side of the graft. In patients with Marfan syndrome and other heritable conditions, separate (8- and 10-mm) grafts to the visceral branches are often used to prevent subsequent “patch aneurysms” that can develop in residual aortic tissue.111 Visceral and renal artery occlusive disease is commonly encountered during aneurysm repair; options for correcting branch-vessel stenosis include endarterectomy, direct arterial stenting, and bypass grafting.

Clamping the descending thoracic aorta causes ischemia of the spinal cord and abdominal viscera. Clinically significant manifestations of hepatic, pancreatic, and bowel ischemia are relatively uncommon. However, both acute renal failure and spinal cord injury resulting in paraplegia or paraparesis remain major causes of morbidity and mortality after these operations. Therefore, several aspects of the operation are devoted to minimizing spinal and renal ischemia (Table 22-3). Our multimodal approach to spinal cord protection includes expeditious repair to minimize aortic clamping time, moderate systemic heparinization (1.0 mg/kg) to prevent small-vessel
Figure 22-13. Illustration of the Crawford classification of thoracoabdominal aortic aneurysm repair, based on the extent of aortic replacement. (Used with permission from Baylor College of Medicine.)

Figure 22-12. Illustration of the Criado landing zones, which are used to describe aortic anatomy during thoracic endovascular repair. The arch is the short segment that includes the origins of the three brachiocephalic arteries—the innominate artery, the left common carotid artery, and the left subclavian artery. Zone 0 includes the ascending aorta and the origin of the innominate artery. Zone 1 includes the origin of the left common carotid artery. Zone 2 includes the left subclavian artery origin. Zone 3 is a short section of the aorta that comprises the 2 cm immediately distal to the origin of the left subclavian artery, and zone 4 begins where zone 3 ends. (Used with permission from Baylor College of Medicine.)

### Table 22-3

**Current strategy for spinal cord and visceral protection during repair of distal thoracic aortic aneurysms**

<table>
<thead>
<tr>
<th>All extents</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Permissive mild hypothermia (32°C to 34°C, nasopharyngeal)</td>
</tr>
<tr>
<td>• Moderate heparinization (1 mg/kg)</td>
</tr>
<tr>
<td>• Aggressive reattachment of segmental arteries, especially between T8 and L1</td>
</tr>
<tr>
<td>• Sequential aortic clamping when possible</td>
</tr>
<tr>
<td>• Perfusion of renal arteries with 4°C crystalloid solution when possible</td>
</tr>
</tbody>
</table>

Crawford extent I and II thoracoabdominal repairs

- Cerebrospinal fluid drainage
- Left heart bypass during proximal anastomosis
- Selective perfusion of celiac axis and superior mesenteric artery during intercostal and visceral anastomoses

Thrombosis, mild permissive hypothermia (32°C to 34°C [89.6°F to 93.2°F] nasopharyngeal temperature), and reattachment of segmental intercostal and lumbar arteries. As the aorta is replaced from proximal to distal, the aortic clamp is moved sequentially to lower positions along the graft to restore perfusion to newly reattached branch vessels. During extensive thoracoabdominal aortic repairs (i.e., Crawford extent I and II repairs), cerebrospinal fluid drainage is used to improve spinal perfusion by reducing cerebrospinal fluid pressure. Because the benefits of this adjunct have been confirmed in a randomized clinical trial, its use is recommended in current guidelines (Class I, Level B recommendation). During cerebral spinal fluid drainage, the cerebral spinal fluid pressure is closely monitored, and the amount of fluid that is removed is carefully limited to avoid the devastating complication of intracranial hemorrhage. Motor evoked potentials are used by some groups to monitor the spinal cord throughout the operation. Left heart bypass, which provides perfusion of the distal aorta and its branches during the clamping period, is also used during extensive thoracoabdominal aortic repairs. Because left heart bypass unloads the heart, it is also useful in patients with poor cardiac reserve. Balloon perfusion cannulas connected to the left heart bypass circuit can be used to deliver blood directly to the celiac axis and superior mesenteric artery during their reattachment. The potential benefits of reducing hepatic and bowel ischemia include reduced risks of postoperative coagulopathy and bacterial translocation, respectively. Whenever possible, renal protection is achieved by perfusing the kidneys with cold (4°C [39.2°F]) crystalloid. In a randomized clinical trial, reduced kidney temperature was found to be associated with renal protection, and the use of cold crystalloid independently predicted preserved renal function. Hypothermic circulatory arrest can also be used during descending thoracic or thoracoabdominal aortic repairs. At our center, the primary indication for this approach is the inability to clamp the aorta because of rupture, extremely large aneurysm size, or extension of the aneurysm into the distal transverse aortic arch, or because a prior endovascular repair hinders clamping.

As discussed previously, complete repair of extensive aneurysm involving the ascending aorta, transverse arch, and descending thoracic aorta generally requires staged open
operations or a hybrid approach. In such procedures, when the descending or thoracoabdominal component is symptomatic (e.g., causes back pain or has ruptured) or is disproportionately large (compared with the ascending aorta), the distal segment is treated during the initial operation, and repair of the ascending aorta and transverse aortic arch is performed as a second procedure. A reversed elephant trunk repair, in which a portion of the proximal end of the aortic graft is inverted down into the lumen, can be performed during the first operation; this technique facilitates the second-stage repair of the ascending aorta and transverse aortic arch (Fig. 22-14).\(^{121}\)

Although spinal cord ischemia and renal failure receive the most attention, several other complications warrant consideration. The most common complication of extensive repairs is pulmonary dysfunction. With aneurysms adjacent to the left subclavian artery, the vagus and left recurrent laryngeal nerves are often adherent to the aortic wall and thus are susceptible to injury. Vocal cord paralysis should be suspected in patients who have postoperative hoarseness, and the presence of nerve damage should be confirmed by endoscopic examination. Vocal cord paralysis can be treated effectively by direct cord medialization (type 1 thyroplasty).\(^{122}\) Injury to the esophagus during

**Figure 22-14.** Illustration of the reversed elephant trunk technique using a traditional “island” approach to total aortic arch replacement. A. Stage 1: The distal aorta is repaired through a left thoracoabdominal approach. The aneurysm is opened after the aorta is clamped between the left common carotid artery and the left subclavian artery, which is also clamped. Before the proximal anastomosis is performed, the end of the graft is partly invaginated to leave a “trunk” for the subsequent repair. Proximal intercostal arteries are oversewn. B. After the proximal suture line is completed, the clamps are repositioned to restore blood flow to the left subclavian artery. The repair is completed by reattaching patent intercostal arteries to an opening in the side of the graft and creating a beveled distal anastomosis at the level of the visceral branches. C. Stage 2: The proximal aorta is repaired through a median sternotomy. The aortic arch is opened under hypothermic circulatory arrest. The “trunk” is pulled out and used to replace the aortic arch and ascending aorta. This eliminates the need for a new distal anastomosis and simplifies the procedure. Circulatory arrest and operative time, along with their attendant risks, are reduced. D. The completed two-stage repair of the entire thoracic aorta. (Modified with permission from Coselli JS, LeMaire SA, Carter SA, et al: The reversed elephant trunk technique used for treatment of complex aneurysms of the entire thoracic aorta, Ann Thorac Surg. 2005 Dec;80(6):2166-2172.)
the proximal anastomosis can have catastrophic consequences. Carefully separating the proximal descending thoracic aorta from the underlying esophagus before performing the proximal anastomosis minimizes the risk of a secondary aortoesophageal fistula. In patients who have previously undergone coronary artery bypass with a left internal thoracic artery graft, clamping proximal to the left subclavian artery can precipitate severe myocardial ischemia and cardiac arrest. When the need to clamp at this location is anticipated in these patients, a left common carotid-to-subclavian bypass is performed to prevent cardiac complications (Fig. 22-15).123

Endovascular Repair

**Descending Thoracic Aortic Aneurysms** Stent graft repair has become the standard treatment for patients with descending thoracic aortic aneurysm.55,56,124 Although aortic repair with a self-fixing endoprosthesis was reported by Volodos103,104 in the mid 1980s, it was the report by Parodi and associates125 of using endovascular stent grafting to repair abdominal aortic aneurysm that launched widespread interest in developing this approach. Only 3 years after this seminal report was published, Dake and colleagues126 reported performing endovascular descending thoracic aortic repair with “homemade” stent grafts in 13 patients.

Guidelines for the use of endovascular repair in thoracic aortic disease have been published,44 and reporting standards to uniformly describe the endovascular repair process have been established.127 Although endografting was initially approved to treat degenerative descending thoracic aortic aneurysm, newer devices have been approved for use in treating various descending thoracic aortic pathologies, including blunt aortic injury, penetrating aortic ulcer (see following section), coarctation, and dissection. Although the use of stent grafts in cases of aortic infection is not ideal, patients with a fistula or mycotic aneurysm are sometimes treated with endovascular devices as a bridge to open repair.

In elderly patients with severe comorbidity and patients who have undergone previous complex thoracic aortic procedures, endovascular repair is a particularly attractive alternative to standard open surgical procedures.128 Patients who undergo endovascular repair tend to have a lower incidence of intraoperative complications, a shorter length of stay, and a higher likelihood of being discharged to home than those who undergo open repair.129 As mentioned previously, appropriate patient selection depends on specific measurements taken from preoperative CT angiograms.

To protect patients against spinal cord ischemia during endovascular repair of the descending thoracic aorta, the most important maneuver is to keep the mean arterial perfusion pressure between 90 and 110 mmHg after the endograft is deployed. In patients who have had previous open or endovascular abdominal aortic aneurysm repair, cerebrospinal fluid drainage is recommended.130 The first step in the repair procedure is to obtain appropriate vascular access for the insertion of the thoracic stent graft. If the femoral artery will not accommodate the necessary

---

**Figure 22-15.** Illustration of a thoracoabdominal aortic aneurysm repair in a patient with a patent left internal thoracic artery-to-left anterior descending coronary artery graft. The proximal anastomosis is being performed while the aorta is clamped between the left common carotid and subclavian arteries. Myocardial perfusion is maintained through the carotid-subclavian bypass graft. (Modified with permission from Jones MM, Akay M, Murariu D, et al: Safe aortic arch clamping in patients with patent internal thoracic artery grafts. Ann Thorac Surg. 2010 Apr;89(4):e31-e32.)
sheath, then an iliac artery is exposed. If necessary, a graft can be sewn to the iliac artery in an end-to-side fashion to facilitate the deployment of the endograft. After 5,000 to 10,000 units of heparin are administered, a guidewire and the delivery sheath are typically inserted into the access artery under fluoroscopic guidance; depending on which endovascular device is used, the stent graft can be advanced through a sheath or with no sheath. The endograft is then advanced into the aorta and suitably positioned. Note that the best view of the distal arch and descending thoracic aorta is usually in the left anterior oblique position at an angle of approximately 40° to 50°. The device is then deployed, and the proximal and distal ends can be ballooned for better apposition of the stent graft to the aortic wall. An aortogram is then performed to rule out any endoleak, and protamine is administered. As an alternative to aortography, intravascular ultrasonography (IVUS) can be used to identify the proximal and distal landing zones, and the entire procedure can be performed with minimal or no contrast.

Although it is not uncommon to cover the left subclavian artery with the endograft to lengthen the proximal landing zone, findings suggest that the risk of spinal cord complications is heightened when the subclavian artery is covered and not revascularized, presumably because of a loss of collateral circulation to the spinal cord. To prevent this complication, a carotid-to-subclavian bypass can be easily constructed to maintain vertebral artery blood flow and minimize neurologic injury (Fig. 22-16). In addition, recent studies suggest that revascularization of the left subclavian artery is associated with lower stroke risk in patients in whom an endograft was deployed in Zone 2 and covered the left subclavian artery. In addition, new generations of stent grafts are being designed with side branches that can be placed within the left subclavian artery. This feature is particularly attractive if the proximal neck is short or if the patient has a patent left internal thoracic artery-to-left anterior descending coronary artery bypass. Indications for left subclavian artery revascularization include previous coronary artery bypass with patent internal thoracic artery, dominant left vertebral artery, aneurysm arising from the left subclavian artery, left arm arterio-venous fistula, and coverage of a long segment of the descending thoracic aorta.

**Elephant Trunk Completion** In select patients, elephant trunk completion repairs can be done with an endovascular approach (see Fig. 22-8C), rather than by the traditional open operation through a thoracotomy. Recall that an elephant trunk is used when an aortic aneurysm extends from the distal arch to the descending thoracic aorta. An endograft can be deployed at the time of elephant trunk construction or during a separate, subsequent procedure. When the stent is deployed in a retrograde manner during a second-stage procedure, the procedure is facilitated by placing radiopaque markers at the end of the elephant trunk during the first-stage procedure. This allows the distal end of the trunk to be identified via fluoroscopy. A guidewire can then be manipulated into the trunk and advanced into the ascending aorta to stabilize it during stent deployment. Note that advancing a wire in retrograde fashion from the femoral artery into the elephant trunk can be challenging. Occasionally, the wire must be advanced in an antegrade fashion from a brachial artery. The frozen elephant trunk technique—in which a short stent graft is delivered antegrade inside the trunk—can be used to perform the entire repair in one stage or to facilitate the second stage.

**Thoracoabdominal Aortic Aneurysms** Although endovascular thoracoabdominal aortic aneurysm repair remains experimental, it has been shown to be feasible in a handful of specialized centers. Endovascular thoracoabdominal aortic aneurysm repairs are quite complex, because at least one of the visceral arteries is incorporated into the repair. The number of visceral branches that need to be addressed varies with the extent of aortic coverage. The types of stent grafts used include fenestrated grafts, reinforced fenestrated grafts, branched or cuffed grafts, modular combinations of grafts, and multilayer stents. Graft fenestrations and branch vessels are typically aligned by using inflatable angioplasty balloons. Procedure time is not insignificant, nor is the amount of contrast medium required to obtain the highly detailed images needed to plan these procedures. In addition, some of the stent grafts used in endovascular thoracoabdominal aortic aneurysm repair are custom-made in advance and thus may take several weeks to obtain; therefore, their use is limited to cases of elective repair. In efforts to hasten repair and utilize off-the-shelf devices, parallel graft approaches, which use a combination of large- and small-diameter stents, have been reported. And, although some centers now propose distal coverage of the celiac axis for extent I thoracoabdominal aortic aneurysms, this potentially risky approach is not widely used.

It should be noted that, like open thoracoabdominal aortic aneurysm repair, endovascular repair carries risks of paraplegia, renal failure, stroke, and death, despite the apparent benefits of its being a less invasive procedure. Notably, reports from centers experienced in endovascular thoracoabdominal aortic repair primarily describe limited extent IV repairs. Although the technology is progressing rapidly, at present endovascular thoracoabdominal aortic aneurysm repair should be considered investigational.

**Hybrid Repair** Extensive hybrid thoracoabdominal aortic aneurysm repair can be a life-saving option in patients at high surgical risk, such as those who have limited physiologic reserve, are of advanced age, or have significant comorbidities. Hybrid procedures use open surgical techniques to reroute blood supply to the visceral arteries so that their aortic origins can be covered by stent grafts without causing visceral ischemia (Fig. 22-17). Endovascular methods are then used (either as part of the same procedure or at a later stage) to repair the aortic aneurysm, often with simple tube stent grafts; such devices are more readily available than the customized, modular stent grafts deployed in strictly endovascular repairs. Overall, results for hybrid thoracoabdominal aortic aneurysm repair have been somewhat disappointing. However, a handful of centers report acceptable outcomes in high-risk patients, particularly when a staged hybrid approach is used.

**Postoperative Considerations**

**Open Procedures** Aortic anastomoses are often extremely fragile during the early postoperative period. Even brief episodes of postoperative hypertension can disrupt suture lines and precipitate severe bleeding or pseudoaneurysm formation. Therefore, during the initial 24 to 48 hours, meticulous blood pressure control is maintained to protect the integrity of the anastomoses. Generally, we liberally use IV vasoactive agents to keep the mean arterial blood pressure between 80 and 90 mmHg. In patients with extremely friable aortic tissue, such as those with Marfan syndrome, we lower the target range to 70 to 80 mmHg. It is a delicate balancing act because one must be
Figure 22-16. Illustration of a “Zone 2” hybrid repair of the proximal descending thoracic aorta. A. The preoperative representation of the aneurysm shows that establishing a 2-cm proximal landing zone for a stent graft will require covering the origin of the left subclavian artery. B. Through a supraclavicular approach, a bypass from the left common carotid artery to the left subclavian artery is performed to reroute circulation and create a landing zone for the stent graft. After the bypass is completed, the left subclavian artery is ligated proximal to the graft. C. In the completed hybrid repair, the aneurysm has been excluded successfully by a stent graft that covers the origin of the left subclavian artery, and the proximal landing zone of the endograft is within zone 2. Importantly, blood flow to the left vertebral artery and arm is preserved by the bypass graft. (Reproduced with permission from Bozinovski J, LeMaire SA, Weldon SA: Hybrid Repairs of the Distal Aortic Arch and Proximal Descending Thoracic Aorta, Oper Tech Thorac Cardiovasc Surg 2007;12(3):167-177.)
mindful of spinal cord perfusion and avoid periods of relative hypotension while maintaining these low pressures.

Endovascular Procedures Many of the complications are directly related to manipulation of the delivery system within the iliac arteries and aorta. Patients with small, calcified, tortuous iliofemoral arteries are at particularly high risk for life-threatening iliac artery rupture. Although relatively uncommon, acute iatrogenic retrograde dissection into the aortic arch and ascending aorta is a life-threatening complication that necessitates emergency repair of the ascending aorta and aortic arch via sternotomy and cardiopulmonary bypass. The most important risk factors for this complication include incautious wire and catheter manipulation, aggressive proximal ballooning (especially in cases of acute descending thoracic aortic dissection), and hybrid arch repair in which the native ascending aorta is dilated (more than 4 cm). Retrograde proximal dissection converts a localized descending thoracic aortic aneurysm into an acute problem involving the entire thoracic aorta. Of note, retrograde aortic dissection may also occur several months after initial repair.

Another significant complication of descending thoracic aortic stent grafting is endoleak. An endoleak occurs when there is a persistent flow of blood (visible on radiologic imaging) into the aneurysm sac, and it may occur during the initial procedure or develop over time. Although endoleaks are a relatively common complication, they are not benign because they lead to continual pressurization of the sac, which can cause expansion or even rupture. These complications are categorized (Table 22-4) according to the site of the leak. Although all endoleaks may progress such that they can be considered life-threatening, type I and type III endoleaks generally necessitate early and aggressive intervention. Recently published reporting guidelines aid standardized reporting.

Other complications include stent graft misdeployment, device migration, endograft kinking or infolding, and stent graft infection, including fistula. Although not all complications related to stent grafts are fatal, endovascular repairs should be performed by expert teams qualified to address the variety of problems that may arise; some patients may need to have these devices removed and replaced with polyester grafts. Complications of endovascular repair are relatively common, so regularly scheduled radiologic imaging surveillance is of the utmost importance.

Table 22-4
Classification of and common treatment strategies for endoleak

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I</td>
<td>Incomplete seal between stent graft and aorta at the proximal landing site (Type Ia), the distal landing site (Type Ib), or branch module, fenestration, or plug (Type Ic)</td>
</tr>
<tr>
<td></td>
<td>Early reintervention to improve seal or conversion to open surgery</td>
</tr>
<tr>
<td>Type II</td>
<td>Retrograde perfusion of sac from excluded collateral arteries</td>
</tr>
<tr>
<td></td>
<td>Surveillance; as-needed occlusion with percutaneous or other interventions</td>
</tr>
<tr>
<td>Type III</td>
<td>Incomplete seal between overlapping stent graft or module (Type IIIa), or tear in graft fabric (Type IIIb)</td>
</tr>
<tr>
<td></td>
<td>Early reintervention to cover or conversion to open surgery</td>
</tr>
<tr>
<td>Type IV</td>
<td>Perfusion of sac due to porosity of material</td>
</tr>
<tr>
<td></td>
<td>Surveillance; as-needed reintervention to reline stent graft</td>
</tr>
<tr>
<td>Type V</td>
<td>Expansion of sac with no identifiable source</td>
</tr>
<tr>
<td></td>
<td>Surveillance; as-needed reintervention to reline stent graft</td>
</tr>
</tbody>
</table>
Aortic Dissection

Pathology and Classification

Aortic dissection, the most common catastrophic event involving the aorta, is a progressive separation of the aortic wall layers that usually occurs after a tear forms in the intima and inner media. As the separation of the layers of the media propagates, two channels are typically formed (Fig. 22-18): the original lumen, which remains lined by the intima and which is called the true lumen, and the newly formed channel within the layers of the media, which is called the false lumen. The dissecting membrane separates the true and false lumens. Additional tears in the dissecting membrane that allow communication between the two channels are called reentry sites. Although the separation of layers primarily progresses distally along the length of the aorta, it can also proceed in a proximal direction; this process often is referred to as proximal extension or retrograde dissection.

The extensive disruption of the aortic wall has severe anatomic consequences (Fig. 22-19). First, the outer wall of the false lumen is extremely thin, inflamed, and fragile, which makes it prone to expansion or rupture in the face of ongoing hemodynamic stress. Second, the expanding false lumen can compress the true lumen and cause malperfusion syndrome by interfering with blood flow in the aorta or any of its branch vessels, including the coronary, carotid, intercostal, visceral, renal, and iliac arteries. Finally, when the separation of layers occurs within the aortic root, the aortic valve commissures can become unhinged, which results in acute valvular regurgitation. The clinical consequences of each of these sequelae are addressed in detail in the section on clinical manifestations.

Dissection vs. Aneurysm. The relationship between dissection and aneurysmal disease requires clarification. Dissection and aneurysm are separate entities, although they often coexist and are mutual risk factors. In some cases, dissection occurs in patients without aneurysms, and the subsequent progressive dilatation of the weakened outer aortic wall ultimately results in an aneurysm. On the other hand, in patients with degenerative aneurysms, the ongoing deterioration of the aortic wall can lead to a superimposed dissection. The overused term dissecting aneurysm should be reserved for this specific situation.

Classification. For management purposes, aortic dissections are classified according to their location and chronicity. Improvements in imaging have increasingly revealed variants of aortic dissection that probably represent different forms along the spectrum of this condition.

Location To guide treatment, dissections are categorized according to their anatomic location and extent. The two traditional classification schemes that remain in common use are the DeBakey and the Stanford classification systems (Fig. 22-20). In their current forms, both of these schemes describe the segments of aorta that are involved in the dissection, rather than the site of the initial intimal tear. The main drawback of the Stanford classification system is that it does not distinguish between patients with isolated ascending aortic dissection and patients with dissection involving the entire aorta. Both types of patients would be classified as having type A dissections, despite the fact that their treatment, follow-up, and prognosis are substantially different.

Additional classification schemas include that by Borst and associates, in which the ascending and descending aorta are considered independently; the recent modification of the DeBakey classification by Tsagakis et al, which extends type II dissection into the aortic arch; and the Penn modification of the Stanford classification, which expands the classification to include the presence of tissue and global malperfusion. These modifications may help to better streamline the primary surgical intervention.

Figure 22-18. Illustration of longitudinal sections of the aortic wall and lumen. Blood flows freely downstream in normal aortic tissue. In classic aortic dissection, blood entering the media through a tear creates a false channel in the wall. Intramural hematomas arise when hemorrhage from the vasa vasorum causes blood to collect within the media; the intima is intact. Penetrating aortic ulcers are deep atherosclerotic lesions that burrow into the aortic wall and allow blood to enter the media. In each of these conditions, the outer aortic wall is severely weakened and prone to rupture.
Figure 22-19. Illustration of the potential anatomic consequences of aortic dissection, with a mapped diagram of affected regions (inset). A. Ascending aortic rupture and cardiac tamponade. B. Disruption of coronary blood flow. C. Injury to the aortic valve causing regurgitation. D, E, and F. Compromised blood flow to branch vessels, causing ischemic complications. (Adapted with permission from Creager MA, Dzau VS, Loscalzo J: Vascular Medicine, 7th ed. Philadelphia, PA: Elsevier/Saunders; 2006.)

Figure 22-20. Illustration of the classification schemes for aortic dissection based on which portions of the aorta are involved. Dissection can be confined to the ascending aorta (left) or the descending aorta (middle), or it can involve the entire aorta (right). (Used with permission from Baylor College of Medicine.)
Regardless of which system is used, patients with isolated ascending aortic dissection usually undergo emergent operation, as do patients with dissection involving both the ascending and descending thoracic segments. Patients with isolated descending thoracic and abdominal aortic dissection are typically treated medically, unless complications requiring surgery develop. Understanding the precise extent of dissection has become increasingly important as some aortic centers consider augmenting traditional ascending aortic repairs with endovascular techniques to treat dissected distal aortic segments.

**Chronicity** Aortic dissection also is categorized according to the time elapsed since the initial tear. Dissection is considered acute within the first 14 days after the initial tear; after 14 days, the dissection is considered chronic. Although arbitrary, the distinction between acute and chronic dissections has important implications, not only for decision making about perioperative management strategies and operative techniques, but also for evaluating surgical results. Figure 22-21 provides an algorithm for the management of acute aortic dissection. In light of the importance of acuity, Borst and associates\(^{154}\) have proposed a third phase—termed *subacute*—to describe the transition between the acute and chronic phases. The subacute period encompasses days 15 through 60 after the initial tear. Although this is past the traditional 14-day acute phase, patients with subacute dissection continue to have extremely fragile aortic

---

**Figure 22-21.** Algorithm used to facilitate decisions regarding treatment of acute aortic dissection. CT = computed tomography; MRA = magnetic resonance angiography; TEE = transesophageal echocardiography.
tissue, which may complicate operative treatment and increase the risks associated with surgery. Recently, the International Registry of Acute Aortic Dissections (IRAD) investigators proposed a new classification system for characterizing the phases of aortic dissection: The period within 24 hours from the onset of symptoms was defined as hyperacute, the period between 2 and 7 days was defined as acute, the period between 8 and 30 days was defined as subacute, and the period beyond 30 days was defined as chronic.159

**Variants** As noted earlier, advancements in noninvasive imaging of the aorta have revealed variants of aortic dissection (see Fig. 22-18). The recently introduced term acute aortic syndrome encompasses classic aortic dissection and its variants. Other aortic syndromes, which were once thought to be rare, include intramural hematoma (IMH) and penetrating aortic ulcer (PAU). Although the issue is somewhat controversial, the current consensus is that, in most cases, these variants of dissection should be treated identically to classic dissection.

An IMH is a collection of blood within the aortic wall, without an intimal tear, that is believed to be due to rupture of the vasa vasorum within the media. The accumulation of blood can result in a secondary intimal tear that ultimately leads to a dissection.160 Because IMH and aortic dissection represent a continuum, it is possible that IMH is seen less frequently than aortic dissection because IMH rapidly progresses to true dissection. The prevalence of IMH among patients with acute aortic syndromes is approximately 6%, and 16% progress to full dissection.161 An IMH can be classified according to its location (i.e., ascending or descending) and should be treated analogously to classic dissection.162

A PAU is essentially a disrupted atherosclerotic plaque that projects into the aortic wall and is associated with surrounding hematoma. Eventually, the ulcer can penetrate the aortic wall, which leads to dissection or rupture. The rate of disease progression is higher than that of IMH alone.163

**Causes and Clinical History**

Aortic dissection is a lethal condition with a reported incidence of 3.5 per 100,000 in the United States.164 Without appropriate modern medical or surgical treatment, most patients (approximately 90%) die within 3 months of dissection, mostly from rupture.165,166

Although several risk factors for aortic dissection have been identified, the specific causes remain unknown. Ultimately, any condition that weakens the aortic wall increases the risk of aortic dissection. Common general cardiovascular risk factors, such as smoking, hypertension, atherosclerosis, and hypercholesterolemia, are associated with aortic dissection. Patients with heritable forms of aortopathy, aortitis, bicuspid aortic valve, or preexisting medial degenerative disease are at increased risk for dissection, especially if they already have a thoracic aortic aneurysm.24 Aortic injury during cardiac catheterization, surgery, or endovascular aortic repair is a common cause of iatrogenic dissection. Other conditions that are associated with aortic dissection include cocaine and amphetamine abuse,167 as well as severe emotional stress or extreme physical exertion such as during weightlifting.168 Advances in the understanding of the molecular mechanisms behind abdominal aortic aneurysms have prompted similar investigations of thoracic aortic dissection.169-171

**Clinical Manifestations**

The onset of dissection often is associated with severe chest or back pain, classically described as “tearing,” that migrates distally as the dissection progresses along the length of the aorta. The location of the pain often indicates which aortic segments are involved. Pain in the anterior chest suggests involvement of the ascending aorta, whereas pain in the back and abdomen generally indicates involvement of the descending and thoracoabdominal aorta. Additional clinical sequelae of acute aortic dissection vary substantially and are best considered in terms of the dissection’s potential anatomic manifestations at each level of the aorta (see Fig. 22-19 and Table 22-5). Thus, potential complications of dissection of the aorta (and involved secondary arteries) may include cardiac ischemia (coronary artery) or tamponade, stroke (brachiocephalic arteries), paraplegia or paraparesis (intercostal arteries), mesenteric ischemia (superior mesenteric artery), kidney failure (renal arteries), and limb ischemia or loss of motor function (brachial or femoral arteries).

Ascending aortic dissection can directly injure the aortic valve, causing regurgitation. The severity of the regurgitation varies with the degree of commissural disruption, which ranges from partial separation of only one commissure, producing mild

**Table 22-5**

<table>
<thead>
<tr>
<th>ANATOMIC MANIFESTATION</th>
<th>SYMPTOMS AND SIGNS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aortic valve insufficiency</td>
<td>Dyspnea Murmur Pulmonary rales Shock</td>
</tr>
<tr>
<td>Coronary malperfusion</td>
<td>Chest pain with characteristics of angina Nausea/vomiting Shock Ischemic changes on electrocardiogram Elevated cardiac enzymes</td>
</tr>
<tr>
<td>Pericardial tamponade</td>
<td>Dyspnea Jugular venous distension Pulsus paradoxus Muffled cardiac tones Shock Low-voltage electrocardiogram</td>
</tr>
<tr>
<td>Subclavian or iliofemoral artery malperfusion</td>
<td>Cold, painful extremity Extremity sensory and motor deficits Peripheral pulse deficit</td>
</tr>
<tr>
<td>Carotid artery malperfusion</td>
<td>Syncope Focal neurologic deficit (transient or persistent) Carotid pulse deficit Coma</td>
</tr>
<tr>
<td>Spinal malperfusion</td>
<td>Paraplegia Incontinence</td>
</tr>
<tr>
<td>Mesenteric malperfusion</td>
<td>Nausea/vomiting Abdominal pain</td>
</tr>
<tr>
<td>Renal malperfusion</td>
<td>Oliguria or anuria Hematuria</td>
</tr>
</tbody>
</table>
SPECIFIC CONSIDERATIONS

PART II
valvular regurgitation, to full separation of all three commissures and complete prolapse of the valve into the left ventricle, producing severe acute heart failure. Patients with acute aortic valve regurgitation may report rapidly worsening dyspnea.

Ascending dissections also can extend into the coronary arteries or shear the coronary ostia off of the true lumen, causing acute coronary occlusion; when this occurs, it most often involves the right coronary artery. The sudden disruption of coronary blood flow can cause a myocardial infarction. This presentation of acute myocardial ischemia can mask the presence of aortic dissection, which results in delayed diagnosis and treatment.172

The thin and inflamed outer wall of a dissected ascending aorta often produces a serosanguineous pericardial effusion that can accumulate and cause tamponade. Suggestive signs include jugular venous distention, muffled heart tones, pulsus paradoxus, and low-voltage electrocardiogram (ECG) tracings. Free rupture into the pericardial space produces rapid tamponade and is generally fatal.

As the dissection progresses, any branch vessel from the aorta can become involved, which results in compromised blood flow and ischemic complications (i.e., malperfusion). Therefore, depending on which arteries are involved, the dissection can produce acute stroke, paraplegia, hepatic failure, bowel infarction, renal failure, or a threatened ischemic limb.

Diagnostic Evaluation
Because of the variations in severity and the wide variety of potential clinical manifestations, the diagnosis of acute aortic dissection can be challenging.173-175 Only 3 out of every 100,000 patients who present to an emergency department with acute chest, back, or abdominal pain are eventually diagnosed with aortic dissection. Not surprisingly, diagnostic delays are common; delays beyond 24 hours after hospitalization occur in up to 39% of cases. Unfortunately, delays in diagnosis lead to delays in treatment, which can have disastrous consequences. The European Society of Cardiology Task Force on Aortic Dissection stated, “The main challenge in managing acute aortic dissection is to suspect and thus diagnose the disease as early as possible.”173 A recent study by the IRAD investigators examined the reasons for delayed diagnosis and found that diagnosis lagged in women, as well as in patients with atypical symptoms, such as fever or mild pain (rather than severe pain).172 A high index of suspicion is critical, particularly in younger, atypical patients, who may have heritable disorders or other, less common risk factors.

Most patients with acute aortic dissection (80% to 90%) experience severe pain in the chest, back, or abdomen.173-175 The pain usually occurs suddenly, has a sharp or tearing quality, and often migrates distally as the dissection progresses along the aorta. For classification purposes (acute vs. subacute vs. chronic), the onset of pain is generally considered to represent the beginning of the dissection process. Most of the other common symptoms either are nonspecific or are caused by the secondary manifestations of dissection.

A discrepancy between the extremities in pulse, blood pressure, or both is the classic physical finding in patients with aortic dissection. It often occurs because of changes in flow in the true and false lumens, and it does not necessarily indicate extension into an extremity branch vessel. Involvement of the aortic arch often creates differences between the right and left arms, whereas descending aortic dissection often causes differences between the upper and lower extremities. Like symptoms, most of the physical signs after dissection are related to the secondary manifestations and therefore vary considerably (see Table 22-5). For example, signs of stroke or a threatened ischemic limb may dominate the physical findings in patients with carotid or iliac malperfusion, respectively.

Unfortunately, laboratory studies are of little help in diagnosing acute aortic dissection. There has been continued interest in using D-dimer level to aid in making this diagnosis.176 Several reports indicate that D-dimer is an extremely sensitive indicator of acute aortic dissection; elevated levels are found in approximately 97% of affected patients.177 Tests that are commonly used to detect acute coronary events—including ECG and tests for serum markers of myocardial injury—deserve special consideration and need to be interpreted carefully. Normal ECGs and serum marker levels in patients with acute chest pain should raise suspicion about the possibility of aortic dissection. It is important to remember that ECG changes and elevated serum marker levels associated with myocardial infarction do not exclude the diagnosis of aortic dissection because dissection can cause coronary malperfusion. Of note, abnormal ECGs have recently been shown to delay the diagnosis of aortic dissection, and the possibility of aortic dissection should not be prematurely ruled out.172,178 Similarly, although CXRs may show a widened mediastinum or abnormal aortic contour, up to 16% of patients with dissection have a normal-appearing CXR.174 The value of the CXR for detecting aortic dissection is limited, with a sensitivity of 67% and a specificity of 86%.179

Once the diagnosis of dissection is considered, the thoracic aorta should be imaged with CT, MRA, or echocardiography. The accuracy of these noninvasive imaging tests has all but eliminated the need for diagnostic aortography in most patients with suspected aortic dissection. Currently, the diagnosis of aortic dissection is usually established with contrast-enhanced CT, which has a sensitivity of 98% and a specificity of 87%, and, most importantly, acquires images swiftly.180 The classic diagnostic feature is a double-lumen aorta (Fig. 22-22). In addition, CT scans provide essential information about the segments of the aorta involved; the acuity of the dissection; aortic dilatation, including the presence of preexisting degenerative aneurysms; and the development of threatening sequelae, including pericardial effusion, early aortic rupture, and branch-vessel compromise. Although MRA also provides excellent imaging (with both a sensitivity and specificity of 98%), the MR suite is not well suited for critically ill patients. In patients who cannot undergo contrast-enhanced CT or MRA, transesophageal echocardiography can be used to establish the diagnosis.

Transesophageal echocardiography (TEE) is excellent for detecting dissection, aneurysm, and IMH in the ascending aorta. In appropriate hands, TEE has a demonstrated sensitivity and specificity as high as 98% and 95%, respectively.181 Furthermore, TEE offers important information about ventricular function and aortic valve competency. Finally, TEE is the diagnostic modality of choice for hemodynamically unstable patients in whom the diagnosis of ascending dissection is suspected; ideally, these patients should be taken to the operating room, where the TEE can be performed and, if the TEE is confirmatory, surgery can be started immediately.

In selected patients with ascending aortic dissection (i.e., those who have evidence of preexisting coronary artery disease), coronary angiography can be considered before surgery. Specific relative indications in these patients include a history
of angina or myocardial infarction, a recent myocardial perfusion study with abnormal results, previous coronary artery bypass or angioplasty, and acute ischemic changes on ECG. Contraindications include hemodynamic instability, aortic rupture, and pericardial effusion. In our practice, patients with acute aortic dissections rarely undergo coronary angiography. However, all patients presenting for elective repair of chronic ascending dissections have diagnostic coronary angiograms taken.

Of note, when malperfusion of the renal, visceral, or lower extremity arteries develops, the patient is usually treated in an angiography suite or hybrid operating room. Although the dissection usually is diagnosed on CT scan, these patients also undergo aortography, during which the mechanism of the malperfusion is ascertained and, if possible, corrected. Hence, catheter-based aortography may be obsolete as a diagnostic test for dissection, but it remains beneficial for patients with malperfusion.

**Treatment**

**Initial Assessment and Management.** Regardless of the location of the dissection, the initial treatment is the same for all patients with suspected or confirmed acute aortic dissection (see Fig. 22-21). Furthermore, because of the potential for rupture before the diagnosis is confirmed, aggressive pharmacologic management is started once there is clinical suspicion of dissection, and this treatment is continued during the diagnostic evaluation. The goals of pharmacologic treatment are to stabilize the dissection and prevent rupture.

Patients are monitored closely in an intensive care unit. Indwelling radial arterial catheters are used to monitor blood pressure and optimize titration of antihypertensive agents. Blood pressures in a malperfused limb can underrepresent the central aortic pressure; therefore, blood pressure is measured in the arm with the better pulse. Central venous catheters assure reliable IV access for delivering vasoactive medications. Pulmonary artery catheters are reserved for patients with severe cardiopulmonary dysfunction.

In addition to confirming the diagnosis of dissection and defining its acuity and extent, the initial evaluation focuses on determining whether any of several life-threatening complications are present. Particular attention is paid to changes in neurologic status, peripheral pulses, and urine output. Serial laboratory studies—including arterial blood gas concentrations, complete blood cell count, prothrombin and partial thromboplastin times, and serum levels of electrolytes, creatinine, blood urea nitrogen, and liver enzymes—are useful for detecting organ ischemia and optimizing management.

The initial management strategy, commonly described as *anti-impulse therapy* or *blood pressure control*, focuses on reducing aortic wall stress, the force of left ventricular ejection, chronotropy, and the rate of change in blood pressure (dP/dT). Reductions in dP/dT are achieved by lowering both cardiac contractility and blood pressure. The drugs initially used to accomplish these goals include IV β-adrenergic blockers, direct vasodilators, calcium channel blockers, and angiotensin-converting enzyme inhibitors. These agents are used to achieve a heart rate between 60 and 80 bpm, a systolic blood pressure between 100 and 110 mmHg, and a mean arterial blood pressure between 60 and 75 mmHg. These hemodynamic targets are maintained as long as urine output remains adequate and neurologic function is not impaired. Achieving adequate pain control with IV opiates, such as morphine and fentanyl, is important for maintaining acceptable blood pressure control.

β-Antagonists are administered to all patients with acute aortic dissections unless there are strong contraindications, such as severe heart failure, bradyarrhythmia, high-grade atrioventricular conduction block, or bronchospastic disease. Esmolol

---

**Figure 22-22.** Computed tomographic scans showing that the aorta has been separated into two channels—the true (T) and false (F) lumens—in two patients with different phases of aortic dissection. A. An acute DeBakey type I aortic dissection. The dissecting membrane appears wavy (arrows) in the early phase of dissection. Here, the true lumen of the proximal aorta can be seen to be extensively compressed. This may lead to malperfusion of the heart. B. A chronic DeBakey type III aortic dissection. In the chronic phase, the membrane appears straighter and less mobile (arrow) because it has stabilized over time. (Used with permission of Baylor College of Medicine.)
can be useful in patients with bronchospastic disease because it is a cardioselective, ultra-fast-acting agent with a short half-life. Labetalol, which causes both nonselective β-blockade and postsynaptic α1-blockade, reduces systemic vascular resistance without impairing cardiac output. Doses of β-agonists are titrated to achieve a heart rate of 60 to 80 bpm. In patients who cannot receive β-agonists, calcium channel blockers such as diltiazem are an effective alternative. Nitroprusside, a direct vasodilator, can be administered once β-blockade is adequate. When used alone, however, nitroprusside can cause reflex increases in heart rate and contractility, elevated dP/dT, and progression of aortic dissection. Enalapril and other angiotensin-converting enzyme inhibitors are useful in patients with renal malperfusion. These drugs inhibit renin release, which may improve renal blood flow.

**Treatment of Ascending Aortic Dissection**

**Acute Dissection** Because of the risk of aortic rupture, acute ascending aortic dissection is usually considered an absolute indication for emergency surgical repair. However, specific patient groups may benefit from nonoperative management or delayed operation.8183 Delayed repair may be considered for patients who (a) present with severe acute stroke or mesenteric ischemia, (b) are elderly and have substantial comorbidity, (c) are in stable condition and may benefit from transfer to specialized centers, or (d) have undergone a cardiac operation in the remote past. Regarding the last group, it is important that the previous operation not be too recent; dissections that occur during the first 3 weeks after cardiac surgery pose a high risk of rupture and tamponade, and such dissections warrant early operation.8184

In the absence of the aforementioned circumstances, most patients with acute ascending aortic dissection undergo emergent graft replacement of the ascending aorta. Operative repair is similar to that for aneurysm of the transverse aortic arch (previously described) because hypothermic circulatory arrest is commonly used regardless of the extent of repair. Immediately before the operation begins, intraoperative TEE is commonly performed to further assess baseline myocardial and valvular function and, if necessary, to confirm the diagnosis. The operation is performed via a median sternotomy with cardiopulmonary bypass and hypothermic circulatory arrest (Fig. 22-23). In preparation for circulatory arrest, cannulas are placed in the right axillary artery (to provide arterial inflow) and in the right atrium (to provide venous drainage).94 The innominate artery is sometimes used for arterial inflow if it is not dissected.985 After an appropriate level of cooling has been achieved (approximately 24°C), cardiopulmonary bypass is stopped, and the ascending aorta is opened. The innominate artery is then occluded with a clamp or snare, and flow from the axillary artery cannula is used to provide ACP.966 Currently as a default, we use bilateral ACP with a separate perfusion catheter in the left common carotid artery to ensure perfusion of the left side of the brain. This strategy of performing the distal anastomosis during a brief period of circulatory arrest, often termed open distal anastomosis, obviates the need to place a clamp across the fragile aorta, avoiding further aortic damage. Also, it allows the surgeon to carefully inspect the aortic arch for intimal tears. Traditionally, the entire arch is replaced only if a primary intimal tear is located in the arch or if the arch is aneurysmal; most commonly, repair is limited to replacement of the entire ascending aorta or to a beveled “hemiarch” repair.187 Conservative repair has been shown to increase the likelihood of early survival.188 The distal aortic cuff is prepared by tacking the inner and outer walls together and occasionally using a small amount of surgical adhesive to obliterate the false lumen and strengthen the tissue. A polyester tube graft is sutured to the distal aortic cuff. The anastomosis between the graft and the aorta is fashioned so that blood flow will be directed into the true lumen; this often alleviates distal malperfusion problems that were present preoperatively. After the distal anastomosis has been completed and adequately reinforced, the graft is deaired and clamped, full cardiopulmonary bypass is resumed, rewarming is initiated, and the proximal portion of the repair is started. In the absence of conditions that generally necessitate aortic root replacement (i.e., anuloaortic ectasia or heritable disorders, particularly Marfan and Loey-Dietz syndromes), aortic valve regurgitation can be corrected by resuspending the commissures onto the outer aortic wall.180 The proximal aortic cuff is prepared with tacking sutures and occasionally a small amount of surgical adhesive before the proximal aortic anastomosis is performed.

In the majority of patients who undergo surgical repair of acute ascending dissection, the dissection persists distal to the site of the operative repair; the residually dissected aorta, which generally includes at least a portion of the transverse aortic arch as well as a large portion of the distal aorta, is susceptible to dilatation over time. Extensive dilatation of the arch or distal aorta develops in 25% to 40% of survivors,190,191 and often necessitates further aortic repair. Additionally, long-term survival after acute proximal aortic dissection is generally poor, and rupture of the dilated distal aorta is a common cause of late death in these patients.188,190-192

The challenges that survivors of acute proximal aortic dissection commonly face over time have led to the development of alternative acute dissection strategies such as total arch replacement193 and hybrid arch strategies to extend proximal aortic repair into the distal aorta. The goal of hybrid arch approaches in acute dissection is to thrombose the residual false lumen by compressing it with the radial force that is exerted by a stent graft placed in the true lumen, thereby facilitating remodeling and preventing late aneurysm formation.194,195 However, in such repairs, the compressed false lumen may continue to be perfused in a retrograde fashion.

In Europe, Japan, and elsewhere, one-piece hybrid prostheses are now available that incorporate a polyester graft for the proximal repair and a stent graft component for the descending aorta. The device enables single-stage “frozen elephant trunk” repair of the ascending aorta, entire aortic arch, and proximal descending thoracic aorta.196 In the United States, such devices are not currently available, so this repair is commonly done by concomitantly deploying a commercially available stent graft in an antegrade fashion after fully replacing the ascending aorta and aortic arch. In some variations of this off-label approach, the stent graft is directly sutured to the distal aspect of the proximal open repair, whereas in others, there may be a gap of native tissue between the open and endovascular repair. Although this technique appears to be extensively used outside the United States, and with early and mid-term success,194,197-199 only a few U.S. reports describe its use.200-202 Emerging reports describe an enhanced risk of spinal cord ischemia, a risk that is not usually associated with open arch repair. This is probably due to the extensive coverage of the intercostal vessels by the stent graft. Uncertainties in the frozen elephant trunk procedure need to be addressed before it becomes a standard
Figure 22-23. Illustration of proximal aortic repair for acute ascending aortic dissection. A. This repair requires a median sternotomy and cardiopulmonary bypass. The ascending aorta is opened during hypothermic circulatory arrest, while antegrade cerebral perfusion is delivered via an axillary artery graft (shown) or via an innominate artery graft, provided that the innominate artery is not dissected (see Fig. 22-10). B. The dissecting membrane is removed to expose the true lumen. C. An open distal anastomosis prevents clamp injury of the friable arch tissue and allows inspection of the arch lumen. A balloon perfusion catheter in the left common carotid artery ensures bilateral antegrade cerebral perfusion. If the origin of the dissection (i.e., intimal tear or disruption) does not extensively involve the greater curvature of the aortic arch, and if there is no evidence of a preexisting arch aneurysm, a beveled, hemiarch repair is carried out, preserving most of the greater curvature of the arch. The aorta is transected, beginning at the greater curvature immediately proximal to the origin of the innominate artery and extending distally toward the lesser curvature to the level of the left subclavian artery. Consequently, most of the transverse aortic arch, except for the dorsal segment containing the brachiocephalic arteries, is removed. An appropriately sized, sealed (with collagen or gelatin) polyester tube graft is selected, and the beveled distal anastomosis is made with continuous 3-0 or 4-0 monofilament suture; the potential space between the true and false lumen can be obliterated with a small amount of surgical adhesive or by using a strip of Teflon felt. To improve hemostasis, the distal anastomosis can be reinforced by placing interrupted mattress sutures with felt pledgets. D. After cardiopulmonary bypass is resumed and a cross-clamp is applied to the hemiarch replacement graft, the aortic valve is assessed. Disrupted commissures are resuspended with pledgeted mattress sutures to restore valvular competence. E. The aorta is generally transected at the sinotubular junction, and a very small amount of surgical adhesive can be applied between the true and false lumens, or more commonly, the false lumen within the proximal aortic stump is obliterated by inserting a semicircle of felt within the false lumen of the noncoronary sinus. The trimmed edges are brought together by using 6-0 polypropylene sutures. F. The proximal anastomosis is carried out at the sinotubular junction, incorporating the distal margin of the commissures. G. In patients with residual distal aortic dissection (such as in DeBakey type I aortic dissection), hemiarch repair can be extended with antegrade stent delivery to the descending thoracic aorta. (Used with permission of Baylor College of Medicine.)
recommendation for this subset of patients. Another alternative employs separate grafts: a standard polyester graft to replace the ascending aorta and proximal hemiarch, and a stent graft delivered antegrade into the descending thoracic aorta (Fig. 22-23G). Although this procedure differs from a formal “frozen elephant trunk” repair in that it does not replace the entire arch, it is meant to achieve the same goal: promoting remodeling of the dissected descending aortic segment.

**Chronic Dissection** Occasionally, patients with ascending aortic dissection present for repair in the chronic phase. In most respects, the operation is similar to that for acute dissection repair. One notable difference is that the tissue is stronger in chronic dissection than in acute dissection, which makes suturing safer. In addition, the false lumen is not obliterated at the distal anastomosis; instead, the dissecting membrane is fenestrated into the arch to assure perfusion of both lumens and to prevent postoperative malperfusion complications. Unlike operations for acute dissection, operations for chronic dissection are often aggressive repairs that extend into the arch and root because the tissues are much less fragile.

**Treatment of Descending Aortic Dissection**

**Nonoperative Management** Nonoperative, pharmacologic management of acute descending aortic dissection results in lower morbidity and mortality rates than traditional open surgical treatment does. The most common causes of death during nonoperative treatment are aortic rupture and end-organ malperfusion. Therefore, patients are continually reassessed for new complications. Serial CT scans are generally obtained during the index hospitalization—usually on day 2 or 3 and on day 8 or 9 of treatment—and compared with the initial scan to rule out significant aortic expansion.

Once the patient’s condition has been stabilized, pharmacologic management is gradually shifted from IV to oral medications. Oral therapy, usually including a β-antagonist, is initiated when systolic pressure is consistently between 100 and 110 mmHg and the neurologic, renal, and cardiovascular systems are stable. Many patients can be discharged after their blood pressure is well controlled with oral agents and after serial CT scans confirm the absence of aortic expansion.

Long-term pharmacologic therapy is important for patients with chronic aortic dissection. β-Blockers remain the drugs of choice. In a 20-year follow-up study, DeBakey and colleagues found that inadequate blood pressure control was associated with late aneurysm formation. Aneurysms developed in only 17% of patients with “good” blood pressure control, compared with 45% of patients with “poor” control.

Aggressive imaging follow-up is recommended for all patients with chronic aortic dissection. Both contrast-enhanced CT and MRA scans provide excellent aortic imaging and facilitate serial comparisons to detect progressive aortic expansion. The first surveillance scan is obtained approximately 6 weeks after the onset of dissection. Subsequent scans are obtained at 3 to 6 months and then at 1 year after onset. If the aorta appears to be stable, imaging is obtained annually thereafter. Scans are obtained more frequently in high-risk patients, such as those with Marfan or Loeys-Dietz syndrome, and in those in whom significant aortic expansion is detected. For patients who have undergone graft repair of descending aortic dissection, annual CT or MRA scans are also obtained to detect false aneurysm formation or dilatation of unrepaird segments of aorta. Early detection of worrisome changes allows timely, elective intervention before rupture or other complications develop; rupture of the distal aorta is relatively common in patients with chronic aortic dissection and often results in death.

**Indications for Open Surgery** In the acute phase of descending aortic dissection, open surgery has been traditionally reserved for patients who experience complications. Complicated acute distal aortic dissections are those with aortic rupture, increasing periaortic or pleural fluid volume, rapidly expanding aortic diameter, uncontrolled hypertension, and persistent pain despite adequate medical therapy and malperfusion. In general terms, emergency open operations were originally intended to prevent or repair rupture and relieve life-threatening ischemic manifestations. However, open operation is associated with high morbidity in such cases; now that stent graft technology is available, endovascular surgical intervention is recommended for patients with complicated acute distal aortic dissection.

Acute dissection superimposed on a preexisting aneurysm is considered a life-threatening condition and is therefore another indication for operation. Finally, patients who have a history of noncompliance with medical therapy may ultimately benefit more from surgical intervention if they are otherwise reasonable operative candidates.

In the chronic phase, the indications for open surgical intervention for aortic dissection are similar to those for degenerative thoracic aortic aneurysm, although a slightly lower threshold of repair is now recommended. Guidelines for thoracic aortic disease recommend elective operation in otherwise healthy patients when the affected segment has reached a diameter of 5.5 cm, especially in patients with heritable disorders. Rapid aortic enlargement and other factors that increase the likelihood of aortic rupture may also be considered.

**Endovascular Treatment**

**Malperfusion Syndrome** Endovascular therapy is routinely used in patients with descending aortic dissection complicated by visceral malperfusion. Abdominal malperfusion syndrome often is fatal; prompt identification of visceral ischemia and expedited treatment to restore hepatic, gastrointestinal, and renal perfusion are imperative for a positive outcome. As described in a later section, several open surgical techniques can be used to reestablish blood flow to compromised organs. However, in acute cases, open surgery is associated with poor outcomes. Therefore, endovascular intervention is the preferred initial approach in such cases. In one endovascular technique known as endovascular fenestration, a balloon is used to create a tear in the dissection flap, which allows blood to flow in both the true and false lumens. Although endovascular fenestration was commonly used in the past, its use has declined in recent years as direct aortic and branch-vessel stenting techniques have evolved and gained favor. Placing a stent graft in the true lumen of the aorta can resolve a “dynamic” malperfusion. Occasionally, a small stent must be placed directly in the lumen of a visceral or renal artery because the dissection has propagated into the branch, resulting in “static” malperfusion at the origin.

Iliofemoral malperfusion causing limb-threatening leg ischemia also can be treated via an endovascular approach. Limb malperfusion usually resolves after the endovascular repair of acute descending thoracic aortic dissection. If the malperfusion does not resolve, then a femoral-to-femoral arterial bypass graft is an effective option.
Acute Dissection Although surgery has been traditionally recommended for patients with complicated acute descending aortic dissection, many centers have shifted toward using endovascular stent grafts as the preferred approach in these cases because of the high morbidity associated with the open operation. Evidence suggests that emergent endovascular repair in patients with true lumen collapse and complications such as rupture or dynamic malperfusion may be lifesaving in these difficult-to-treat patients. However, these patients remain at risk of further complication or future reintervention. Although endovascular repair in patients with heritable aortic disorders is generally not recommended, this technique can be used as a bridge to later, definitive repair in such life-threatening circumstances.66

Controversy exists regarding the use of endovascular stent grafts to treat uncomplicated acute descending dissection; some encouraging data have been published in the last couple of years.210 The goal of this treatment strategy is to use the stent graft to cover the intimal tear, seal the entry site of the dissection, and eventually cause thrombosis of the false lumen to aid in aortic remodeling and reduce late aortic expansion. Such procedures take place in a hybrid operating room. After the true lumen is accessed through the femoral arteries, an aortogram is taken, and the intimal tear is identified. Note that the diameter of the true lumen is measured on the preoperative contrast-enhanced CT scan. The use of IVUS is encouraged to help access the true lumen, verify navigation of the wire inside the true lumen, and confirm measurements. For these cases, a stent graft is selected with a diameter no more than 10% greater than that of the true lumen. Unlike stents deployed to treat most descending thoracic aortic aneurysms, stents deployed to treat descending thoracic aortic dissections must not be ballooned, because ballooning can cause a new intimal tear, retrograde dissection into the ascending aorta, or even aortic rupture. The ideal length of the descending thoracic aorta that should be covered in patients with acute distal descending dissection remains unclear. Close monitoring with serial imaging is necessary after endovascular repair because the false lumen remains at risk for retrograde perfusion or pressurization.

Chronic Dissection Endovascular treatment of chronic descending aortic dissection is supported by the 5-year data of the INSTEAD-XL trial, which showed that endovascular repair combined with optimal medical treatment was associated with slower disease progression and greater aorta-specific survival than optimal medical treatment alone.211 Importantly, patients in the trial had dissections in the early chronic phase, many within 10 to 12 weeks of onset and all within 1 year of onset. Endovascular repair of chronic dissection is particularly challenging because the relative rigidity of the dissecting membrane—which increases over time during the chronic phase—and the presence of multiple reentry sites make it difficult to exclude the false lumen.

Penetrating Aortic Ulcer Patients with PAUs appear to be good candidates for endovascular intervention. Covering the focal ulceration with a stent graft has been shown to be an effective treatment.212 In a recent study by Patel and colleagues,213 endovascular repair of PAU was associated with better early outcomes than open repair. However, when PAU was associated with adjacent hematoma within the aortic wall, rates of subsequent reintervention were increased.

Open Repair

Acute Dissection In patients with acute aortic dissection, open surgical repair of the descending thoracic or thoracoabdominal aorta has been traditionally associated with high morbidity and mortality.174 Therefore, surgery was generally only performed to prevent fatal rupture or to restore branch-vessel perfusion in patients with complicated dissection.208 With the evolution of endovascular technology, open repair has fallen out of favor in recent years.

Malperfusion Syndrome In patients with malperfusion, when an endovascular approach is unavailable or unsuccessful, open surgery is necessary. Lower-extremity ischemia can be readily addressed with surgical extra-anatomic revascularization techniques, such as femoral-to-femoral bypass grafting. In patients with abdominal organ ischemia, flow to the compromised bed must be reestablished swiftly. Although they are considered second-line therapies, multiple techniques are available, including graft replacement of the aorta (with flow redirected into the true lumen), open aortic fenestration, and visceral or renal artery bypass.

Chronic Dissection A more aggressive replacement usually is performed during elective aortic repairs in patients with chronic dissection. In many regards, the operative approach used in these patients is identical to that used for descending thoracic and thoracoabdominal aortic aneurysms, as described in the first half of this chapter (Fig. 22-24). One key difference is the need to excise as much dissecting membrane as possible to clearly identify the true and false lumens and to locate all important branch vessels. When the dissection extends into the visceral or renal arteries, the membrane can be fenestrated, or the false lumen can be obliterated with sutures or intraluminal stents. Asymmetric expansion of the false lumen can create wide separation of the renal arteries. This problem is addressed by reattaching the mobilized left renal artery to a separate opening in the graft or by performing a left renal artery bypass with a side graft. Wedges of dissecting membrane also are excised from the aorta adjacent to the proximal and distal anastomoses, which allows blood to flow through both true and false lumens. When placing the proximal clamp is not technically feasible, hypothermic circulatory arrest can be used to facilitate the proximal portion of the repair.

OUTCOMES

Improvements in anesthesia, surgical techniques, and perioperative care have led to substantial improvements in outcome after thoracic aortic aneurysm repair. When performed in specialized centers, these operations are associated with excellent survival rates and acceptable morbidity rates. The interpretation of outcomes data is complicated by site-specific variables, such as the number of years reported and whether data are taken from single-practice centers or from pooled, multicenter, or national registries, and by patient-specific variables, such as type of enrollment, urgency and extent of repair, concomitant procedures performed, and the presence of preexisting risk factors such as advanced age, previous cardiovascular repair, disease of any system or organ, or heritable conditions.

Repair of Proximal Aortic Aneurysms

Risks associated with the open repair of the proximal aorta vary by extent of repair and are greatest for repairs involving total arch replacement.71,214 All varieties of aortic root replacement have shown acceptable early mortality rates and few complications. Two groups with 20 and 27 years’ experience with composite valve graft replacement reported early mortality rates of 5.6% and 1.9%, respectively; the more recent repairs had better outcomes.215,216 Early mortality rates for stentless porcine
tissue root replacements are also low, ranging from 3.6% to 6.0%. Early mortality rates for contemporary valve-sparing approaches to aortic root replacement are quite low (1%–2%) in experienced centers. Late survival rates after valve-sparing root procedures range from 97% to 99% at 5 years and approach 94% at 10 years.

Repairs incorporating the ascending aorta and aortic arch have acceptable outcomes; risk increases with patient-specific factors such as severe atherosclerosis or as larger sections of the aortic arch are incorporated into the repair. A revised surgical strategy—such as the use of hypothermic circulatory arrest—is often needed to avoid clamping atherosclerotic sections in the “porcelain” aorta. In Zingone and colleagues’ series of 64 patients who underwent replacement of atherosclerotic ascending aorta, hypothermic circulatory arrest was used in 61 patients (95%). Even though these patients had substantial comorbidity and 83% underwent concomitant cardiac repairs, acceptable rates of early mortality (11%) and stroke (6%) were obtained. Other

---

**Figure 22-24.** Illustration of distal aortic repair of a chronic dissection. A. Thoracoabdominal incision. B. Extent II thoracoabdominal aortic aneurysm resulting from chronic aortic dissection. The patient has previously undergone composite valve graft replacement of the aortic root and ascending aorta. After left heart bypass is initiated, the proximal portion of the aneurysm is isolated by placing clamps on the left subclavian artery, between the left common carotid and left subclavian arteries, and across the middle descending thoracic aorta. C. The isolated segment of aorta is opened by using electrocautery. D. The dissecting membrane is excised, and bleeding intercostal arteries are oversewn. The aorta is prepared for proximal anastomosis by transecting it distal to the proximal clamp and separating this portion from the esophagus (not shown). E. The proximal anastomosis between the aorta and an appropriately sized polyester graft is completed with continuous polypropylene suture. F. After left heart bypass has been stopped and the distal aortic cannula has been removed, the proximal clamp is repositioned onto the graft, the other two clamps are removed, and the remainder of the aneurysm is opened. G. The rest of the dissecting membrane is excised, and the openings to the celiac, superior mesenteric, and renal arteries are identified. H. Selective visceral perfusion with oxygenated blood from the bypass circuit is delivered through balloon perfusion catheters placed in the celiac and superior mesenteric arterial ostia. Cold crystalloid is delivered to the renal arteries. The critical intercostal arteries are reattached to an opening cut in the graft. I. To minimize spinal cord ischemia, the proximal clamp is repositioned distal to the intercostal reattachment site. A second oval opening is fashioned in the graft adjacent to the visceral vessels. Selective perfusion of the visceral arteries continues during their reattachment to the graft. A separate anastomosis is often required to reattach the left renal artery. J. After the balloon perfusion catheters are removed and the visceral anastomosis is completed, the clamp is again moved distally, restoring blood flow to the celiac, renal, and superior mesenteric arteries. The final anastomosis is created between the graft and the distal aorta. (Reproduced with permission from Creager MA, Dzau VS, Loscalzo J: Vascular Medicine, 7th ed. Philadelphia, PA: Elsevier/Saunders; 2006.)
studies indicate that the enhanced risk of neurocognitive disturbances in ascending repairs using circulatory arrest are not offset by lower rates of early mortality. Regarding extended proximal repair, reported early mortality rates after traditional stage 1 elephant trunk repairs (primarily using island reattachment strategies) range from 2.3 to 13.9%.229-233

Contemporary mortality rates for extensive proximal aortic repair have improved as new strategies and modified adjuncts have been adopted. For example, by adopting contemporary approaches, we have reduced early mortality for stage 1 elephant trunk repairs from 12% to 2% in our patients.86,230 Similarly, in a report by Kazui and colleagues234 covering 20 years of experience and 472 consecutive patients who underwent aortic arch repair with selective ACP, operative mortality was 16.0% for early repairs and 4.1% for more recent repairs. Other contemporary reports of the use of techniques such as moderate hypothermia and Y-graft approaches235-238 indicate similarly improved outcomes; early mortality ranges from 1% to 7%, stroke rates range from 1% to 6%, and no cases of paraplegia are reported. Although paraplegia has traditionally been an unusual and infrequent complication of aortic arch repair, it has been reported as a complication of “long” elephant trunk approaches and frozen elephant trunk approaches.239

Because of the heterogeneity of hybrid arch approaches and the tendency to use these approaches in high-risk patients, results of hybrid arch repair are difficult to interpret. In a recent report from our group,49 among 319 consecutive patients who underwent total arch replacement in the last 8.5 years, 274 patients had traditional open repair and 45 patients had hybrid zone 0 exclusion repairs. The rate of permanent adverse outcome (death, persistent neurologic deficit at discharge, or persistent hemodialysis at discharge) was not significantly different between the two groups. A higher overall stroke rate was noticed in the hybrid group, reinforcing the importance of catheter skills and careful wire manipulation. A meta-analysis conducted by Koulliias and Wheatley241 of data from 15 studies with 463 patients found an average 30-day mortality rate of 8.3%; stroke, 4.4%; paraplegia, 3.9%; and endoleak, 9.2%. Of note, relatively few repairs (30%) were performed “off-pump,” and the majority of repairs used cardiopulmonary bypass or hypothermic circulatory arrest. Additionally, several reports of small series have documented increased risk of acute retrograde aortic dissection during hybrid arch repairs; rates range from 0% to 7.5%, and these patients face significant mortality risk (ranging from 33% to 100%) should this occur.
Treatment of Acute Ascending Aortic Dissection
The International Registry of Acute Aortic Dissection (IRAD) provides the most comprehensive data on contemporary outcomes in patients with acute aortic dissection. This registry was established in 1996 and has accumulated data from >7000 patients treated for acute aortic dissection at 51 centers in 12 countries. An IRAD analysis of data from 776 patients who underwent surgical repair of acute ascending aortic dissection revealed an in-hospital mortality rate of 23.8%. The investigators identified several preoperative predictors of early mortality, including age >70 years, previous cardiac surgery, hypotension or shock at presentation, abrupt onset of symptoms, migrating pain, cardiac tamponade, preoperative renal failure, pulse deficit, and evidence of myocardial ischemia or infarction on ECG. In a report from IRAD, in-hospital mortality after surgical treatment had decreased from 25% in 1995 to 18% in 2013. The German Registry for Acute Aortic Dissection (GERAADA) has collected data on more than 3300 patients from 56 centers since 2006. In a report of 1436 patients with acute proximal dissection that was surgically repaired using hypothermic circulatory arrest with or without unilateral and bilateral ACP, the early mortality rates ranged from 13.9% to 19.4%; the 628 patients with unilateral ACP had the lowest rate of early death. Operative mortality reported by North American centers varies from 5% to 17%; improvements in outcomes may be related to the implementation of protocol-based management and the assembly of thoracic aortic teams.

Repair of Distal Aortic Aneurysms
Endovascular Repair of Descending Thoracic Aortic Aneurysms. In the earliest series of endovascular repairs of descending thoracic aortic aneurysms, mortality and morbidity were difficult to assess. Most of the reported series were small and included a large proportion of high-risk patients with substantial comorbidity. Subsequent evidence from pivotal, nonrandomized trials that compared patients who underwent endograft exclusion with historical or concurrent patients who underwent open repair showed that the stent graft groups had significantly less morbidity and early mortality than the open repair groups, although in two of the trials, a nonsignificant between-group difference was observed in the rate of stroke. Five-year comparative data show that the two groups differed significantly in their aneurysm-related mortality rates (2.8% for endovascular patients and 11.7% for open repair patients) but not in their rates of all-cause mortality (which were 32% and 31%, respectively). Additional pivotal trial 5-year outcomes indicate the growing disparity between aneurysm-related (96.1%) and all-cause survival (58.5%) in patients with endovascular repair, leading some to comment on the possible futility of repair in many patients. Among 8967 patients identified in the National Inpatient Sample database (8255 with open repair and 712 with endovascular repair), the odds of death were 46% lower among patients who underwent endovascular repair rather than open repair. The endovascular repair group also had lower odds of postoperative neurologic, cardiac, and respiratory complications.

Open Repair of Descending Thoracic and Thoracoabdominal Aortic Aneurysms. Contemporary results of open repairs of descending thoracic aortic aneurysms, including those performed in select patients with chronic dissection, indicate that early mortality rates range from 4.1% to 8.0%, renal failure rates range from 4.2% to 7.5%, and paraplegia rates range from 2.3% to 5.7%; stroke rates are generally lower, ranging from 1.8% to 2.1%. In our series, although the risk of paraplegia increased with the extent of repair, the risk of mortality was greatest for those undergoing repair of the proximal two thirds of the descending aorta. As expected, stroke rates after distal aortic repairs were highest when the clamp site was near the left subclavian artery.

Contemporary series of open thoracoabdominal aortic repairs show acceptable survival. Reported outcome rates range from 5% to 12% for early mortality, 3.8% to 9.5% for paraplegia, 1.7% to 5.2% for stroke, and 6% to 12% for renal complications. Many of these series summarize 10 to 20 years of surgical experience, although some present a shorter but more contemporary experience. Even for complex thoracoabdominal aortic aneurysms, mortality and morbidity after thoracoabdominal aortic aneurysm repair, reported that patients treated at low-volume centers fared less well. Replacing the entire thoracoabdominal aorta (i.e., performing an extent II repair) carries the highest risk of death, bleeding, renal failure, and paraplegia. Early survival has been estimated at 79% at 2 years, and mid-term survival has been estimated at 63% at 5 years. In our recent report of 3309 repairs, the overall mortality rate was 7.5%, and the rate of operative death was higher in extent II and III repairs than in extent I and IV. Permanent paraplegia and paraparesis occurred in 2.9% and 2.4% of patients, respectively, and the incidence of paraplegia in patients 50 years of age or younger was only 1.1%. Estimated survival after repair was 84% ± 1% at 1 year, 64% ± 1% at 5 years, 37% ± 1% at 10 years, and 18% ± 1% at 15 years.

Treatment of Descending Thoracic Aortic Dissection
Nonoperative Management. The in-hospital mortality rate is 8.7% for patients with acute descending aortic dissection who receive nonoperative treatment; however, when IRAD stratified patients according to clinical presentation, the mortality rate for patients with uncomplicated dissection was less than 4%, whereas the mortality rate for patients with complicated dissection was more than 20%. The primary causes of death during nonoperative management are rupture, malperfusion, and cardiac failure. Risk factors associated with treatment failure—defined as death or need for surgery—include an enlarged aorta, persistent hypertension despite maximal treatment, oliguria, and peripheral ischemia. Among patients who receive nonoperative treatment for descending aortic dissection and who survive the acute period, approximately 90% remain alive 1 year later, and approximately 76% are alive 3 years later.

Endovascular Treatment. For patients with complicated acute descending thoracic aortic dissection, including rupture and malperfusion of the visceral or renal arteries, an endovascular approach is ideal. The Stanford group reported a 93% technical success rate for endovascular reperfusion of an ischemic bed. Their experience with the use of first-generation stent grafts to treat acute complicated descending dissections...
was also encouraging: Complete thrombosis of the false lumen occurred in 79% of patients. The early mortality rate was 16%, comparable to that associated with open techniques.\textsuperscript{273} A meta-analysis of observational studies of endovascular repair, which included 248 patients with acute descending thoracic aortic dissection, found a 30-day mortality rate of 9.8%.\textsuperscript{274} When compared with early mortality rates obtained from IRAD data,\textsuperscript{174} this rate is substantially lower than the rate associated with open surgical treatment and is similar to the rate achieved with nonoperative management. However, patients with complicated acute descending dissection remain susceptible to late events; at 1 year, survival is approximately 70%, and reintervention is needed in about 10% of survivors.\textsuperscript{275}

The ADSORB trial\textsuperscript{276} focused on patients with uncomplicated acute descending thoracic aortic dissection. Patients were randomly assigned to optimal medical therapy alone (n = 31) or endovascular repair plus optimal medical therapy (n = 30).\textsuperscript{277} The 1-year results showed aortic remodeling with false lumen thrombosis and reduced diameter in the group treated with endovascular repair.

The INSTED-XL trial involved 140 patients with stable, early-chronic descending thoracic aortic dissection who were randomly assigned to either endovascular repair plus optimal medical treatment or optimal medical therapy alone.\textsuperscript{211} The eagerly anticipated 5-year data showed that endovascular repair was associated with greater survival and slower disease progression.

**CONCLUSIONS**

Aortic aneurysm may present as localized or extensive disease. The availability and development of adjuncts and endovascular techniques have supported the constant evolution of surgical strategies to tackle these complex problems. Repair strategies range from isolated, totally endovascular aortic repair for descending thoracic aneurysms to extensive total aortic and staged replacements with a combination of both open and endovascular techniques. Regardless of the difficulty of accurately assessing the risks associated with aortic repair, surgical repair of the thoracoabdominal aorta clearly remains the most challenging aortic repair in terms of mortality and morbidity. Accordingly, replacing the entire thoracoabdominal aorta (i.e., performing an extent II repair) carries the highest risk of death, renal failure, and paraplegia.\textsuperscript{69,70,261,263,269}

**ACKNOWLEDGMENTS**

The authors wish to thank Susan Y. Green, MPH, and Stephen N. Palmer, PhD, ELS, for editorial assistance; Scott A. Weldon, MA, CMI, and Carol P. Larson, CMI, for creating the illustrations; and Kapil Sharma, MD, and Raja Gopaldas, MD, for their substantial contributions to the chapters published in the 9th and 10th editions of this book, on which this updated chapter is based.

**REFERENCES**

Entries highlighted in bright blue are key references.


# Arterial Disease

Peter H. Lin, Carlos F. Bechara, Changyi Chen, and Frank J. Veith

## General Approach to the Vascular Patient
- **898**
  - The Vascular History
  - The Vascular Physical Examination
  - Noninvasive Diagnostic Evaluation of the Vascular Patient
  - Radiologic Evaluation of the Vascular Patient
  - Preoperative Cardiac Evaluation

## Basic Principles of Endovascular Therapy
- **903**
  - Needles and Access
  - Guidewires
  - Hemostatic Sheaths
  - Catheters
  - Angioplasty Balloons
  - Stents
  - Stent Grafts

## Carotid Artery Disease
- **907**
  - Epidemiology and Etiology of Carotid Occlusive Disease
  - Clinical Manifestations of Cerebral Ischemia
  - Diagnostic Evaluation
  - Treatment of Carotid Occlusive Disease
  - Carotid Endarterectomy Versus Angioplasty and Stenting

## Renal Artery Disease
- **935**
  - Etiology
  - Clinical Manifestations
  - Diagnostic Evaluation
  - Treatment Indications
  - Surgical Reconstruction
  - Clinical Results of Surgical Repair
  - Endovascular Treatment

## Aortoiliac Occlusive Disease
- **940**
  - Diagnostic Evaluation
  - Differential Diagnosis
  - Collateral Arterial Network
  - Disease Classification
  - General Considerations
  - Surgical Treatment for Aortic Disease
  - Endovascular Treatment for Iliac Artery Disease
  - Complications of Endovascular Aortoiliac Interventions
  - Clinical Results Comparing Surgical and Endovascular Treatment of Aortoiliac Disease

## Lower Extremity Occlusive Disease
- **950**
  - Epidemiology
  - Diagnostic Evaluation
  - Differential Diagnosis

## Nonatherosclerotic Disorders of Blood Vessels
- **969**
  - Giant Cell Arteritis (Temporal Arteritis)
  - Takayasu’s Arteritis
  - Ehlers-Danlos Syndrome
  - Marfan’s Syndrome
  - Pseudoxanthoma Elasticum
  - Kawasaki’s Disease
  - Inflammatory Arteritis and Vasculitis
  - Behçet’s Disease
  - Polyarteritis Nodosa
  - Radiation-Induced Arteritis
  - Raynaud’s Syndrome
  - Fibromuscular Dysplasia
  - Nonatherosclerotic Disease Affecting the Popliteal Artery
  - Buerger’s Disease (Thromboangiitis Obliterans)


Key Points

1. Carotid intervention as a preventive strategy should be performed in patients with 60% or greater symptomatic internal carotid artery stenosis and those with 80% or greater asymptomatic internal carotid artery stenosis. Carotid intervention for asymptomatic stenosis between 60% and 79% remains controversial. The modality of carotid intervention—carotid endarterectomy versus carotid stenting—remains controversial; currently, carotid endarterectomy appears to be associated with lower stroke rate with long term durability, whereas carotid stenting is more suitable under certain challenging anatomic or physiologic conditions.

2. Abdominal aortic aneurysms should be repaired when the risk of rupture, determined mainly by aneurysm size, exceeds the risk of death due to perioperative complications or concurrent illness. Endovascular repair with less perioperative morbidity and mortality compared with open reconstruction is preferred in patients with suitable anatomic morphology for stent-graft placement.

3. Treatment objectives for symptomatic mesenteric ischemia are to improve quality of life and prevent bowel infarction. Endovascular intervention with stenting has similar treatment efficacy comparative with less perioperative morbidity compared to open mesenteric bypass. Surgical reconstruction has a proven durability and patency rate compared with endovascular intervention.

4. Aortoiliac occlusive disease can be treated with either endovascular means or open reconstruction, depending on patient risk stratification, occlusion characteristics, and symptomatology.

5. Claudication is a marker of extensive atherosclerosis and is mainly managed with risk factor modification and pharmacotherapy. Only 5% of patients with claudication will need intervention because of disabling extremity pain. The 5-year mortality of a patient with claudication approaches 30%. Patients with rest pain or tissue loss need expedient evaluation and vascular reconstruction to ameliorate the severe extremity pain and prevent limb loss. Endovascular intervention is preferred as the first line of therapy for lower extremity occlusive disease, whereas bypass reconstruction should be considered in failed endovascular therapy or long segment femoropopliteal occlusive disease.

GENERAL APPROACH TO THE VASCULAR PATIENT

Since the vascular system involves every organ system in our body, the symptoms of vascular disease are as varied as those encountered in any medical specialty. Lack of adequate blood supply to target organs typically presents with pain, for example, calf pain with lower extremity claudication, postprandial abdominal pain from mesenteric ischemia, and arm pain with axillo-subclavian arterial occlusion. In contrast, stroke and transient ischemic attack (TIA) are the presenting symptoms from middle cerebral embolization as a consequence of a stenosed internal carotid artery. The pain syndrome of arterial disease is usually divided clinically into acute and chronic types, with all shades of severity between the two extremes. Sudden onset of pain can indicate complete occlusion of a critical vessel, leading to more severe pain and critical ischemia in the target organ, resulting in lower limb gangrene or intestinal infarction. Chronic pain results from a slower, more progressive atherosclerotic occlusion, which can be totally or partially compensated by developing collateral vessels. Acute on chronic is another pain pattern in which a patient most likely has an underlying arterial stenosis that suddenly occludes, for example, the patient with a history of calf claudication who now presents with sudden, severe acute limb-threatening ischemia. The clinician should always try to understand and relate the clinical manifestations to the underlying pathologic process.

The Vascular History

Appropriate history should be focused based on the presenting symptoms related to the vascular system (Table 23-1). Of particular importance in the previous medical history is noting prior vascular interventions (endovascular or open surgical), and all vascular patients should have inquiry made about their prior cardiac history and current cardiac symptoms. Approximately 30% of vascular patients will be diabetic. A history of prior and current smoking status should be noted.

The patient with carotid disease in most cases is completely asymptomatic, having been referred based on the finding of a cervical bruit or duplex finding of stenosis. Symptoms of carotid territory TIs include transient monocular blindness (amaurosis), contralateral weakness or numbness, and dysphasia. Symptoms persisting longer than 24 hours constitute a stroke. In contrast, the patient with chronic mesenteric ischemia is likely to present with postprandial abdominal pain and weight loss. The patient fears eating because of the pain, avoids food, and loses weight. It is very unlikely that a patient with abdominal pain who has not lost weight has chronic mesenteric ischemia.

The patient with lower extremity pain on ambulation has intermittent claudication that occurs in certain muscle groups; for example, calf pain upon exercise usually reflects superficial femoral artery disease, while pain in the buttocks reflects iliac disease. In most cases, the pain manifests in one muscle group below the level of the affected artery, occurs only with exercise, and is relieved with rest only to recur at the same location, hence the term “window gazer’s disease.” Rest pain (a manifestation of severe underlying occlusive disease) is constant and occurs in the foot (not the muscle groups), typically at the metatarsophalangeal junction, and is relieved by dependency. Often the

Table 23-1

<table>
<thead>
<tr>
<th>Pertinent elements in vascular history</th>
</tr>
</thead>
<tbody>
<tr>
<td>• History of stroke or transient ischemic attack</td>
</tr>
<tr>
<td>• History of coronary artery disease, including previous myocardial infarction and angina</td>
</tr>
<tr>
<td>• History of peripheral arterial disease</td>
</tr>
<tr>
<td>• History of diabetes</td>
</tr>
<tr>
<td>• History of hypertension</td>
</tr>
<tr>
<td>• History of tobacco use</td>
</tr>
<tr>
<td>• History of hyperlipidemia</td>
</tr>
</tbody>
</table>
patient is prompted to sleep with their foot hanging off one side of the bed to increase the hydrostatic pressure.

The Vascular Physical Examination

Specific vascular examination should include abdominal aortic palpation, carotid artery examination, and pulse examination of the lower extremity (femoral, popliteal, posterior tibial, and dorsalis pedis arteries). The abdomen should be palpated for an abdominal aortic aneurysm, detected as an expansile pulse above the level of the umbilicus. It should also be examined for the presence of bruits. Because the aorta typically divides at the level of the umbilicus, an aortic aneurysm is most frequently palpable in the epigastrium. In thin individuals, a normal aortic pulsation is palpable, while in obese patients, even large aortic aneurysms may not be detectable. Suspicion of a clinically enlarged aorta should lead to the performance of an ultrasound scan for a more accurate definition of aortic diameter.

The carotids should be auscultated for the presence of bruits, although there is a higher correlation with coronary artery disease than underlying carotid stenosis. A bruit at the angle of the mandible is a significant finding, leading to follow-up duplex scanning. The differential diagnosis is a transmitted murmur from a sclerotic or stenotic aortic valve. The carotid is palpable deep to the sternocleidomastoid muscle in the neck. Palpation, however, should be gentle and rarely yields clinically useful information.

Upper extremity examination is necessary when an arteriovenous graft is to be inserted in patients who have symptoms of arm pain with exercise. Thoracic outlet syndrome (TOS) can result in occlusion or aneurysm formation of the subclavian artery. Distal embolization is a manifestation of TOS; consequently, the fingers should be examined for signs of ischemia and ulceration. The axillary artery enters the limb below the middle of the clavicle, where it can be palpated in thin patients. It is usually easily palpable in the axilla and medial upper arm. The brachial artery is most easily located at the antecubital fossa immediately medial to the biceps tendon. The radial artery is palpable at the wrist anterior to the radius.

For lower extremity vascular examination, the femoral pulse is usually palpable midway between the anterior superior iliac spine and the pubic tubercle. The popliteal artery is palpated in the popliteal fossa with the knee flexed to 45° and the foot supported on the examination table to relax the calf muscles. Palpation of the popliteal artery is a bimanual technique. Both thumbs are placed on the tibial tuberosity anteriorly, and the fingers are placed into the popliteal fossa between the two heads of the gastrocnemius muscle. The popliteal artery is palpated by compressing it against the posterior aspect of the tibia just below the knee. The posterior tibial pulse is detected by palpation 2 cm posterior to the medial malleolus. The dorsalis pedis is detected 1 cm lateral to the hallucis longus extensor tendon, which dorsiflexes the great toe and is clearly visible on the dorsum of the foot. Pulses can be graded using either the traditional four-point scale or the basic two-point scale system (Table 23-2). The foot should also be carefully examined for pallor on elevation and rubor on dependency, as these findings are indicative of chronic ischemia. Note should also be made of nail changes and loss of hair. Ulceration and other findings specific to disease states are described in relevant sections later in this chapter.

After reconstructive vascular surgery, the graft may be available for examination, depending on its type and course.

Table 23-2
Grading scales for peripheral pulses

<table>
<thead>
<tr>
<th>TRADITIONAL SCALE</th>
<th>BASIC SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>4+</td>
<td>Normal</td>
</tr>
<tr>
<td>3+</td>
<td>Slightly reduced</td>
</tr>
<tr>
<td>2+</td>
<td>Markedly reduced</td>
</tr>
<tr>
<td>1+</td>
<td>Barely palpable</td>
</tr>
<tr>
<td>0 Absent</td>
<td>Absent</td>
</tr>
</tbody>
</table>

The in situ lower extremity graft runs in the subcutaneous fat and can be palpated along most of its length. A change in pulse quality, aneurysmal enlargement, or a new bruit should be carefully noted. Axillofemoral grafts, femoral-to-femoral grafts, and arteriovenous access grafts can usually be easily palpated as well.

Noninvasive Diagnostic Evaluation of the Vascular Patient

Ankle-Brachial Index. There is increasing interest in the use of the ankle-brachial index (ABI) to evaluate patients at risk for cardiovascular events. An ABI less than 0.9 correlates with increased risk of myocardial infarction and indicates significant, although perhaps asymptomatic, underlying peripheral vascular disease. The ABI is determined in the following ways. Blood pressure is measured in both upper extremities using the highest systolic blood pressure as the denominator for the ABI. The ankle pressure is determined by placing a blood pressure cuff above the ankle and measuring the return to flow of the posterior tibial and dorsalis pedis arteries using a pencil Doppler probe over each artery. The ratio of the systolic pressure in each vessel divided by the highest arm systolic pressure can be used to express the ABI in both the posterior tibial and dorsalis pedis arteries (Fig. 23-1). Normal is more than 1. Patients with claudication typically have an ABI in the 0.5 to 0.7 range, and those with rest pain are in the 0.3 to 0.5 range. Those with gangrene have an ABI of less than 0.3. These ranges can vary depending on the degree of compressibility of the vessel. The test is less reliable in patients with heavily calcified vessels. Due to noncompressibility, some patients, such as diabetics and those with end-stage renal disease, may have ABI ≥1.40 and require additional noninvasive diagnostic testing to evaluate for peripheral artery disease. Alternative tests include toe-brachial pressures, pulse volume recordings, transcutaneous oxygen measurements, or vascular imaging (duplex ultrasound).

Segmental Limb Pressures. By placing serial blood pressure cuffs down the lower extremity and then measuring the pressure with a Doppler probe as flow returns to the artery below the cuff, it is possible to determine segmental pressures down the leg. This data can then be used to infer the level of the occlusion. The systolic pressure at each level is expressed as a ratio, with the highest systolic pressure in the upper extremities as the denominator. Normal segmental pressures commonly show high thigh pressures 20 mmHg or greater in comparison to the brachial artery pressures. The low thigh pressure should be equivalent to brachial pressures. Subsequent pressures should fall by no more than 10 mmHg at each level. A pressure gradient of 20 mmHg between two subsequent levels is usually indicative of occlusive disease at that level. The most frequently used index is the ratio of the ankle pressure to the brachial pressure,
Underlying arterial occlusive disease. Cuffs placed at different levels on the leg detect changes in blood volume and produce a pulse volume recording (PVR) when connected to a plethysmograph (Fig. 23-2). To obtain accurate PVR waveforms, the cuff is inflated to 60 to 65 mmHg, so as to detect volume changes without causing arterial occlusion. Pulse volume tracings are suggestive of proximal disease if the upstroke of the pulse is not brisk, the peak of the wave tracing is rounded, and there is disappearance of the dicrotic notch.

Although isolated segmental limb pressures and PVR measurements are 85% accurate when compared with angiography in detecting and localizing significant atherosclerotic lesions, when used in combination, accuracy approaches 95%. For this reason, it is suggested that these two diagnostic modalities be used in combination when evaluating peripheral artery disease.

Radiologic Evaluation of the Vascular Patient

Ultrasound. Ultrasound examinations are relatively time consuming, require experienced technicians, and may not visualize all arterial segments. Doppler waveform analysis can suggest atherosclerotic occlusive disease if the waveforms in the insonated arteries are biphasic, monophasic, or asymmetrical. B-mode ultrasonography provides black and white, real-time images. B-mode ultrasonography does not evaluate blood flow; thus, it cannot differentiate between fresh thrombus and flowing blood, which have the same echogenicity. Calcification in atherosclerotic plaques will cause acoustic shadowing. B-mode ultrasound probes cannot be sterilized. Use of the B-mode probe intraoperatively requires a sterile covering and gel to maintain an acoustic interface. Experience is needed to obtain and interpret images accurately. Duplex ultrasonography entails performance of B-mode imaging, spectral Doppler scanning, and color-flow duplex scanning. The caveat to performance of duplex ultrasonography is meticulous technique by a certified vascular ultrasound technician, so that the appropriate 60° Doppler angle is maintained during insonation with the ultrasound probe. Alteration of this angle can markedly alter waveform appearance and subsequent interpretation of velocity measurements. Direct imaging of intra-abdominal vessels with duplex ultrasound is less reliable because of the difficulty in visualizing the vessels through overlying bowel. These disadvantages currently limit the applicability of duplex scanning in the evaluation of aortoiliac and infrapopliteal disease. A clinical study has shown that duplex ultrasonography had lower sensitivity in the calculation of infrapopliteal vessel stenosis in comparison to conventional digital subtraction or computed tomography angiography. Few surgeons rely solely on duplex ultrasonography for preoperative planning in lower extremity revascularizations; but with experience, lower extremity arteries can be insonated to determine anatomy, and the functional significance of lesions can be determined by calculation of degree of stenosis from velocity ratios. Duplex scanning is unable to evaluate recently implanted polytetrafluoroethylene (PTFE) and polyester (Dacron) grafts because they contain air, which prevents ultrasound penetration.

Computed Tomography Angiography. Computed tomography angiography (CTA) is a noninvasive, contrast-dependent method for imaging the arterial system. It depends on intravenous infusion of iodine-based contrast agents. The patient is advanced through a rotating gantry, which images serial transverse slices. The contrast-filled vessels can be extracted from the slices and rendered in three-dimensional format (Fig. 23-3).
Figure 23-2. Typical report of peripheral vascular study with arterial segmental pressure measurement plus Doppler evaluation of the lower extremity.

The extracted images can also be rotated and viewed from several different directions during postacquisition image processing. This technology has been advanced as a consequence of aortic endografting. CTA provides images for postprocessing that can be used to display the aneurysm in a format that demonstrates thrombus, calcium, lumen, and the outer wall, and allows “fitting” of a proposed endograft into the aneurysm (Fig. 23-4). CTA is increasingly being used to image the carotid bifurcation, and as computing power increases, the speed of image acquisition and resolution will continue to increase. The major limitations of multidetector CTA are use of contrast and presence of artifacts caused by calcification and stents. CTA can

Figure 23-3. A multidetector computed tomography angiography with three-dimensional reconstruction of the iliofemoral arterial circulation in two patients with lower leg claudication. A. A 50-year-old man with an occluded right superficial femoral artery (single long arrow) with reconstituted superficial femoral artery at the level of mid-thigh. Arterial calcifications (single short arrows) are present in the bilateral distal superficial femoral arteries. B. A 53-year-old man with occluded right common iliac artery (double arrows).
overestimate the degree of in-stent stenosis, while heavy calcification can limit the diagnostic accuracy of the method by causing a “blooming artifact.” The artifacts can be overcome with alteration in image acquisition technique. There are no randomized trials to document the superiority of multidetector CTA over traditional angiography, but there is emerging evidence to support the claim that multidetector CTA has sensitivity, specificity, and accuracy that rival invasive angiography.

**Magnetic Resonance Angiography.** Magnetic resonance angiography (MRA) has the advantage of not requiring iodinated contrast agents to provide vessel opacification (Fig. 23-5). Gadolinium is used as a contrast agent for MRA studies, and because it is generally not nephrotoxic, it can be used in patients with elevated creatinine. MRA is contraindicated in patients with pacemakers, defibrillators, spinal cord stimulators, intracerebral shunts, cochlear implants, and cranial clips. Patients with claustrophobia may require sedation to be able to complete the test. The presence of metallic stents causes artifacts and signal drop-out; however, these can be dealt with using alternations in image acquisition and processing. Nitinol stents produce minimal artifact. Compared to other modalities, MRA is relatively slow and expensive. However, due to its noninvasive nature and decreased nephrotoxicity, MRA is being used more frequently for imaging vasculature in various anatomic distributions.

**Diagnostic Angiography.** Diagnostic angiography is considered the gold standard in vascular imaging. In many centers, its use is rapidly decreasing due to the development of noninvasive imaging modalities such as duplex arterial mapping, CTA, and MRA. Nevertheless, contrast angiography still remains in widespread use. The essential aspects of angiography are vascular access and catheter placement in the vascular bed that requires examination. The imaging system and the contrast agent are used to opacify the target vessel. Although in the past this function has largely been delegated to the interventional radiology service, an increasing number of surgeons are performing this procedure and following the diagnostic imaging with immediate surgical or endovascular intervention. There are several considerations when relying on angiography for imaging.

Approximately 70% of atherosclerotic plaques occur in an eccentric location within the blood vessel; therefore, images can be misleading when trying to evaluate stenoses because angiography is limited to a uniplanar “lumenogram.” With increased use of intravascular stent deployment, it has also been noted that assessment of stent apposition and stent position in relation to surrounding branches may be inaccurate. Furthermore, angiography exposes the patient to the risks of both ionizing radiation and intravascular contrast. Nevertheless, contrast angiography remains the most common invasive method of vascular investigation for both diagnostic and therapeutic intervention. The angiogram usually provides the final information needed to decide whether or not to proceed with operation or endovascular interventions.

Digital subtraction angiography (DSA) offers some advantages over conventional cut-film angiography such as excellent visualization despite use of lower volumes of contrast media. In particular, when multilevel occlusive lesions limit the amount of contrast reaching distal vessels, supplemental use of digital subtraction angiographic techniques may enhance visualization and definition of anatomy. Intra-arterial DSA uses a portable, axially
rotatable imaging device that can obtain views from different angles. DSA also allows for real-time video replay (Fig. 23-6). An entire extremity can be filmed with DSA using repeated injections of small amounts of contrast agent to obtain sequential angiographic images, the so-called pulse-chase technique.

**Preoperative Cardiac Evaluation**

The most important and most controversial aspect of preoperative evaluation in patients with atherosclerotic disease requiring surgical intervention is the detection and subsequent management of associated coronary artery disease. Several studies have documented the existence of significant coronary artery disease in 40% to 50% or more of patients requiring peripheral vascular reconstructive procedures, 10% to 20% of whom may be relatively asymptomatic largely because of their inability to exercise. Myocardial infarction is responsible for the majority of both early and late postoperative deaths. Most available screening methods lack sensitivity and specificity to predict postoperative cardiac complications. There have been conflicting reports regarding the utility of preoperative dipyridamole-thallium nuclear imaging or dobutamine-echocardiography to stratify vascular patients in terms of perioperative cardiac morbidity and mortality. In nearly half of patients, thallium imaging proves to be unnecessary because cardiac risk can be predicted by clinical information alone. Even with coronary angiography, it is difficult to relate anatomic findings to functional significance and, hence, surgical risk. There are no data confirming that percutaneous coronary interventions or surgical revascularization prior to vascular surgical procedures impact mortality or incidence of myocardial infarctions. In fact, coronary angiography is associated with its own inherent risks, and patients undergoing coronary artery bypass grafting or coronary percutaneous transluminal angioplasty (PTA) before needed aortoiliac reconstructions are subjected to the risks and complications of both procedures.

The Coronary Artery Revascularization Prophylaxis (CARP) trial showed that coronary revascularization in patients with peripheral vascular disease and significant coronary artery disease, who are considered high risk for perioperative complications, did not reduce overall mortality or perioperative myocardial infarction. Additionally, patients who underwent prophylactic coronary revascularization had significant delays prior to undergoing their vascular procedure and increased limb morbidity compared to patients who did not. Studies do support improvement in cardiovascular and overall prognosis with medical optimization of patients. Therefore, use of perioperative β-blockade, as well as use of antiplatelet medication, statins, and angiotensin-converting enzyme inhibitors, is encouraged in vascular patients.

![Figure 23-6. Digital subtraction angiography (DSA) provides excellent visualization of intravascular circulation with intra-arterial contrast administration. As depicted in this DSA study, multilevel lesions are demonstrated, which include a focal left iliac artery stenosis (large arrow), right superficial femoral occlusion (curved arrows), left superficial femoral stenosis (small arrow), and multiple tibial artery stenoses (arrowheads).](image-url)
in a manner that would enable an accurate comparison with the more traditional methods of open surgical intervention. Long-term follow-up for these procedures is frequently lacking; however, because of the potential to treat patients with decreased mortality and morbidity, endovascular skills and techniques are being adopted into mainstream vascular surgery.

**Needles and Access**

Needles are used to achieve percutaneous vascular access. The size of the needle will be dictated by the diameter of the guidewire used. Most often, an 18-gauge needle is used, as it will accept a 0.035-inch guidewire. A 21-gauge micropuncture needle will accept a 0.018-inch guidewire. The most popular access needle is the Seldinger needle, which can be used for single- and double-wall puncture techniques.

Femoral arterial puncture is the most common site for access. The common femoral artery (CFA) is punctured over the medial third of the femoral head, which is landmarked using fluoroscopy. The single-wall puncture technique requires a sharp, beveled needle tip and no central stylet. The anterior wall of the vessel is punctured with the bevel of the needle pointing up, and pulsatile back-bleeding indicates an intraluminal position. This method is most useful for graft punctures, patients with abnormal clotting profiles, or if thrombolytic therapy is anticipated. Once the needle assumes an intraluminal position, verified by pulsatile back-bleeding, the guidewire may be advanced. This is always passed gently and under fluoroscopic guidance to avoid subintimal dissection or plaque disruption. Double-wall puncture techniques are performed with a blunt needle that has a removable inner cannula. The introducer needle punctures both walls of the artery and is withdrawn until bleeding is obtained to confirm intraluminal position prior to advancing a guidewire. There can be troublesome bleeding from the posterior arterial wall puncture; therefore, single puncture techniques are preferred.

Retrograde femoral access is the most common arterial access technique (Fig. 23-7). The advantages of this technique include the size and fixed position of the CFA, as well as the relative ease of compression against the femoral head at the end of the procedure. Care should be taken to avoid puncturing the external iliac artery above the inguinal ligament because this can result in retroperitoneal hemorrhage secondary to ineffective compression of the puncture site. Likewise, puncturing too low, at or below the CFA bifurcation, can result in thrombosis or pseudoaneurysm formation of the superficial femoral artery (SFA) or profunda femoris artery (PFA). Antegrade femoral access is more difficult than retrograde femoral access, and there is a greater tendency to puncture the SFA, but it is invaluable when the aortic bifurcation cannot be traversed or when devices are not long enough to reach a lesion from a contralateral femoral access approach. Occasionally, when the distal aorta or bilateral iliac arteries are inaccessible because of the extent of atherosclerotic lesions, scarring, or presence of bypass conduits, the brachial artery must be used to obtain access for diagnostic and therapeutic interventions. The left brachial artery is punctured because this avoids the origin of the carotid artery and thus decreases the risk of catheter-related emboli to the brain. The artery is accessed with a micropuncture needle just proximal to the antecubital crease. The use of brachial access is associated with a higher risk of thrombosis and nerve injuries than femoral access.

**Guidewires**

Guidewires are used to introduce, position, and exchange catheters. A guidewire generally has a flexible and stiff end. In general, only the flexible end of the guidewire is placed in the vessel. All guidewires are composed of a stiff inner core and an outer tightly coiled spring that allows a catheter to track over the guidewire. There are five essential characteristics of guidewires: size, length, stiffness, coating, and tip configuration.

Guidewires come in different maximum transverse diameters, ranging from 0.011 to 0.038 inches. For most aortoiliac procedures, a 0.035-inch wire is most commonly used, whereas the smaller diameter 0.018-inch guidewires are reserved for selective small vessel angiography such as infrageniculate or carotid lesions. In addition to diameter size, guidewires come in varying lengths, usually ranging from 180 to 260 cm in length. Increasing the length of the wire always makes it more difficult to handle and increases the risk of contamination. While performing a procedure, it is important to maintain the guidewire across the lesion until the completion arteriogram has been satisfactorily completed.

The stiffness of the guidewire is also an important characteristic. Stiff wires allow for passage of large aortic stent graft devices without kinking. They are also useful when trying to perform sheath or catheter exchanges around a tortuous artery. An example of a stiff guidewire is the Amplatz wire. Hydrophilic coated guidewires, such as the Glidewire, have become invaluable tools for assisting in difficult catheterizations. The coating is primed by bathing the guidewire in saline solution. The slippery nature of this guidewire along with its torque capability significantly facilitate in difficult catheterizations. Guidewires also come in various tip configurations. Angled tip wires like the angled Glidewire can be steered to manipulate a catheter across a tight stenosis or to select a specific branch of a vessel. The Rosen wire has a soft curved end, which makes it ideal for renal artery stenting. The soft curl of this wire prevents it from perforating small renal branch vessels.
Hemostatic Sheaths
The hemostatic sheath is a device through which endovascular procedures are performed. The sheath acts to protect the vessel from injury as wires and catheters are introduced (Fig. 23-8). A one-way valve prevents bleeding through the sheath, and a side-port allows contrast or heparin flushes to be administered during the procedure. Sheaths are sized by their inner diameter. The most commonly used sheaths for percutaneous access have a 5- to 9-French inner diameter, but with open surgical exposure of the CFA, sheaths as large as 26 French can be introduced. Sheaths also vary in length, and long sheaths are available so that interventions remote from the site of arterial access can be performed.

Catheters
A wide variety of catheters exist that differ primarily in the configuration of the tip. The multiple shapes permit access to vessels of varying dimensions and angulations. Catheters are used to perform angiography and protect the passage of balloons and stents, and they can be used to direct the guidewire through tight stenoses or tortuous vessels.

Angioplasty Balloons
Angioplasty balloons differ primarily in their length and diameter, as well as the length of the catheter shaft. As balloon technology has advanced, lower profiles have been manufactured (i.e., the size that the balloon assumes upon deflation). Balloons are used to perform angioplasty on vascular stenoses, to deploy stents, and to assist with additional expansion after insertion of self-expanding stents (Fig. 23-9). Besides length and diameter, operators need to be familiar with several other balloon characteristics. Noncompliant and low-compliance balloons tend to be inflated to their preset diameter and offer greater dilating force at the site of stenosis. Low-compliance balloons are the mainstay for peripheral intervention. Lower profile balloons are less likely to get caught during passage through stents and are easier to pull out of sheaths. Under fluoroscopic guidance, balloon inflation is performed until the waist of the atherosclerotic lesion disappears and the balloon is at the full profile. The duration of balloon inflation and pressures used for the angioplasty depend on the indication for the intervention and the location and characteristics of the lesion being treated. Frequently, several inflations are required to achieve a full profile of the balloon. Occasionally, a lower profile balloon is needed to predilate the tight stenosis so that the selected balloon catheter can cross the lesion. After inflation, most balloons do not regain their preinflation diameter and assume a larger profile. Trackability, pushability, and cross-ability of the balloon should all be considered when choosing a particular balloon. Lastly, shoulder length is an important characteristic to consider when selecting a balloon because of the potential to cause injury during performance of PTA in adjacent arterial segments. There is always risk of causing dissection or rupture during PTA; thus, a completion angiogram is performed while the wire is still in place. Leaving the wire in place provides access for repeating the procedure, placing a stent or stent graft if warranted.

Stents
Vascular stents are commonly used after an inadequate angioplasty with dissection or elastic recoil of an arterial stenosis. They serve to buttress collapsible vessels and help prevent atherosclerotic restenosis. Appropriate indications for primary stenting of a lesion without an initial trial of angioplasty alone are evolving in manners that are dependent on the extent and site of the lesion. Stents are manufactured from a variety of metals including stainless steel, tantalum, cobalt-based alloy,

Figure 23-8. All percutaneous endovascular procedures are performed through an introducer sheath (large arrow), which provides an access conduit from skin to intravascular compartment. The sheath also acts to protect the vessel from injury as guidewires (small arrows) and catheters are introduced.

Figure 23-9. A. An artery with luminal narrowing caused by plaque. B. A balloon angioplasty catheter is positioned within the diseased artery, which is inflated to enlarge the intravascular channel. C. The plaque is compressed with widened flow lumen as the result of balloon angioplasty.
and nitinol. Vascular stents are classified into two basic categories: balloon-expandable stents and self-expanding stents.

Self-expanding stents (Fig. 23-10) are deployed by retracting a restraining sheath and usually consist of Elgiloy (a cobalt, chromium, nickel alloy) or nitinol (a shape memory alloy composed of nickel and titanium), the latter of which will contract and assume a heat-treated shape above a transition temperature that depends on the composition of the alloy. Self-expanding stents will expand to a final diameter that is determined by stent geometry, hoop strength, and vessel size. The self-expanding stent is mounted on a central shaft and is placed inside an outer sheath. It relies on a mechanical spring-like action to achieve expansion. With deployment of these stents, there is some degree of foreshortening that has to be taken into account when choosing the area of deployment. In this way, self-expanding stents are more difficult to place with absolute precision. There are several advantages related to self-expanding stents. Self-expanding stents generally come in longer lengths than balloon-expandable stents and are therefore used to treat long and tortuous lesions. Their ability to continually expand after delivery allows them to accommodate adjacent vessels of different size. This makes these stents ideal for placement in the internal carotid artery. These stents are always oversized by 1 to 2 mm relative to the largest diameter of normal vessel adjacent to the lesion in order to prevent immediate migration.

Balloon-expandable stents are usually composed of stainless steel, mounted on an angioplasty balloon, and deployed by balloon inflation (Fig. 23-11). They can be manually placed on a chosen balloon catheter or obtained premounted on a balloon catheter. The capacity of a balloon-expandable stent to shorten in length during deployment depends on both stent geometry and the final diameter to which the balloon is expanded. These stents are more rigid and are associated with a shorter time to complete endothelialization. They are often of limited flexibility and have a higher degree of crush resistance when compared to self-expanding stents. This makes them ideal for short-segment lesions, especially those that involve the ostia such as proximal common iliac or renal artery stenosis.

An important area of evolution in endovascular therapy in recent years is the development of drug-eluting stents (DES). These stents are usually composed of nitinol and have various anti-inflammatory drugs bonded to them. Over time, the stents release the drug into the surrounding arterial wall and help prevent restenosis. Numerous randomized controlled trials have proven their benefit in coronary arteries. Clinical studies have similarly proved early efficacy of DES in the treatment of peripheral arterial disease.

**Stent Grafts**

The combination of a metal stent covered with fabric gave birth to the first stent grafts. Covered stents have been designed with either a surrounding PTFE or polyester fabric and have been used predominantly for treatment of traumatic vascular lesions, including arterial disruption and arteriovenous fistulas (Fig. 23-12). However, these devices may well find a growing role in treatment of iliac or femoral arterial occlusive disease as well as popliteal aneurysms.

Endovascular aneurysm repair using the concept of stent grafts was initiated by Parodi in 1991. Since that time, a large number of endografts have been inserted under the auspice of clinical trials initially and now as Food and Drug Administration (FDA)-approved devices. Currently there are more than

![Figure 23-10](image1.png) **Figure 23-10.** Self-expanding stents are made of tempered stainless steel or nitinol, an alloy of nickel and titanium, and are restrained when folded inside a delivery catheter. After being released from the restraining catheter, the self-expanding stents will expand to a final diameter that is determined by stent geometry, hoop strength, and vessel size.

![Figure 23-11](image2.png) **Figure 23-11.** In a balloon-expandable stent, the stent is pre-mounted on a balloon catheter. The balloon stretches the stent members beyond their elastic limit. The stent is deployed by full balloon expansion. This type of stent has a higher degree of crush resistance when compared to self-expanding stents, which is ideal for short-segment calcified ostial lesions.
eighth FDA-approved endovascular devices for abdominal aortic aneurysm repair. In general, majority of these devices require that patients have an infrarenal aneurysm with at least a 15-mm proximal aortic neck below the renal arteries and not greater than 60° of angulation. For those patients with associated common iliac artery aneurysmal disease, endovascular treatment can be achieved by initial coil embolization of the ipsilateral hypogastric artery with extension of the endovascular device into the external iliac artery. Newer generation devices with branched endograft can be deployed in the internal iliac artery while maintain in-line flow from the common iliac to external iliac artery to exclude the common iliac artery aneurysm. Recent clinical trials have demonstrated clinical efficacy of fenestrated aortic endograft in treating aneurysm involving the visceral segment of the abdominal aorta. The FDA has similarly approved several thoracic endograft devices for the treatment of descending thoracic aortic aneurysm. Clinical studies have similarly demonstrated durability and efficacy of thoracic aortic devices in patients with traumatic aortic transections and aortic dissections.

CAROTID ARTERY DISEASE

Atherosclerotic occlusive plaque is by far the most common pathology seen in the carotid artery bifurcation. Thirty percent to 60% of all ischemic strokes are related to atherosclerotic carotid bifurcation occlusive disease. In the following section, we first focus our discussion on the clinical presentation, diagnosis, and management—including medical therapy, surgical carotid endarterectomy, and stenting—of atherosclerotic carotid occlusive disease. In the second part of the section, we provide a review on other less common nonatherosclerotic diseases involving the extracranial carotid artery, including kink and coil, fibromuscular dysplasia, arterial dissection, aneurysm, radiation arteritis, Takayasu’s arteritis, and carotid body tumor.

Epidemiology and Etiology of Carotid Occlusive Disease

Approximately 700,000 Americans suffer a new or recurrent stroke each year. Eighty-five percent of all strokes are ischemic, and 15% are hemorrhagic. Hemorrhagic strokes are caused by head trauma or spontaneous disruption of intracerebral blood vessels. Ischemic strokes are due to hypoperfusion from arterial occlusion or, less commonly, to decreased flow resulting from proximal arterial stenosis and poor collateral network. Common causes of ischemic strokes are cardiogenic emboli in 35%, carotid artery disease in 30%, lacunar in 10%, miscellaneous in 10%, and idiopathic in 15%. The term cerebrovascular accident is often used interchangeably to refer to an ischemic stroke. A transient ischemic attack (TIA) is defined as a temporary focal cerebral or retinal hypoperfusion state that resolves spontaneously within 24 hours after its onset. However, the majority of TIs resolve within minutes, and longer-lasting neurologic deficits more likely represent a stroke. Recently, the term brain attack has been coined to refer to an acute stroke or TIA, denoting the condition as a medical emergency requiring immediate attention, similar to a heart attack.

Stroke due to carotid bifurcation occlusive disease is usually caused by atheroemboli (Fig. 23-13). The carotid bifurcation is an area of low flow velocity and low shear stress. As the blood circulates through the carotid bifurcation, there is separation of flow into the low-resistance internal carotid artery and the high-resistance external carotid artery. Characteristically, atherosclerotic plaque forms in the outer wall opposite to the flow divider (Fig. 23-14). Atherosclerotic plaque formation is complex, beginning with intimal injury, platelet deposition, smooth muscle cell proliferation, and fibroplasia, and leading to subsequent luminal narrowing. With increasing degree of stenosis in the internal carotid artery, flow becomes more turbulent, and the risk of atheroembolization escalates. The severity of stenosis is commonly divided into three categories according to the luminal diameter reduction: mild (<50%), moderate (50–69%), and severe (70–99%). Severe carotid stenosis is a strong predictor for stroke. In turn, a prior history of neurologic symptoms (TIA or stroke) is an important determinant for recurrent ipsilateral stroke. The risk factors for the development of carotid artery bifurcation disease are similar to those causing atherosclerotic occlusive disease in other vascular beds. Increasing age, male gender, hypertension, tobacco smoking, diabetes mellitus, homocysteinemia, and hyperlipidemia are well-known predisposing factors for the development of atherosclerotic occlusive disease.

Clinical Manifestations of Cerebral Ischemia

TIA is a focal loss of neurologic function, lasting for less than 24 hours. Crescendo TIAs refer to a syndrome comprising repeated TIAs within a short period of time that is characterized by complete neurologic recovery in between. At a minimum, the term should probably be reserved for those with either daily events or multiple resolving attacks within 24 hours.
extracranial disease and poor intracranial collateral recruitment. Reversible ischemic neurologic deficits refer to ischemic focal neurologic symptoms lasting longer than 24 hours but resolving within 3 weeks. When a neurologic deficit lasts longer than 3 weeks, it is considered a completed stroke. Stroke in evolution refers to progressive worsening of the neurologic deficit, either linearly over a 24-hour period or interspersed with transient periods of stabilization and/or partial clinical improvement.

Patients who suffer cerebrovascular accidents typically present with three categories of symptoms including ocular symptoms, sensory/motor deficit, and/or higher cortical dysfunction. The common ocular symptoms associated with extracranial carotid artery occlusive disease include amaurosis fugax and presence of Hollenhorst plaques. Amaurosis fugax, commonly referred to as transient monocular blindness, is a temporary loss of vision in one eye that patients typically describe as a window shutter coming down or grey shedding of the vision. This partial blindness usually lasts for a few minutes and then resolves. Most of these phenomena (>90%) are due to embolic occlusion of the main artery or the upper or lower divisions. Monocular blindness progressing over a 20-minute period suggests a migrainous etiology. Occasionally, the patient will recall no visual symptoms while the optician notes a yellowish plaque within the retinal vessels, which is also known as Hollenhorst plaque. These plaques are frequently derived from cholesterol embolization from the carotid bifurcation and warrant further investigation. Additionally, several ocular symptoms may be caused by microembolization from extracranial carotid diseases including monocular visual loss due to retinal artery or optic nerve ischemia, the ocular ischemia syndrome, and visual field deficits secondary to cortical infarction and ischemia of the optic tracts. Typical motor and/or sensory symptoms associated with cerebrovascular accidents are lateralized or focal neurologic deficits. Ischemic events tend to have an abrupt onset, with the severity of the insult being apparent from the onset and not usually associated with seizures or paresthesia. In contrast, they represent loss or diminution of neurologic function. Furthermore, motor or sensory deficits can be unilateral or

Hemodynamic TIAs represent focal cerebral events that are aggravated by exercise or hemodynamic stress and typically occur after short bursts of physical activity, postprandially, or after getting out of a hot bath. It is implied that these are due to severe

Figure 23-13. Stroke due to carotid bifurcation occlusive disease is usually caused by atheroemboli arising from the internal carotid artery, which provides the majority of blood flow to the cerebral hemisphere. With increasing degree of stenosis in the carotid artery, flow becomes more turbulent, and the risk of atheroembolization escalates.

Figure 23-14. A. The carotid bifurcation is an area of low flow velocity and low shear stress. As the blood circulates through the carotid bifurcation, there is separation of flow into the low-resistance internal carotid artery and the high-resistance external carotid artery. B. The carotid atherosclerotic plaque typically forms in the outer wall opposite to the flow divider due in part to the effect of the low shear stress region, which also creates a transient reversal of flow during the cardiac cycle.
bilateral, with the upper and lower limbs being variably affected depending on the site of the cerebral lesion. The combination of a motor and sensory deficit in the same body territory is suggestive of a cortical thromboembolic event as opposed to lacunar lesions secondary to small vessel disease of the penetrating arterioles. However, a small proportion of the latter may present with a sensorimotor stroke secondary to small vessel occlusion within the posterior limb of the internal capsule. Pure sensory and pure motor strokes and those strokes where the weakness affects one limb only or does not involve the face are more typically seen with lacunar as opposed to cortical infarction. A number of higher cortical functions, including speech and language disturbances, can be affected by thromboembolic phenomena from the carotid artery, with the most important clinical example for the dominant hemisphere being dysphasia or aphasia and visuospatial neglect being an example of nondominant hemisphere injury.

**Diagnostic Evaluation**

Duplex ultrasonography is the most widely used screening tool to evaluate for atherosclerotic plaque and stenosis of the extracranial carotid artery. It is also commonly used to monitor patients serially for progression of disease or after intervention (carotid endarterectomy or angioplasty). Duplex ultrasound of the carotid artery combines B-mode gray-scale imaging and Doppler waveform analysis. Characterization of the carotid plaque on gray-scale imaging provides useful information about its composition. However, there are currently no universal recommendations that can be made based solely on the sonographic appearance of the plaque. On the other hand, criteria have been developed and well refined for grading the degree of carotid stenosis based primarily on Doppler-derived velocity waveforms.

The external carotid artery has a high-resistance flow pattern with a sharp systolic peak and a small amount of flow in diastole. In contrast, a normal internal carotid artery will have a low-resistance flow pattern with a broad systolic peak and a large amount of flow during diastole. The flow pattern in the common carotid artery resembles that in the internal carotid artery, as 80% of the flow is directed to the internal carotid artery, with waveforms that have broad systolic peaks and moderate amount of flow during diastole. Conventionally, velocity measurements are recorded in the common, external, carotid bulb, and the proximal, mid, and distal portions of the internal carotid artery. Characteristically, the peak systolic velocity is increased at the site of the vessel stenosis. The end-diastolic velocity is increased with greater degree of stenosis. In addition, stenosis of the internal carotid artery can lead to color shifts with color mosaics indicating a poststenotic turbulence. DAMPENING of the Doppler velocity waveforms is typically seen in areas distal to severe carotid stenosis where blood flow is reduced. It is well known that occlusion of the ipsilateral internal carotid artery can lead to a “falsely” elevated velocity on the contralateral side due to an increase in compensatory blood flow. In the presence of a high-grade stenosis or occlusion of the internal carotid artery, the ipsilateral common carotid artery displays high flow resistance waveforms, similar to those seen in the external carotid artery. If there is a significant stenosis in the proximal common carotid artery, its waveforms may be damped with low velocities.

The Doppler grading systems of carotid stenosis were initially established by comparison to angiographic findings of disease. Studies have shown variability in the measurements of the duplex properties by different laboratories, as well as heterogeneity in the patient population, study design, and techniques. One of the most commonly used classifications was established at the University of Washington School of Medicine in Seattle, Washington. Diameter reduction of 50% to 79% is defined by peak systolic velocity greater than 125 cm/s with extensive spectral broadening. For stenosis in the range of 80% to 99%, the peak systolic velocity is greater than 125 cm/s, and peak diastolic velocity is greater than 140 cm/s. The ratio of internal carotid to common carotid artery peak systolic velocity has also been part of various ultrasound diagnostic classifications. A ratio greater than 4 is a great predictor of angiographic stenosis of 70% to 99%. A multispecialty consensus panel has developed a set of criteria for grading carotid stenosis by duplex examination (Table 23-3).

MRA is increasingly being used to evaluate for atherosclerotic carotid occlusive disease and intracranial circulation. MRA is noninvasive and does not require iodinated contrast agents. MRA uses phase contrast or time-of-flight, with either two-dimensional or three-dimensional data sets for greater accuracy. Three-dimensional contrast-enhanced MRA allows data to be obtained in coronal and sagittal planes with improved image qualities due to shorter study time. In addition, the new MRA techniques allow for better reformation of images in various planes to allow better grading of stenosis. There have been numerous studies comparing the sensitivity and specificity of

---

**Table 23-3**

<table>
<thead>
<tr>
<th>DEGREE OF STENOSIS (%)</th>
<th>ICA PSV (CM/S)</th>
<th>ICA/CCA PSV RATIO</th>
<th>ICA EDV (CM/S)</th>
<th>PLAQUE ESTIMATE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>&lt;125</td>
<td>&lt;2.0</td>
<td>&lt;40</td>
<td>None</td>
</tr>
<tr>
<td>&lt;50–69</td>
<td>&lt;125</td>
<td>&lt;2.0</td>
<td>&lt;40</td>
<td>&lt;50</td>
</tr>
<tr>
<td>50–79</td>
<td>125–230</td>
<td>2.0–4.0</td>
<td>40–100</td>
<td>≥50</td>
</tr>
<tr>
<td>≥70 to less than near occlusion</td>
<td>&gt;230</td>
<td>&gt;4.0</td>
<td>&gt;100</td>
<td>≥50</td>
</tr>
<tr>
<td>Near occlusion</td>
<td>High, low, or not detected</td>
<td>Variable</td>
<td>Variable</td>
<td>Visible</td>
</tr>
<tr>
<td>Total occlusion</td>
<td>Not detected</td>
<td>Not applicable</td>
<td>Not detected</td>
<td>Visible, no lumen</td>
</tr>
</tbody>
</table>

*Plaque estimate (diameter reduction) with gray-scale and color Doppler ultrasound.
CCA = common carotid artery; EDV = end-diastolic velocity; ICA = internal carotid artery; PSV = peak systolic velocity.
MRA imaging for carotid disease to duplex and selective contrast angiography. Magnetic resonance imaging (MRI) of the brain is essential in the assessment of acute stroke patients. MRI with diffusion-weighted imaging can differentiate areas of acute ischemia, areas still at risk for ischemia (penumbra), and chronic cerebral ischemic changes. However, computed tomography (CT) imaging remains the most expeditious test in the evaluation of acute stroke patients to rule out intracerebral hemorrhage. Recently, multidetector CTA has gained increasing popularity in the evaluation of carotid disease. This imaging modality can provide volume rendering, which allows rotation of the object with accurate anatomic structures from all angles (Fig. 23-15). The advantages of CTA over MRA include faster data acquisition time and better spatial resolution. However, grading of carotid stenosis by CTA requires further validation at the time of this writing before it can be widely applied.

Historically, DSA has been the gold standard test to evaluate the extra- and intracranial circulation (Fig. 23-16). This is an invasive procedure, typically performed via a transfemoral puncture, and involves selective imaging of the carotid and vertebral arteries using iodinated contrast. The risk of stroke during cerebral angiography is generally reported at approximately 1% and is typically due to atheroembolization related to wire and catheter manipulation in the arch aorta or proximal branch vessels. Over the last few decades, however, the incidence of neurologic complications following angiography has been reduced, due to the use of improved guidewires and catheters, better resolution digital imaging, and increased experience. Local access complications of angiography are infrequent and include development of hematoma, pseudoaneurysm, distal embolization, and acute vessel thrombosis. Currently, selective angiography is particularly used for patients with suspected intracranial disease and for patients in whom percutaneous revascularization is considered. The techniques of carotid angioplasty and stenting for carotid bifurcation occlusive disease are described in detail later in this chapter. We generally use CTA or MRA to get information about the aortic arch anatomy and presence of concomitant intracranial disease and collateral pathway in planning our strategy for carotid stenting or endarterectomy.

Treatment of Carotid Occlusive Disease
Conventionally, patients with carotid bifurcation occlusive disease are divided into two broad categories: patients without prior history of ipsilateral stroke or TIA (asymptomatic) and those with prior or current ipsilateral neurologic symptoms (symptomatic). It is estimated that 15% of all strokes are preceded by a TIA. The 90-day risk of a stroke in a patient presenting with a TIA is 3% to 17%. According to the Cardiovascular Health Study, a longitudinal population-based study of coronary artery disease and stroke in men and women, the prevalence of TIA in men was 2.7% for ages of 65 and 69 and 3.6% for ages 75 to 79; the prevalence in women was 1.4% and 4.1%, respectively. There have been several studies reporting on the effectiveness of stroke prevention with medical treatment and carotid endarterectomy for symptomatic patients with moderate to severe carotid stenosis. Early and chronic aspirin therapy has been shown to reduce stroke recurrence rate in several large clinical trials.

Symptomatic Carotid Stenosis
Currently, most stroke neurologists prescribe both aspirin and clopidogrel for secondary

Figure 23-15. A. Carotid computed tomography angiography is a valuable imaging modality that can provide a three-dimensional image reconstruction with high image resolution. A carotid artery occlusion is noted in the internal carotid artery B. The entire segment of extracranial carotid artery is visualized from the thoracic compartment to the base of skull.

Figure 23-16. A carotid angiogram reveals an ulcerated carotid plaque (arrow) in the proximal internal carotid artery, which also resulted in a high-grade internal carotid artery stenosis.
stroke prevention in patients who have experienced a TIA or stroke. In patients with symptomatic carotid stenosis, the degree of stenosis appears to be the most important predictor in determining risk for an ipsilateral stroke. The risk of a recurrent ipsilateral stroke in patients with severe carotid stenosis approaches 40%. Two large multicenter randomized clinical trials, the European Carotid Surgery Trial (ECST) and the North American Symptomatic Carotid Endarterectomy Trial (NASCET), have both shown a significant risk reduction in stroke for patients with symptomatic high-grade stenosis (70–99%) undergoing carotid endarterectomy when compared to medical therapy alone.20,21 Although there has been much discussion regarding the different methodologies used in the measurement of carotid stenosis and calculation of the lifetable data between the two studies, both of these studies had similar clinical results. The findings of these two landmark trials have also been reanalyzed in many subsequent publications. The main conclusions of the trials remain validated and widely acknowledged. Briefly, the NASCET study showed that for high-grade carotid stenosis, the cumulative risk of ipsilateral stroke was 26% in the medically treated group and 9% in the surgically treated group at 2 years. For patients with moderate carotid artery stenosis (50–69%), the benefit of carotid endarterectomy is less but still favorable when compared to medical treatment alone; the 5-year fatal or nonfatal ipsilateral stroke rate was 16% in the surgically treated group versus 22% in the medically treated group.22 The risk of stroke was similar for the remaining group of symptomatic patients with less than 50% carotid stenosis, whether they had endarterectomy or medical treatment alone. The ECST reported similar stroke risk reduction for patients with severe symptomatic carotid stenosis and no benefit in patients with mild stenosis when carotid endarterectomy was performed versus medical therapy.21

The optimal timing of carotid intervention after acute stroke, however, remains debatable. Earlier studies showed an increased rate of postoperative stroke exacerbation and conversion of a bland to hemorrhagic infarction when carotid endarterectomy was carried out within 5 to 6 weeks after acute stroke. The dismal outcome reported in the early experience was likely related to poor patient selection. The rate of stroke recurrence is not insignificant during the interval period and may be reduced with early intervention for symptomatic carotid stenosis. Contemporary series have demonstrated acceptable low rates of perioperative complications in patients undergoing carotid endarterectomy within 4 weeks after acute stroke.22 In a recent retrospective series, carotid artery stenting when performed early (<2 weeks) after the acute stroke was associated with higher mortality than when delayed (>2 weeks).23

Asymptomatic Carotid Stenosis. Whereas there is universal agreement that carotid revascularization (endarterectomy or stenting) is effective in secondary stroke prevention for patients with symptomatic moderate and severe carotid stenosis, the management of asymptomatic patients remains an important controversy to be resolved. Generally, the detection of carotid stenosis in asymptomatic patients is related to the presence of a cervical bruit or based on screening duplex ultrasound findings. In one of the earlier observational studies, the authors showed that the annual occurrence rate of neurologic symptoms was 4% in a cohort of 167 patients with asymptomatic cervical bruits followed prospectively by serial carotid duplex scan.24 The mean annual rate of carotid stenosis progression to a greater than 50% stenosis was 8%. The presence of or progression to a greater than 80% stenosis correlated highly with either the development of a total occlusion of the internal carotid artery or new symptoms. The major risk factors associated with disease progression were cigarette smoking, diabetes mellitus, and age. This study supported the contention that it is prudent to follow a conservative course in the management of asymptomatic patients presenting with a cervical bruit.

One of the first randomized clinical trials on the treatment of asymptomatic carotid artery stenosis was the Asymptomatic Carotid Atherosclerosis Study (ACAS), which evaluated the benefits of medical management with antiplatelet therapy versus carotid endarterectomy.25 Over a 5-year period, the risk of ipsilateral stroke in individuals with a carotid artery stenosis greater than 60% was 5.1% in the surgical arm. On the other hand, the risk of ipsilateral stroke in patients treated with medical management was 11%. Carotid endarterectomy produced a relative risk reduction of 53% over medical management alone. The results of a larger randomized trial from Europe, the Asymptomatic Carotid Surgery Trial (ACST), recently confirmed similar beneficial stroke risk reduction for patients with asymptomatic, greater than 70% carotid stenosis undergoing endarterectomy versus medical therapy.26 An important point derived from this latter trial was that even with improved medical therapy, including the addition of statin drugs and clopidogrel, medical therapy was still inferior to endarterectomy in the primary stroke prevention for patients with high-grade carotid artery stenosis. It is generally agreed that asymptomatic patients with severe carotid stenosis (80–99%) are at significantly increased risk for stroke and stand to benefit from either surgical or endovascular revascularization. However, revascularization for asymptomatic patients with a less severe degree of stenosis (60–79%) remains controversial.

**Carotid Endarterectomy Versus Angioplasty and Stenting**

Currently, the argument is no longer whether medical therapy alone is inferior to surgical endarterectomy in stroke prevention for severe carotid stenosis. Rather, the debate now revolves around whether carotid angioplasty and stenting produce the same benefits demonstrated by carotid endarterectomy. Since carotid artery stenting was approved by the FDA for clinical application in 2004, this percutaneous procedure has become a treatment alternative in patients who are deemed “high risk” for endarterectomy (Table 23-4). In contrast to many endovascular peripheral arterial interventions, percutaneous carotid stenting represents a much more challenging procedure, because it requires complex catheter-based skills using the 0.014-inch guidewire system and distal protection device. Moreover, current carotid stent devices predominantly use the monorail guidewire system, which requires more technical agility compared with the over-the-wire catheter system that is routinely used in peripheral interventions. This percutaneous intervention often requires balloon angioplasty and stent placement through a long carotid guiding sheath via a groin approach. Poor technical skills can result in devastating treatment complications such as stroke, which can occur in part due to plaque embolization during the balloon angioplasty and stenting of the carotid artery. Because of these various procedural components that require high technical proficiency, many early clinical investigations of carotid artery stenting, which included physicians with little or no carotid stenting experience, resulted in alarmingly poor
Table 23-4

Conditions qualifying patients as high surgical risk for carotid endarterectomy

<table>
<thead>
<tr>
<th>ANATOMIC FACTORS</th>
<th>PHYSIOLOGIC FACTORS</th>
</tr>
</thead>
<tbody>
<tr>
<td>• High carotid bifurcation (above C2 vertebral body)</td>
<td>• Age ≥80 years</td>
</tr>
<tr>
<td>• Low common carotid artery (below clavicle)</td>
<td>• Left ventricular ejection fraction ≤30%</td>
</tr>
<tr>
<td>• Contralateral carotid occlusion</td>
<td>• New York Heart Association class III/IV congestive heart failure</td>
</tr>
<tr>
<td>• Restenosis of ipsilateral prior carotid endarterectomy</td>
<td>• Unstable angina: Canadian Cardiovascular Society class III/IV angina pectoris</td>
</tr>
<tr>
<td>• Previous neck irradiation</td>
<td>• Recent myocardial infarction</td>
</tr>
<tr>
<td>• Prior radical neck dissection</td>
<td>• Clinically significant cardiac disease (congestive heart failure, abnormal stress test, or need for coronary revascularization)</td>
</tr>
<tr>
<td>• Contralateral laryngeal nerve palsy</td>
<td>• Severe chronic obstructive pulmonary disease</td>
</tr>
<tr>
<td>• Presence of tracheostomy</td>
<td>• End-stage renal disease on dialysis</td>
</tr>
</tbody>
</table>

Clinical outcomes. A Cochrane review noted that, before 2006, a total of 1269 patients had been studied in five randomized controlled trials comparing percutaneous carotid intervention and surgical carotid reconstruction.27 Taken together, these trials revealed that carotid artery stenting had a greater procedural risk of stroke and death when compared to carotid endarterectomy (odds ratio, 1.33; 95% confidence interval, 0.86–2.04). Additionally, a greater incidence of carotid restenosis was noted in the stenting group than the endarterectomy cohorts.

However, the constant improvement of endovascular devices, procedural techniques, and adjunctive pharmacologic therapy will likely improve the treatment success of percutaneous carotid intervention. Critical appraisals of several prospective randomized trials comparing the efficacy of carotid stenting versus endarterectomy are available for review.28 Two recently published randomized controlled trial, the Carotid Revascularization Endarterectomy Versus Stent Trial (CREST) and the International Carotid Stenting Study (ICSS) have reported somewhat differing results.29 CREST compared the efficacy of carotid endarterectomy and carotid stenting in both symptomatic and asymptomatic patients.30 Primary end points included 30-day periprocedural composite death, stroke, myocardial infarction, or any ipsilateral stroke up to 4 years. CREST investigators reported no difference between stenting (5.2%) and endarterectomy (4.5%) in terms of primary end point. When each variable was independently analyzed, there was a higher rate of stroke in the stenting group at 30 days (4.1% vs. 2.3%) and a higher rate of myocardial infarction in the endarterectomy group (2.3% vs. 1.1%). The ICSS was a multicenter, international, randomized controlled trial comparing carotid stenting versus endarterectomy in patients with symptomatic carotid stenosis.31 The risk of stroke, death, and myocardial infarction in the stenting group (8.5%) was significantly higher than in the surgical arm (5.2%). The finding that carotid endarterectomy is safer than carotid stenting is also supported by the results of an MRI substudy, which showed significantly more new lesions by diffusion-weighted imaging in the carotid stenting than the carotid endarterectomy patients.

All available randomized studies have provided some answers and raised some questions. Some ongoing clinical trials will undoubtedly provide more insights on the efficacy of carotid stenting in the near future. Currently, the Society for Vascular Surgeons recommends carotid endarterectomy as first-line treatment for most symptomatic patients with stenosis of 50% to 99% and asymptomatic patients with stenosis of 60% to 99%.32 The perioperative risk of stroke and death in asymptomatic patients must be below 3% to ensure benefit for the patient. Carotid artery stenting should be reserved for symptomatic patients with stenosis of 50% to 99% at high risk for carotid endarterectomy for anatomic or medical reasons. Carotid artery stenting is not recommended for asymptomatic patients at this time. Asymptomatic patients at high risk for intervention or with a life expectancy of less than 3 years should be considered for medical management as the first-line therapy.

Surgical Techniques of Carotid Endarterectomy

Although carotid endarterectomy is one of the earliest vascular operations ever described and its techniques have been perfected in the last two decades, surgeons continue to debate many aspects of this procedure. For instance, there is no universal agreement with regard to the best anesthetic of choice, the best intraoperative cerebral monitoring, whether to “routinely” shunt, open versus eversion endarterectomy, and patch versus primary closure. Various anesthetic options are available for patient undergoing carotid endarterectomy including general, local, and regional anesthesia. Typically, the anesthesia of choice depends on the preference of the surgeon, anesthesiologist, and patient. However, depending on the anesthetic given, the surgeon must decide whether intraoperative cerebral monitoring is necessary or intra-arterial carotid shunting will be used. In general, if the patient is awake, then his or her abilities to respond to commands during carotid clamp period determine the adequacy of collateral flow to the ipsilateral hemisphere. On the other hand, intraoperative eneroencephalogram (EEG) or transcranial power Doppler (TCD) has been used to monitor for adequacy of cerebral perfusion during the clamp period for patients undergoing surgery under general anesthesia. Focal ipsilateral decreases in amplitudes and slowing of EEG waves are indicative of cerebral ischemia. Similarly, a decrease to less than 50% of baseline velocity in the ipsilateral middle cerebral artery is a sign of cerebral ischemia. For patients with poor collateral flow exhibiting signs of cerebral ischemia, intra-arterial carotid shunting with removal of the clamp will restore cerebral flow for the remaining part of the surgery. Stump pressures have been used to determine the need for intra-arterial carotid shunting. Some surgeons prefer to shunt all patients on a routine basis and do not use intraoperative cerebral monitoring.
The patient’s neck is slightly hyperextended and turned to the contralateral side, with a roll placed between the shoulder blades. An oblique incision is made along the anterior border of the sternocleidomastoid muscle centered on top of the carotid bifurcation (Fig. 23-17). The platysma is divided completely. Typically, tributaries of the anterior jugular vein are ligated and divided. The dissection is carried medial to the sternocleidomastoid. The superior belly of the omohyoid muscle is usually encountered just anterior to the common carotid artery. This muscle can be divided. The carotid fascia is incised, and the common carotid artery is exposed. The common carotid artery is mobilized cephalad toward the bifurcation. The dissection of the carotid bifurcation can cause reactive bradycardia related to stimulation of the carotid body. This reflex can be blunted with injection of lidocaine 1% into the carotid body or reversed with administration of intravenous atropine. A useful landmark in the dissection of the carotid bifurcation is the common facial vein. This vein can be ligated and divided. Frequently, the 12th cranial nerve (hypoglossal nerve) traverses the carotid bifurcation just behind the common facial vein. The external carotid artery is mobilized just enough to get a clamp across. Often, a branch of the external carotid artery crossing to the sternocleidomastoid can be divided to allow further cephalad mobilization of the internal carotid artery. For high bifurcation, division of the posterior belly of the digastric muscle is helpful in establishing distal exposure of the internal carotid artery.

Intravenous heparin sulfate (1 mg/kg) is routinely administered just prior to carotid clamping. The internal carotid artery is clamped first using a soft noncrushing vascular clamp to prevent distal embolization. The external and common carotid arteries are clamped subsequently. A longitudinal arteriotomy is made in the distal common carotid artery and extended into the bulb and past the occlusive plaque into the normal part of the internal carotid artery. Endarterectomy is carried out to remove the occlusive plaque (Fig. 23-18). If necessary, a temporary shunt can be inserted from the common carotid artery to the internal carotid artery to maintain continuous antegrade cerebral blood flow (Fig. 23-19). Typically, a plane is teased out from the vessel wall, and the entire plaque is elevated and removed. The distal transition line in the internal carotid artery where the plaque had been removed must be examined carefully and should be smooth. Tacking sutures are placed when an intimal flap remains in this transition to ensure no obstruction to flow (Fig. 23-20). The occlusive plaque is usually removed from the origin of the external carotid artery using the eversion technique. The endarterectomized surface is then irrigated and any debris removed. A patch (autogenous saphenous vein, synthetic such as polyester, PTFE, or biologic material) is sewn to close the arteriotomy (Fig. 23-21). Whether patch closure is necessary in all patients and which patch is the best remain controversial. However, most surgeons agree that patch closure is
SPECIFIC CONSIDERATIONS

B. Indications for Endarterectomy

The distal transition line (left side of the picture) in the internal carotid artery where the plaque had been removed must be examined carefully and should be smooth. Tacking sutures (arrows) are placed when an intimal flap remains in this transition to ensure no obstruction to flow.

Figure 23-20. The distal transition line (left side of the picture) in the internal carotid artery where the plaque had been removed must be examined carefully and should be smooth. Tacking sutures (arrows) are placed when an intimal flap remains in this transition to ensure no obstruction to flow.

Indicated particularly for the small vessel (<7 mm). The eversion technique has also been advocated for removing the plaque from the internal carotid artery. In the eversion technique, the internal carotid artery is transected at the bulb, the edges of the divided vessel are everted, and the occluding plaque is “peeled” off the vessel wall. The purported advantages of the eversion technique are no need for patch closure and a clear visualization of the distal transition area. Reported series have not shown a clear superiority of one technique over the others. Surgeons will likely continue to use the technique of their choice. Just prior to completion of the anastomosis to close the arteriotomy, we thoroughly flush the vessels of any potential debris. When the arteriotomy is closed, flow is restored to the external carotid artery first and to the internal carotid artery second. Intravenous protamine sulfate can be given to reverse the effect of heparin anticoagulation following carotid endarterectomy. The wound is closed in layers. After surgery, the patient’s neurologic condition is asserted in the operating room prior to transfer to the recovery area.

Complications of Carotid Endarterectomy. Most patients tolerate carotid endarterectomy very well and typically are discharged home within 24 hours after surgery. Complications after endarterectomy are infrequent but can be potentially life-threatening or disabling. Acute ipsilateral stroke is a dreaded complication following carotid endarterectomy.

Cerebral ischemia can be due to either intraoperative or postoperative events. Embolizations from the occlusive plaque or prolonged cerebral ischemia are potential causes of intraoperative stroke. The most common cause of postoperative stroke is due to embolization. Less frequently, acute carotid artery occlusion can cause acute postoperative stroke. This is usually due to carotid artery thrombosis related to closure of the arteriotomy, an occluding intimal flap, or distal carotid dissection. When patients experience acute symptoms of neurologic ischemia after endarterectomy, immediate intervention may be indicated. Carotid duplex scan can be done expeditiously to assess patency of the extracranial internal carotid artery. Reexploration is mandated for acute carotid artery occlusion. Cerebral angiography can be useful if intracranial revascularization is considered.

Local complications related to surgery include excessive bleeding and cranial nerve palsies. Postoperative hematoma in the neck after carotid endarterectomy can lead to devastating airway compromise. Any expanding hematoma should be evacuated and active bleeding stopped. Securing an airway is critical and can be extremely difficult in patients with large postoperative neck hematoma. The reported incidence of postoperative cranial nerve palsies after carotid endarterectomy varies from 1% to 30%. Well-recognized injuries involve the marginal mandibular, vagus, hypoglossal, superior laryngeal, and recurrent laryngeal nerves. Often these are traction injuries but can also be due to severance of the respective nerves.

Techniques of Carotid Angioplasty and Stenting

Percutaneous carotid artery stenting has become an accepted alternative treatment in the management of patients with carotid bifurcation disease (Fig. 23-22). The perceived advantages of percutaneous carotid revascularization are related to the minimal invasiveness of the procedure compared to surgery. There are anatomic conditions based on angiographic evaluation in which carotid artery stenting should be avoided due to increased procedure-related risks (Table 23-5). In preparation for carotid stenting, the patient should be given oral clopidogrel 3 days prior to the intervention if the patient was not already taking the drug. The procedure is done in either the operating room with angiographic capabilities or in a dedicated angiography room. The patient is placed in the supine position. The patient’s blood pressure and cardiac rhythm are closely monitored.

To gain access to the carotid artery, a retrograde transfemoral approach is most commonly used as the access site for carotid intervention. Using the Seldinger technique, we insert a diagnostic 5- or 6-French sheath in the CFA. A diagnostic arch

Figure 23-21. A. An autologous or synthetic patch can be used to close the carotid arteriotomy incision, which maintains the luminal patency. B. A completion closure of carotid endarterectomy incision using a synthetic patch.
aortogram is obtained. The carotid artery to be treated is then selected using a 5-French diagnostic catheter, and contrast is injected to show the carotid anatomy. It is important to assess the contralateral carotid artery, vertebrobasilar, and intracranial circulation if these are not known based on the preoperative noninvasive studies. Once the decision is made to proceed with carotid artery stenting, with the tip of the diagnostic catheter still in the common carotid artery, a 0.035-inch, 260-cm long stiff guidewire is placed in the ipsilateral external carotid artery. Anticoagulation with intravenous bivalirudin bolus (0.75 mg/kg) followed by an infusion rate of 2.5 mg/kg per hour for the remainder of the procedure is routinely administered. Next the diagnostic catheter is withdrawn and a 90-cm, 6-French guiding sheath is advanced into the common carotid artery over the stiff guide wire. It is critical not to advance the sheath beyond the occlusive plaque in the carotid bulb. The stiff wire is then removed, and preparation is made to deploy the distal embolic protection device (EPD). Several distal EPDs are available (Table 23-6). The EPD device is carefully deployed beyond the target lesion. With regard to the carotid stents, there are several stents that have received approval from the FDA and are commercially available for carotid revascularization (Table 23-7). All current carotid stents use the rapid-exchange monorail 0.014-inch platform. The size selection is typically based on the size of common carotid artery. Predilatation using a 4-mm balloon may be necessary to allow passage of the stent delivery catheter. Once the stent is deployed across the occlusive plaque, postdilatation is usually performed using a ≤5.5-mm balloon. It is noteworthy that balloon dilation of the carotid bulb may lead to immediate bradycardia due to stimulation of the glossopharyngeal nerve. The EPD is then retrieved and the procedure is completed with removal of the sheath from the femoral artery. The puncture site is closed using available closure device or with manual compression. Throughout the procedure, the patient’s neurologic function is closely monitored. The bivalirudin infusion is

Table 23-5

Unfavorable carotid angiographic appearance in which carotid stenting should be avoided

- Extensive carotid calcification
- Polypoid or globular carotid lesions
- Severe tortuosity of the common carotid artery
- Long-segment stenoses (>2 cm in length)
- Carotid artery occlusion
- Severe intraluminal thrombus (angiographic defects)
- Extensive middle cerebral artery atherosclerosis

Figure 23-22. A. Carotid angiogram demonstrating a high-grade stenosis of the left internal carotid artery. B. Completion angiogram demonstrating a satisfactory result following a carotid stent placement.

<table>
<thead>
<tr>
<th>MECHANISM</th>
<th>NAME OF EPD</th>
<th>PORE SIZE (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distal balloon occlusion</td>
<td>PercuSurge Guard Wire,</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Export catheter (Medtronic)</td>
<td></td>
</tr>
<tr>
<td>Distal filter</td>
<td>Angioguard (Cordis)</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Accunet (Abbott)</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td>Emboshield (Abbott)</td>
<td>140</td>
</tr>
<tr>
<td></td>
<td>FilterWire (Boston Scientific)</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>SpiderRx (EV3)</td>
<td>&lt;100</td>
</tr>
<tr>
<td>Flow reversal</td>
<td>Paroli Neuro Protection (Gore)</td>
<td>NA</td>
</tr>
</tbody>
</table>

*Clinical trial (EMPIRE) in United States.
NA = not applicable.
stopped and the patient is kept on clopidogrel (75 mg daily) for at least 1 month and aspirin indefinitely.

**Complications of Carotid Stenting.** Although there have been no randomized trials comparing carotid stenting with and without EPD, the availability of EPDs appears to have reduced the risk of distal embolization and stroke. The results of the various clinical trials and registries of carotid stenting have been reported and compared. It is well known that distal embolization as detected by TCD is much more frequent with carotid stenting even with EPD, when compared with carotid endarterectomy. However, the clinical significance of the distal embolization detected by TCD is not clear because most are asymptomatic. Acute carotid stent thrombosis is rare. The incidence of in-stent carotid restenosis is not well known but is estimated at 10% to 30%. Duplex surveillance shows elevated peak systolic velocities within the stent after carotid stenting not infrequently. However, velocity criteria are being formulated to determine the severity of in-stent restenosis after carotid stenting by ultrasound duplex. It appears that systolic velocities exceeding 300 to 400 cm/s would represent >70% to 80% restenosis. Bradycardia and hypotension occur in up to 20% of patients undergoing carotid stenting. Systemic administration of atropine is usually effective in reversing the bradycardia. Other technical complications of carotid stenting are infrequent and include carotid artery dissection and access site complications, such as groin hematoma, femoral artery pseudoaneurysm, distal embolization, and acute femoral artery thrombosis.

**Nonatherosclerotic Disease of the Carotid Artery**

**Carotid Coil and Kink.** A carotid coil consists of an excessive elongation of the internal carotid artery producing tortuosity of the vessel (Fig. 23-23). Embryologically, the carotid artery is derived from the third aortic arch and dorsal aortic root and is uncoiled as the heart and great vessels descend into the mediastinum. In children, carotid coils appear to be congenital in origin. In contrast, elongation and kinking of the carotid artery in adults are associated with the loss of elasticity and an abrupt angulation of the vessel. Kinking is more common in women than men. Cerebral ischemic symptoms caused by kinks of the carotid artery are similar to those from atherosclerotic carotid lesions but are more likely due to cerebral hypoperfusion than embolic episodes. Classically, sudden head rotation, flexion, or extension can accentuate the kink and provoke ischemic symptoms. Most carotid kinks and coils are found incidentally on carotid duplex scan. However, interpretation of the Doppler frequency shifts and spectral analysis in tortuous carotid arteries can be difficult because of the uncertain angle of insonation. Cerebral angiography, with multiple views taken in neck flexion, extension, and rotation, is useful in the determination of the clinical significance of kinks and coils.

**Fibromuscular Dysplasia.** Fibromuscular dysplasia (FMD) usually involves medium-sized arteries that are long and have few branches (Fig. 23-24). Women in the fourth or fifth decade of life are more commonly affected than men. Hormonal effects on the vessel wall are thought to play a role in the pathogenesis of FMD. FMD of the carotid artery is commonly bilateral, and in about 20% of patients, the vertebral artery is also involved.37

---

**Table 23-7**

Currently approved carotid stents in the United States

<table>
<thead>
<tr>
<th>NAME OF STENT</th>
<th>MANUFACTURER</th>
<th>CELL DESIGN</th>
<th>TAPERED STENT</th>
<th>DELIVERY SYSTEM SIZE (FRENCH)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acclulink</td>
<td>Abbott</td>
<td>Open</td>
<td>Yes</td>
<td>6</td>
</tr>
<tr>
<td>Exact</td>
<td>Abbott</td>
<td>Closed</td>
<td>Yes</td>
<td>6</td>
</tr>
<tr>
<td>NexStent</td>
<td>Boston Scientific</td>
<td>Closed</td>
<td>Self-tapering</td>
<td>5</td>
</tr>
<tr>
<td>Protégé RX</td>
<td>EV3</td>
<td>Open</td>
<td>Yes</td>
<td>6</td>
</tr>
<tr>
<td>Precise RX</td>
<td>Cordis</td>
<td>Open</td>
<td>No</td>
<td>6</td>
</tr>
<tr>
<td>Exponent</td>
<td>Medtronic</td>
<td>Open</td>
<td>No</td>
<td>6</td>
</tr>
</tbody>
</table>

---

**Figure 23-23.** Excessive elongation of the carotid artery can result in carotid kinking (arrow), which can compromise cerebral blood flow and lead to cerebral ischemia.
An intracranial saccular aneurysm of the carotid siphon or middle cerebral artery can be identified in up to 50% of the patients with FMD. Four histologic types of FMD have been described in the literature. The most common type is medial fibroplasia, which may present as a focal stenosis or multiple lesions with intervening aneurysmal outpouchings. The disease involves the media with the smooth muscle being replaced by fibrous connective tissue. Commonly, mural dilations and microaneurysms can be seen with this type of FMD. Medial hyperplasia is a rare type of FMD, with the media demonstrating excessive amounts of smooth muscle. Intimal fibroplasia accounts for 5% of all cases and occurs equally in both sexes. The media and adventitia remain normal, and there is accumulation of subendothelial mesenchymal cells with a loose matrix of connective tissue causing a focal stenosis in adults. Finally, premedial dysplasia represents a type of FMD with elastic tissue accumulating between the media and adventitia. FMD can also involve the renal and external iliac arteries. It is estimated that approximately 40% of patients with FMD present with a TIA due to embolization of platelet aggregates. DSA demonstrates the characteristic “string of beads” pattern, which represents alternating segments of stenosis and dilatation. The string of beads can also be shown noninvasively by CTA or MRA. FMD should be suspected when an increased velocity is detected across a stenotic segment without associated atherosclerotic changes on carotid duplex ultrasound. Antiplatelet medication is the generally accepted therapy for asymptomatic lesions. Endovascular treatment is recommended for patients with documented lateralizing symptoms. Surgical correction is rarely indicated.

**Carotid Artery Dissection.** Dissection of the carotid artery accounts for approximately 20% of strokes in patients younger than 45 years of age. The etiology and pathogenesis of spontaneous carotid artery dissection remain incompletely understood. Arterial dissection involves hemorrhage within the media, which can extend into the subadventitial and subintimal layers. When the dissection extends into the subadventitial space, there is an increased risk of aneurysm formation. Subintimal dissections can lead to intramural clot or thrombosis. Traumatic dissection is typically a result of hyperextension of the neck during blunt trauma, neck manipulation, strangulation, or penetrating injuries to the neck. Even in supposedly spontaneous cases, a history of preceding unrecognized minor neck trauma is not uncommon. Connective disorders, such as Ehlers-Danlos syndrome, Marfan’s syndrome, α₁-antitrypsin deficiency, or FMD, may predispose to carotid artery dissection. Iatrogenic dissections can also occur due to catheter manipulation or balloon angioplasty.

Typical clinical features of carotid artery dissection include unilateral neck pain, headache, and ipsilateral Horner’s syndrome in up to 50% of patients, followed by manifestations of the cerebral or ocular ischemia and cranial nerve palsies. Neurologic deficits can result either because of hemodynamic failure (caused by luminal stenosis) or by an artery-to-artery thromboembolism. The ischemia may cause TIAs or infarctions, or both. Catheter angiography has been the method of choice to diagnose arterial dissections, but with the advent of duplex ultrasonography, MRI/MRA, and CTA, most dissections can now be diagnosed using noninvasive imaging modalities (Fig. 23-25).

The dissection typically starts in the internal carotid artery distal to the bulb. Uncommonly, the dissection can start in the common carotid artery or is an extension of a more proximal aortic dissection. Medical therapy has been the accepted primary treatment of symptomatic carotid artery dissection. Anticoagulation (heparin and warfarin) and antiplatelet therapy have been commonly used, although there have not been any randomized studies to evaluate their effectiveness. The prognosis depends on the severity of neurologic deficit but is generally good in extracranial dissections. The recurrence rate is low. Therapeutic interventions have been reserved for recurrent TIAs or strokes or failure of medical treatment. Endovascular options include intra-arterial stenting, coiling of associated pseudoaneurysms, or, more recently, deployment of covered stents.

**Figure 23-24.** A carotid fibromuscular dysplasia with typical characteristics of multiple stenoses with intervening aneurysmal outpouching dilations. The disease involves the media with the smooth muscle being replaced by fibrous connective tissue.

**Figure 23-25.** Carotid ultrasound reveals a patient with a carotid artery dissection in which carotid flow is separated in the true flow lumen (long arrow) from the false lumen (short arrow).
Carotid Artery Aneurysms. Carotid artery aneurysms are rare, encountered in less than 1% of all carotid operations (Fig. 23-26). The true carotid artery aneurysm is generally due to atherosclerosis or medial degeneration. The carotid bulb is involved in most carotid aneurysms, and bilateralism is present in 12% of the patients. Patients typically present with a pulsatile neck mass. The available data suggest that, untreated, these aneurysms lead to neurologic symptoms from embolization. Thrombosis and rupture of the carotid aneurysm are rare. Pseudoaneurysms of the carotid artery can result from injury or infection. Mycotic aneurysms often involved syphilis in the past, but they are now more commonly associated with peritonsillar abscesses caused by *Staphylococcus aureus* infection. FMD and spontaneous dissection of the carotid artery can lead to the formation of true aneurysms or pseudoaneurysms. Whereas conventional surgery has been the primary mode of treatment in the past, carotid aneurysms are currently being treated more commonly using endovascular approaches.38

Carotid Body Tumor. The carotid body originates from the third branchial arch and from neuroectodermal-derived neural crest lineage. The normal carotid body is located in the adventitia or periadventitial tissue at the bifurcation of the common carotid artery (Fig. 23-27). The gland is innervated by the glossopharyngeal nerve. Its blood supply is derived predominantly from the external carotid artery but can also come from the vertebral artery. Carotid body tumor is a rare lesion of the neuroendocrine system. Other glands of neural crest origin are seen in

Figure 23-26. **A.** An anteroposterior angiogram of the neck revealing a carotid artery aneurysm. **B.** A lateral projection of the carotid artery aneurysm. **C.** Following endovascular placement, the carotid artery aneurysm is successfully excluded.

Figure 23-27. **A.** A carotid body tumor (*arrow*) located adjacent to the carotid bulb. **B.** Following periadventitial dissection, the carotid body tumor is removed.
the neck, parapharyngeal spaces, mediastinum, retroperitoneum, and adrenal medulla. Tumors involving these structures have been referred to as paraganglioma, glomus tumor, or chemodectoma. Approximately 5% to 7% of carotid body tumors are malignant. Although chronic hypoxemia has been invoked as a stimulus for hyperplasia of carotid body, approximately 35% of carotid body tumors are hereditary. The risk of malignancy is greatest in young patients with familial tumors.

Symptoms related to the endocrine products of the carotid body tumor are rare. Patients usually present between the fifth and seventh decades of life with an asymptomatic lateral neck mass. The diagnosis of carotid body tumor requires confirmation on imaging studies. Carotid duplex scan can localize the tumor to the carotid bifurcation, but CT or MRI is usually required to further delineate the relationship of the tumor to the adjacent structures. Classically, a carotid body tumor will widen the carotid bifurcation. The Shamban classification describes the tumor extent: I, tumor is less than 5 cm and relatively free of vessel involvement; II, tumor is intimately involved but does not encase the vessel wall; and III, tumor is intramural and encases the carotid vessels and adjacent nerves. With good-resolution CT and MRI, arteriography is usually not required. However, arteriography can provide an assessment of the vessel invasion and intracranial circulation and allows for preoperative embolization of the feeder vessels, which has been reported to reduce intraoperative blood loss. Surgical resection is the recommended treatment for suspected carotid body tumor.

Carotid Trauma. Blunt or penetrating trauma to the neck can cause injury to the carotid artery. Notwithstanding the massive bleeding from carotid artery transection, injury to the carotid artery can result in carotid dissection, thrombosis, or pseudoaneurysm formation. Carotid duplex ultrasound can be useful to locate the site of injury in the cervical segment of the carotid artery. Spiral CTA has become the modality of choice to detect extracranial carotid artery injury. Confirmation of carotid injury by contrast cerebral angiography remains the gold standard diagnostic test. Injuries to the cervical segment of the common and internal carotid arteries can be repaired surgically. Acute carotid artery thrombosis is usually treated medically with anticoagulation if the patient is asymptomatic. Revascularization should be considered for patients presenting with ongoing cerebral ischemia related to carotid artery thrombosis. Traumatic carotid artery dissection can cause cerebral ischemia due to thromboembolization, decreased flow, or thrombosis. Commonly, the dissection involves the distal portion of the cervical and petrous segment of the internal carotid artery. Medical management with antiplatelet or anticoagulation therapy is usually adequate for uncomplicated traumatic carotid dissection. In patients with pseudoaneurysms of the carotid artery that are located in a segment that is out of surgical reach, the use of selective coil embolization of the pseudoaneurysm or exclusion of the pseudoaneurysm by a covered stent graft has been reported. Bare metal stent has been used with success in the treatment of traumatic carotid artery dissection.

ABDOMINAL AORTIC ANEURYSM

Despite more than 50,000 patients undergoing elective repair of abdominal aortic aneurysm (AAA) each year in the United States, approximately 15,000 patients die annually as a result of ruptured aneurysm, making it the 10th leading cause of death in men in this country. The incidence appears to be increasing, and this is due in part to improvements in diagnostic imaging and, more importantly, a growing elderly population. With early diagnosis and timely intervention, aneurysm rupture-related death is largely preventable. Conventional treatment of an AAA involves replacing the aneurysmal segment of the aorta with a prosthetic graft, with the operation performed through a large abdominal incision. Techniques for this open abdominal surgery have been refined, adapted, and extensively studied by vascular surgeons over the past four decades. Despite a well-documented low perioperative mortality rate of 2% to 3% in large academic institutions, the thought of undergoing an open abdominal aortic operation often provokes a sense of anxiety in many patients due in part to the postoperative pain associated with the large abdominal incision as well as the long recovery time needed before the patient can return to normal physical activity.

The most common location of aortic aneurysms is the infrarenal aorta. Endovascular stent graft placement represents a revolutionary and minimally invasive treatment for infrarenal AAAs that only requires 1 to 2 days of hospitalization, and the patient can return to normal physical activity within 1 week. The concept of using an endoluminal device in the management of vascular disease was first proposed by Dotter and colleagues, who successfully treated a patient with iliac occlusion using transluminal angioplasty in 1964. Nearly two decades later, Parodi and colleagues reported the first successful endovascular repair of AAA using a stent graft device. Since then, a variety of stent graft technologies have been developed to treat AAA. The rapid innovation of this new treatment modality has undoubtedly captured the attention of patients with aortic aneurysms as well as physicians who practice endovascular therapy. Physicians in general should be knowledgeable regarding available treatment options of AAA in order to provide adequate evaluation and education to patients and their families. The purpose of this section is to outline the treatment options for AAAs, including conventional repair and endovascular approach. Advantages and potential complications of these treatments will also be addressed.

Causes and Risk Factors

The pathogenesis of aneurysmal disease of the aorta is complex and multifactorial. A degenerative process in the aortic wall is the most common cause of AAA development. Matrix metalloproteinases (MMP), proteolytic enzymes, are found abundantly in the wall of AAA. Atherosclerotic disease, age, male sex, smoking history, family history, hypertension, coronary artery disease, and chronic obstructive pulmonary disease are associated with the development of AAA. Diabetes and black race have negative association with AAA. Other less common causes include inflammation, infection, and connective tissue disease. Inflammatory AAA accounts for 5% to 10% of all AAAs. In contrast to atherosclerotic AAA, the inflammatory variant is characterized pathologically by marked thickening of the aneurysm wall, fibrosis of the adjacent retroperitoneum, and rigid adherence of the adjacent structures to the anterior aneurysm wall. Male sex and smoking are even stronger risk factors in inflammatory AAA. Smoking cessation is the first step of medical therapy, followed by surgical repair. Infectious or mycotic AAA is rare but is associated with high mortality. Patients with connective tissue disorders such as Marfan’s syndrome and Ehlers-Danlos syndrome tend to have more extensive and larger aneurysms at a younger age.
Natural History of Aortic Aneurysm

The natural history of an AAA is to expand and rupture. AAA exhibits a “staccato” pattern of growth, where periods of relative quiescence may alternate with expansion. Therefore, although an individual pattern of growth cannot be predicted, average aggregate growth is approximately 3 to 4 mm/year. There is some evidence to suggest that larger aneurysms may expand faster than smaller aneurysms, but there is significant overlap between the ranges of growth rates at each strata of size.

Rupture risk appears to be directly related to aneurysm size as predicted by Laplace's Law. Although more sophisticated methods of assessing rupture risk based on finite element analysis of wall stress are under active investigation, maximum transverse diameter remains the standard method of risk assessment for aneurysm rupture. In the past, AAA rupture risk has been overestimated. More recently, two landmark studies have served to better define the natural history of AAA. Based on best available evidence, the annualized risk of rupture is given in Table 23-8. The rupture risk is quite low below 5.5 cm and begins to rise exponentially thereafter. This size can serve as an appropriate threshold for recommending elective repair provided one’s surgical mortality is below 5%. For each size strata, however, women appear to be at higher risk for rupture than men, and a lower threshold of 4.5 to 5.0 cm may be reasonable in good-risk patients. Although data are less compelling, a pattern of rapid expansion of >0.5 cm within 6 months can be considered a relative indication for elective repair. Aneurysms that fall below these indications may safely be followed with CT or ultrasound at 6-month intervals, with long-term outcomes equivalent to earlier surgical repair. Interestingly, in the Aneurysm Detection and Management (ADAM) study, 80% of all AAAs that were followed in this manner eventually came to repair within 5 years.

Unless symptomatic or ruptured, AAA repair is a prophylactic repair. The rationale for recommending repair is predicated on the assumption that the risk of aneurysm rupture exceeds the combined risk of death from all other causes such as cardiopulmonary disease and cancer. On the other hand, our limitation in predicting timing and cause of death is underscored by the observation that over 25% of patients who were deemed unfit for surgical repair because of their comorbidities died from rupture of their aneurysms within 5 years.

Clinical Manifestations

Most AAAs are asymptomatic and are usually found incidentally during workup for chronic back pain or kidney stones. Physical examination is neither sensitive nor specific except in thin patients. Large aneurysms may be missed in the obese, while normal aortic pulsations may be mistaken for an aneurysm in thin individuals. Rarely patients present with back pain and/or abdominal pain with a tender pulsatile mass. Patients with these symptoms must be treated as a rupture until proven otherwise. If the patient is hemodynamically stable and the aneurysm is intact on a CT scan, the patient is admitted for blood pressure control with intravenous antihypertensive agents and undergoes repair usually within 12 to 24 hours or at least during the same hospitalization. In contrast, patients who are hemodynamically unstable with a history of acute back pain and/or syncopal and a known unrepaired AAA or a pulsatile abdominal mass should be immediately taken to the operating room with a presumed diagnosis of a ruptured AAA.

Overall mortality of AAA rupture is 71% to 77%, which includes all out-of-hospital and in-hospital deaths, as compared with 2% to 6% for elective open surgical repair. Nearly half of all patients with ruptured AAA will die before reaching the hospital. For the remainder, surgical mortality is 45% to 50% and has not substantially changed in the last 30 years.

Relevant Anatomy

An AAA is defined as a pathologic focal dilation of the aorta that is greater than 30 mm or 1.5 times the adjacent diameter of the normal aorta (Fig. 23-28). Male aortas tend to be larger than female aortas, and there is generalized growth of the aortic diameter with each decade of life. Ninety percent of AAAs are infrarenal in location and have a fusiform morphology. There is a higher predilection for juxtarenal and suprarenal AAAs in women compared with men. Concomitant common iliac and/or hypogastric artery aneurysms can be found in 20% to 25% of patients. Although the etiology of most aortic aneurysms is atherosclerotic, clinically significant peripheral occlusive disease is unusual and present in less than 10% of all cases.

Although extravascular anatomy is important for open surgical repair of AAA, intravascular anatomy and aortoiliac morphology are important for endovascular repair. Pertinent anatomic dimensions include the diameter of the proximal nondilated infrarenal aortic neck, which can range from 18 to 30 mm; common iliac artery, which can range from 8 to 16 mm; and external iliac arteries, which can range from 6 to 10 mm. Morphologically, the aortic neck can manifest complex angulation above and below the renal arteries due to combination of elongation and anterolateral displacement by the posterior bulge of the aneurysmal aorta. Furthermore, the shape of the proximal neck is rarely tubular, but often is conical, reverse conical, or barrel-shaped. Distally, the iliac arteries can have severe

Table 23-8

<table>
<thead>
<tr>
<th>DESCRIPTION</th>
<th>DIAMETER OF AORTA (CM)</th>
<th>ESTIMATED ANNUAL RISK OF RUPTURE (%)</th>
<th>ESTIMATED 5-YEAR RISK OF RUPTURE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal aorta</td>
<td>2–3</td>
<td>0</td>
<td>0 (unless AAA develops)</td>
</tr>
<tr>
<td>Small AAA</td>
<td>4–5</td>
<td>1</td>
<td>5–10</td>
</tr>
<tr>
<td>Moderate AAA</td>
<td>5–6</td>
<td>2–5</td>
<td>30–40</td>
</tr>
<tr>
<td>Large AAA</td>
<td>6–7</td>
<td>3–10</td>
<td>&gt;50</td>
</tr>
<tr>
<td>Very large AAA</td>
<td>&gt;7</td>
<td>&gt;10</td>
<td>Approaching 100</td>
</tr>
</tbody>
</table>

*The estimated 5-year risk is more than five times the estimated annual risk because over that 5 years, the AAA, if left untreated, will continue to grow in size.
tortuosity with multiple compound turns. Although significant from hemodynamic standpoint, severe iliac calcifications combined with extreme tortuosity can pose a formidable challenge during endovascular repair.

**Diagnostic Evaluation**

Preoperative evaluation should include routine history and physical exam with particular attention to (a) any symptoms referable to the aneurysm, which may impact the timing of repair; (b) history of pelvic surgery or radiation, in the event retroperitoneal exposure is required or interruption of hypogastric circulation is planned; (c) claudication suggestive of significant iliac occlusive disease; (d) lower extremity bypass or other femoral reconstructive procedures; and (e) chronic renal insufficiency or contrast allergy.

Cross-sectional imaging is required for definitive evaluation of AAA. Although ultrasound is safe, widely available, relatively accurate, and inexpensive and thus the screening modality of choice, CT scan remains the gold standard for determination of anatomic eligibility for endovascular repair. Size of AAA may differ up to 1 cm between CT and ultrasound, and during longitudinal follow-up, comparisons should be made between identical modalities. With modern multirow detector scanners, a timed-bolus intravenous contrast-enhanced, 2.5- to 3.0-mm slice spiral CT of the chest, abdomen, and pelvis can be performed in less than 30 seconds with a single breath hold. Extremely high-resolution images are obtained with submillimeter spatial resolution (Fig. 23-29). Proper window level and width (brightness and contrast) are important for discrimination among aortic wall, calcific plaque, thrombus, and lumen. The only major drawback to CT is the risk of contrast nephropathy in diabetics and in patients with renal insufficiency.

The spiral technique further affords the ability for three-dimensional reconstruction. Three-dimensional reconstructions can yield important morphologic information that is critical to endovascular therapy. Using third-party software, these images can be viewed and manipulated on one’s desktop computer, and so-called “center-line” (transverse slices perpendicular to the central flow lumen of the aorta) diameter and length measurements obtained. Conventional angiography has a minimal role in the current management of AAA. Angiography is invasive with an increased risk of complications. Indications for angiography are isolated to concomitant iliac occlusive disease (present in <10% of patients with AAA) and unusual renovascular anatomy.

**Surgical Repair of Abdominal Aortic Aneurysm**

General anesthesia is necessary when performing a conventional open AAA repair. While a retroperitoneal incision is a well-accepted surgical approach, a midline transabdominal incision remains the more common approach for open aortic aneurysm operation. Since the abdominal incision can lead to significant pain and discomfort, an epidural catheter can be placed prior to the operation for postoperative analgesic infusion to provide pain control. Once the abdominal cavity is opened, the small intestines and transverse colon are retracted to expose the retroperitoneum overlying the AAA. The retroperitoneum is next divided, followed by isolation of both proximal and distal segments of the AAA. Intravenous heparin (100 IU/kg) is given followed by clamping of the proximal and distal segments of the aneurysm. The aneurysm sac is open next, and a prosthetic graft is used to reconstruct the aorta. If the aneurysm only involves the abdominal aorta, a tube graft can be used to replace the aorta (Fig. 23-30). If the aneurysm extends distally to the iliac arteries, a prosthetic bifurcated graft is used for either an aorto-bi-iliac or aorto-bi-femoral bypass reconstruction (Fig. 23-31). The overlying aneurysm sac and the retroperitoneum are closed to cover the prosthetic bypass graft to minimize potential bowel contact to the graft. Small and large intestines are returned to the abdominal cavity followed by the closure of the abdominal fascia and skin.
Advantages and Risks of Open Abdominal Aortic Aneurysm Repair. The main advantage of a conventional open repair is that the AAA is permanently eliminated because it is entirely replaced by a prosthetic aortic graft. The risk of aneurysm recurrence or delayed rupture no longer exists. As a result, long-term imaging surveillance is not needed with these patients. In contrast, the long-term efficacy of endovascular repair remains unclear. Consequently, long-term imaging surveillance is critical to ensure that the aortic aneurysm remains properly sealed by the stent graft. Other potential advantages of open repair include direct assessment of the circulatory integrity of the colon. If signs of colonic ischemia become evident after aortic bypass grafting, a concomitant mesenteric artery bypass can be performed to revascularize the colonic circulation. In addition, open repair permits the surgeons to explore for other abdominal pathologies, such as gastrointestinal tumors, liver mass, or cholelithiasis.

As for the risks associated with open repair, cardiac complications, in the form of either myocardial infarction or arrhythmias, remain the most common morbidity, with an incidence between 2% and 6%.51 Another significant complication is renal failure or transient renal insufficiency as a result of perioperative hypotension, atheromatous embolization, inadvertent injury to the ureter, preoperative contrast-induced nephropathy, or suprarenal aortic clamping. Although the incidence of renal failure is less than 2% in elective aneurysm repair, it can occur in more than 20% of patients after repair of a ruptured AAA.49

Ischemic colitis is a devastating potential complication after open repair. The likelihood of such a complication is highest in those who had a prior colon resection and undergo repair of a ruptured AAA, due to the loss of collateral blood supply to the rectosigmoid colon. It is estimated that 5% of patients who undergo elective aneurysm repair will develop partial-thickness ischemic colitis but without significant clinical sequelae.52 However, if the partial-thickness ischemia progresses to full-thickness gangrene and peritonitis, mortality can be as high as 90%.52

The incidence of prosthetic graft infection ranges between 1% and 4% after open repair.53 It is more common in those who undergo repair of a ruptured AAA. If the prosthetic graft is not fully covered by the aneurysm sac or retroperitoneum, intestinal adhesion with subsequent bowel erosion may occur, resulting in an aortoenteric fistula. The predominant sign of such a complication is massive hematemesis, and it typically occurs years after the operation. Despite these potential complications, however, the majority of patients who undergo successful elective open repair have an uneventful recovery.

Endovascular Repair of Abdominal Aortic Aneurysm
Over a decade has passed since the first report of human implantation of a homemade stent graft for endovascular repair of an AAA by Parodi in 1991.11 Several prospective clinical trials across different devices and analysis of large Medicare administrative databases and meta-analyses of published literature have consistently demonstrated significantly decreased operative time, blood loss, hospital length of stay, and overall perioperative morbidity and mortality of endovascular repair compared
sac is not resected, which is subjected for potential aneurysm expansion or even rupture. Importantly, aortic branches, such as lumbar arteries or the inferior mesenteric artery (IMA), are occluded, which can lead to persistent aneurysm pressurization and aneurysm expansion. Currently, there are more than fifteen different endovascular devices approved for clinical use for infrarenal aortic aneurysm implantation throughout the world. Despite some differences in physical appearance, mechanical properties, and endograft materials, these endovascular devices will be discussed collectively for this chapter. Most of these devices are modular devices consisting of a primary device or main body and one or two iliac limbs that insert into the main body to complete the repair. Depending on the device, there are varying degrees of flexibility in the choice of iliac limbs that can be matched to the main body, which can impact the customizability for a particular anatomy.

A severe limitation of the endovascular repair devices is the need for adequate proximal neck to achieve a durable sealing zone. Several techniques have been proposed to overcome this limitation. These include fenestrated or branched endografts and the “chimney,” “snorkel,” and “periscope” techniques. The fenestrated stent grafts rely on precise alignment between the fenestration and the corresponding visceral artery. Multiple clinical trials using customized fenestrated stent graft for the treatment of short-necked and juxtarenal aortic aneurysm repair have shown promising short- and mid-term results. However, fenestrated stent graft generally requires device customization which is accessible only to high volume tertiary institutions, and not widely available to all hospital facilities. Alternatively, some centers have reported good results with intraoperative surgeon-modified endograft to create fenestrations for the treatment of complex aortic aneurysms in high-risk patients. Further development of the fenestrated techniques also opens the way for endovascular treatment of suprarenal and thoracoabdominal aneurysm. The review of literature showed that open surgery remains a safe and effective treatment option for good-risk patients with juxtarenal aortic aneurysm. Fenestrated endovascular repair is associated with low mortality and compares favorably with open surgery in terms of morbidity, especially renal function impairment and cardiac complications.

**Figure 23-31.** Intraoperative view of a bifurcated graft used to repair an aortic aneurysm.

with open surgical repair. For patients who are at increased risk for surgery because of age or comorbidity, endovascular repair is a superior minimally invasive alternative.

The principle of endovascular repair of AAA involves the implantation of an aortic stent graft that is fixed proximally and distally to nonaneurysmal aortoiliac segment and thereby endoluminally excluding the aneurysm from the aortic circulation (Fig. 23-32). Unlike open surgical repair, the aneurysm

**Figure 23-32.** A. An aortogram demonstrating a large infrarenal abdominal aortic aneurysm. B. Following endovascular stent graft implantation, the aortic aneurysm is successfully excluded.
Patient Selection for Endovascular Aortic Aneurysm Repair. Anatomic eligibility for endovascular repair is mainly based on three areas: the proximal aortic neck, common iliac arteries, and external iliac and common femoral arteries, which relate to the proximal and distal landing zones or fixation sites and the access vessels, respectively. The requirements for the proximal aortic neck are a diameter of 18 to 28 mm and a minimum length of 15 mm (Table 23-9). Usually, multiple measurements of the diameter are taken along the length of the neck to assess its shape. All diameter measurements are mid-wall to mid-wall of the vessel. Secondary considerations include mural calcifications (<50% circumference), luminal thrombus (<50% circumference), and angulation (<45°). Presence of a significant amount of any one of these secondary features in combination with a relatively short proximal neck may compromise successful short- and long-term fixation of the stent graft and exclusion of the aneurysm. The usual distal landing zone is the common iliac artery. The external iliac artery may serve as an alternate site when the ipsilateral common iliac artery is aneurysmal or ectatic. The treatable diameters of common iliac arteries range from 8 to 20 mm, and there should be at least 20 mm of patent artery of uniform diameter to allow adequate fixation. Finally, at least one of two common femoral and external iliac arteries must be at least 7 mm in diameter in order to safely introduce the main delivery sheath. Slightly smaller iliac diameters may be tolerated depending on the specific device and in the absence of severe tortuosity and calcific disease. Difficult access is one of the main causes of increased procedural time and intraoperative complications. Using these criteria, approximately 60% of all AAAs are anatomic candidates for endovascular repair.

The next step in the preoperative planning is device selection. Typically, the proximal diameter of the main device is oversized by 10% to 20% of the nominal diameter of the aortic neck. Distally, the iliac limbs are oversized by 1 to 4 mm depending on the individual device’s instructions for use. The biggest challenge to proper device selection remains determining the optimal length from the renal arteries to the hypogastric arteries. Despite availability of sophisticated three-dimensional reconstructions, the exact path that a device will take from the proximal aortic neck to the distal iliac arteries is difficult to predict. It is dependent on a host of factors related to the mechanical properties of the stent graft and the morphology of the aortoiliac flow lumen. “Plumb-line” measurements of axial CT images can be quite inaccurate, typically grossly underestimating the length, whereas center-line measurements usually overestimate the length. angiographic measurements using a marker catheter are invasive, require contrast and radiation exposure, and are also inaccurate because they fail to account for the stiffness of the stent graft. The consequences of not choosing the correct length of the device include inadvertent coverage of the hypogastric artery if too long and the need for additional devices if too short.

Advantages and Risks of Endovascular Repair. The obvious advantage of an endovascular AAA repair is its minimally invasive nature. Typically, patients who undergo this procedure stay in the hospital for only 1 to 3 days, in contrast to the 5- to 10-day stay required after conventional open surgical repair. In our institution, patients who have had an endovascular repair are routinely transferred to a general vascular ward from the postanesthesia recovery unit, avoiding admission to a more costly intensive care unit.

Because an abdominal incision is not necessary in endovascular repair, the procedure is particularly beneficial in patients with severe pulmonary disease, such as chronic obstructive pulmonary disease or emphysema. Patients can sustain adequate breathing in the postoperative period, avoiding respiratory complications or prolonged mechanical ventilation. Because the abdominal cavity has not been entered, the risk of gastrointestinal complications, such as ileus, ventral hernia, or bowel obstruction due to intestinal adhesion, is also greatly reduced. Moreover, regional or epidural anesthesia can be used, avoiding the risks associated with general anesthesia in patients with severe cardiopulmonary dysfunction.

Despite its many advantages, endovascular repair does have potential complications. Since the stent graft device is attached endoluminally within the abdominal aorta, an endoleak due to incomplete stent graft exclusion of the aneurysm can occur. With this type of leak, blood flow persists outside the lumen of the endoluminal graft but within an aneurysm sac. A meta-analysis of 1118 patients who underwent successful endovascular repair found an endoleak incidence of 24%. Although a small endoleak usually poses little clinical significance because it will typically become thrombosed spontaneously, a large or persistent endoleak may lead to continuous aneurysm perfusion and ultimately to aneurysm rupture. The rupture rate following an endovascular AAA repair has been reported to be less than 0.8%.

Stent graft iliac limb dysfunction resulting in thrombosis has been reported following endovascular repair. One possible cause is aneurysm remodeling, resulting in a shortening in the aortic length, which can cause the stent graft to kink. Alternatively, progression of an underlying iliac atherosclerotic lesion may cause compression of the iliac limb and ultimately result in graft-limb occlusion. Treatment options include thrombolysis or graft thrombectomy to determine the underlying cause and possibly additional stent graft placement. Renal artery occlusion may occur due to improper stent graft positioning or migration.82

Graft limb separation or dislocation has also been reported.82

In patients with AAA and concurrent iliac artery aneurysms who undergo preoperative coil embolization of the internal iliac artery, 20% to 45% experience symptoms of pelvic ischemia.83 These symptoms may include buttock claudication, impotence, gluteal skin sloughing, and colonic ischemia. Other complications pertaining to endovascular repair relate to the access site and include groin hematoma and wound infection.

<table>
<thead>
<tr>
<th>Table 23-9</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ideal characteristics of an aneurysm for endovascular abdominal aortic aneurysm repair</strong></td>
</tr>
<tr>
<td>Neck length (mm)</td>
</tr>
<tr>
<td>Neck diameter (mm)</td>
</tr>
<tr>
<td>Aortic Neck angle (degrees)</td>
</tr>
<tr>
<td>Neck mural calcification (% circumference)</td>
</tr>
<tr>
<td>Neck luminal thrombus (% circumference)</td>
</tr>
<tr>
<td>Common iliac artery diameter (mm)</td>
</tr>
<tr>
<td>Common iliac artery length (mm)</td>
</tr>
<tr>
<td>External iliac artery diameter (mm)</td>
</tr>
</tbody>
</table>
Occasionally, the stent graft device can malfunction by either failing to deploy or dislodging during the deployment procedure. If the device cannot be salvaged or rescued endoluminally, open surgical repair of the aneurysm may be necessary.

**Technical Considerations of Endovascular Aortic Aneurysm Repair.** Although endovascular AAA repair may be performed in any venue with appropriate digital fluoroscopic imaging capability, due to the need for absolute sterility and aseptic technique, it is most safely performed in a surgical suite. The patient is prepped and draped just as in open AAA repair. Patients with renal insufficiency should be started on perioperative oral N-acetylcysteine (Mucomyst) and sodium bicarbonate infusion to reduce the risk of contrast nephropathy. A variety of anesthetic options may be used. Regional anesthesia may be appropriate for patients with pulmonary disease. There are reports of success with local anesthetics alone, as the incisions are typically smaller than a typical open inguinal hernia repair.64

Groin access for endovascular aortic aneurysm repair can be achieved by either surgical cutdown for femoral artery exposure or percutaneous approach using “preclose” technique with the Perclose suture-mediated vascular closure device (Abbott Perclose, Redwood City, CA). Review of reported series on this percutaneous technique suggest a technical success rate of 95% for medium-size sheaths ranging from 12 to 16 French, and 87% success for 18- to 24-French sizes.65 Once femoral artery access is obtained followed by introducer sheath placement, initial soft-tipped starter guidewires are exchanged for stiff guidewires that are advanced to the thoracic arch. Intravenous heparin at 80 IU/kg are administered, and the activated clotting time is maintained at 200 to 250 seconds. These guidewires provide the necessary support for the subsequent introduction of the large-diameter delivery catheters and devices. In the absence of special anatomic considerations, the primary device is inserted through the right side and the contralateral iliac limb is inserted through the left side. After administration of heparin, the delivery catheter or the introducer sheath is advanced to the L1-L2 vertebral space, which typically marks the location of the renal arteries. An angiographic catheter is advanced from the contralateral femoral artery to the same level.

A road-mapping aortogram is obtained to localize the renal arteries. The primary device is rotated to the desired orientation and deployed immediately below the lowest renal artery (Fig. 23-33). The angiographic catheter is replaced with a directional catheter and an angled guidewire, and the opening for the contralateral limb on the main device is cannulated. Intrastent passage of the guidewire is confirmed, and the angled guidewire is replaced with a stiff guidewire. The contralateral iliac limb is inserted into the docking opening of the primary device and deployed. A completion angiogram is performed looking for patency of the renal and hypogastric arteries, the device limbs, proximal and distal fixation, and endoleak. Adjunctive interventions including additional devices, balloons, and bare stents are performed as needed. The procedure is concluded with routine repairs of the femoral arteries and closure of the groin incisions. The patients recover in the recovery room for 2 to 4 hours and admitted to the general care floor. Although in the past, patients were admitted to the intensive care unit, this is rarely needed. Most patients can be started on a regular diet that evening and discharged the next morning.

**Surveillance Following Endovascular Aortic Aneurysm Repair.** Life-long follow-up is essential to the long-term success after endovascular AAA repair. Indeed, one may go so far as to say that absence of appropriate follow-up is tantamount to not having had a repair at all. A triple-phase (noncontrast, contrast, and delayed) spiral CT scan and a four-view (anteroposterior, lateral, and two obliques) abdominal X-ray should be obtained within the first month. Subsequent imaging can be obtained at 6-month intervals in the first 1 to 2 years and yearly thereafter. After the first 6 months, patients who cannot travel easily may obtain their studies locally and submit them for review. The CT scan is for detection of endoleaks, subtle proximal migrations, and changes in aneurysm size. The abdominal X-ray gives a “birds-eye” view of the overall morphology of the stent graft. Subtle changes in conformation of the iliac limbs relative to each other and/or the spine can provide early signs of impending component separation or loss of fixation. Further, stent fractures and/or suture breaks that can compromise long-term device integrity can sometimes only be detected on a plain film and not on a CT scan.

**Results From Clinical Studies Comparing Endovascular Versus Open Repair**

The primary success rate after endovascular repair of AAA has been reported to be as high as 95%.41 The less invasive nature
SPECIFIC CONSIDERATIONS

by Mary
perioperative repair both endoluminal and endovascular repair. Randomization is also difficult because most patients who anatomically qualify for endovascular repair would withdraw from the study if randomized to open repair. Consequently, there are very few randomized controlled trials that have compared outcomes in patients with similar risk factors and anatomy who are eligible for both types of repair. Two such European trials have recently published short-term outcome data that are unbiased in design.

The DREAM trial is a multicenter randomized trial that compared open versus endovascular repair among a group of 345 patients at 28 European centers using multiple different devices including Gore, AneurRx, and Zenith. Patients were included only if they were considered to be candidates for both types of repairs. The operative mortality rate was 4.6% in the operative group versus 1.2% in the endoluminal group at 30 days. When looking at the combined rate of operative mortality and severe complications, there was an incidence of 9.8% in the open repair group versus 4.7% in the endoluminal group. The difference here was largely due to the higher frequency of pulmonary complications seen in the open group. There was a higher incidence of graft-related complications in the endoluminal group. There was no difference in the nonvascular local complication rate among the two groups. The Endovascular Repair-1 (EVAR-1) trial is also a multicenter randomized trial that compared open to endoluminal repair. This study was conducted on 1082 patients at 34 centers in the United Kingdom using all available devices. Short-term mortality at 30 days was 4.7% in the open group and 1.7% in the endoluminal group. The in-hospital mortality rate was also increased in the open group compared to the endoluminal group (6.2% vs. 2.1%). As expected, the secondary intervention rate was higher in the endoluminal group (9.8% vs. 5.8%). Complication rates were not reported in the EVAR-1 trial. Criticisms can be applied to both of these trials. Patients had to be eligible for either type of repair in order to be included in the study. Consequently, these findings cannot be generalized to patients who are too sick to undergo open surgery or to patients whose anatomy precludes them from undergoing endovascular repair.

The Open Versus Endovascular Repair (OVER) Veterans Affairs Cooperative Study Group randomly assigned 881 patients with asymptomatic AAAs who were candidates for both procedures to either endovascular repair (n = 444) or open repair (n = 437) and followed them for up to 9 years. Reduction in perioperative mortality with endovascular repair was sustained at 3 years but not thereafter. There was no difference in primary outcome of all-cause mortality. Endovascular repair and open repair resulted in similar long-term survival. Six aneurysm ruptures were confirmed in the endovascular repair group versus none in the open repair group. Rupture after endovascular repair remains a concern. A significant interaction was observed between age and type of treatment. Endovascular repair led to increased long-term survival among younger patients but not among older patients, for whom a greater benefit from the endovascular approach had been expected.

Device-Specific Outcome. Matsumura and associates compared endoluminal versus open repair using the Excluder device. In their review, they demonstrated a 30-day mortality rate of 1% along with endoleak rates of 17% and 20% at 1- and 2-year intervals, respectively. The limb narrowing, limb migration, and trunk migration were all 1% at 2 years. There were no deployment failures or early conversions. There was an annual 7% reintervention rate. Aneurysm growth was demonstrated in 14% of patients at 2 years. The Zenith device by Cook has been studied by Greenberg and associates, who compared standard surgical repair with endoluminal repair in low-risk patients and endoluminal repair in high-risk patients. They reported a 30-day mortality rate of 3.5%, which was equal to the open group. The endoleak rates were 7.4% and 5.4% at 1- and 2-year intervals, respectively. There was a 5.3% migration of 5 mm at 1 year. Freedom from rupture was 100% in the low-risk group and 98.9% in the high-risk endoluminal group at 2 years. Experience with the AneurRx device has been reported by Zarins. In this 4-year review, they found a 30-day mortality rate of 2.8%. Endoleak rate at 4 years was 13.9%, aneurysm enlargement was 11.5%, and stent graft migration was 9.5%. Freedom from rupture was noted to be 98.4% at 4 years. Criado and associates have reported on their 1-year experience with the Talent LPS device by Medtronic. They report a 30-day mortality rate of 0.8%. Endoleak rate was 10%. Three deployment failures were noted, and freedom from rupture was 100%. Aneurysm growth and migration rates were divided into three different neck size groups. Patients with a wide neck (>26 mm) had a 3% growth and migration rate. Narrow-neck patients (<26 mm) had a 1% growth rate and a 2% migration neck. Interestingly, short-neck patients (<15 mm) had no aneurysm growths and a 2% migration rate.

Cost Analysis. The current climate of cost containment and limited reimbursement for healthcare services mandates a critical analysis of the economic impact of any new medical technology on the market. The in-hospital costs for both endovascular and open repair include graft cost, operating room fees, radiology, pharmacy, ancillary care, intensive care unit charges, and floor charges. Despite the improved morbidity and mortality rates, several early studies have reported no cost benefit with the application of endovascular repair. The limiting factor appears to be the cost of the device. Despite commercialization of endovascular repair, the device costs are still in the range of $5000 to $6000 with no signs of abating. A report by Angle and associated further corroborates previous studies. In their review, despite decreased hospital and intensive care unit stays and utilization of pharmacy and respiratory services, cost of endovascular repair was 1.74 times greater than the standard surgical approach. In addition, these cost analysis studies are centered on in-hospital costs and do not even begin to address secondary costs such as postoperative surveillance that is required with endovascular repair. In the OVER trial, endovascular repair was found to be a cost-effective alternative to...
open repair in the U.S. Veterans Affairs healthcare system for at least the first 2 years. The primary outcomes were mean total healthcare cost per life-year and per quality-adjusted life-year. There were no differences found in survival, quality of life, and costs after 2 years between the endovascular and the open group. Although graft costs were higher in the endovascular group, length of stay was shorter, resulting in lower cost of AAA repair hospitalization in the endovascular group. Costs remained lower after 2 years in the endovascular group, but the difference was no longer significant.

**Classification and Management of Endoleak**

An endoleak is an extravasation of contrast outside the stent graft and within the aneurysm sac (Fig. 23-34). It can be present in up to 20% to 30% of all endovascular AAA repairs in the early postoperative period. In general, over half of these endoleaks will resolve spontaneously during the first 6 months, resulting in a 10% incidence of chronic endoleaks in all cases beyond the first year of follow-up. Endoleaks can be detected using conventional angiography, contrast CT (Fig. 23-35), MRA, and color-flow duplex ultrasound. Although there is no recognized gold standard, in practice, angiography is considered the least sensitive but most specific for characterizing the source of the endoleak, whereas the CT scan is the most sensitive but least specific. Widespread availability and reliability that is relatively independent of technique have made the CT scan the de facto standard imaging modality for postoperative surveillance. Conversely, routine use of duplex ultrasound and MRA has been limited by the lack of proper equipment and local expertise. On the other hand, investigational techniques such as time-resolved MRA may provide greater sensitivity and specificity than either angiography or CT in the future.

Four types of endoleaks have been described (Table 23-10). Type I endoleak refers to fixation-related leaks that occur at the proximal or distal attachment sites. These represent less than 5% of all endoleaks and are seen as an early blush of contrast into the aneurysm sac from the proximal or distal ends of the device during completion angiography. Although seen as marker of poor patient selection or inadequate repair, over 80% of these leaks spontaneouly seal in the first 6 months. Persistent type I endoleaks, on the other hand, require prompt treatment. Type II endoleak refers to retrograde flow originating from a lumbar, inferior mesenteric, accessory renal, or hypogastric artery. They are the most common type of endoleak, accounting for 20% to 30% of all cases, and about half resolve spontaneously. On angiography, they are seen as a late filling of the aneurysm sac from a branch vessel(s). Type II endoleaks carry a relatively benign natural history and do not merit intervention unless associated with aneurysm growth. Type III endoleaks refer to failure of device integrity or component separation from modular systems. If detected intraoperatively or in the early perioperative period, it is usually from inadequate overlap between two stent

**Table 23-10**

<table>
<thead>
<tr>
<th>Endoleak classification</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I endoleak</td>
<td>Attachment site leak</td>
</tr>
<tr>
<td>Type II endoleak</td>
<td>Side branch leak caused by lumbar or inferior mesenteric arteries</td>
</tr>
<tr>
<td>Type III endoleak</td>
<td>Junctional leak (of overlapping endograft components) and graft fabric defect</td>
</tr>
<tr>
<td>Type IV endoleak</td>
<td>Endograft fabric porosity leak</td>
</tr>
</tbody>
</table>
grants, whereas in the late period, the endoleak may be from a fabric tear or junctional separation from conformational changes of the aneurysm. Regardless of the etiology or timing, these should be promptly repaired. Finally, type IV endoleak refers to the diffuse, early blush seen during completion angiography due to graft porosity and/or suture holes of some Dacron-based devices. It does not have any clinical significance and usually cannot be seen after 48 hours and heparin reversal. Endoleaks that are initially considered type IV but persist become type III endoleaks by definition because this indicates a more significant material defect than simple porosity or a suture hole.

**Endotension Following Endovascular Aortic Aneurysm Repair.** In approximately 5% of cases after an apparently successful endovascular repair, the aneurysm continues to grow without any demonstrable endoleak. This phenomenon has been described as endotension. Although it was initially thought that an endoleak was really present but simply not detected, case have been reported where the aneurysm has been surgically opened and the contents were completely devoid of any blood and no extravasation could be found. The mechanism of continued pressurization of the aneurysm sac following successful exclusion from the arterial circulation remains unsolved at this time. One putative mechanism has been linked to a transudative process related to the exudation of PTFE graft materials. More importantly, however, the natural history of these enlarging aneurysms without endoleaks is unknown, but to date, there has been no evidence to suggest that they carry an increased risk of rupture. Conservatively speaking, until further long-term data become available, if the patient is a suitable surgical risk, elective open conversion should be considered.

**Secondary Interventions Following Endovascular Aortic Aneurysm Repair.** There is approximately 10% to 15% per year risk of secondary interventions following endovascular AAA repair. These procedures are critical in the long-term success of the primary procedure in prevention of aneurysm rupture and aneurysm-related death. These secondary procedures, in order of frequency, include proximal or distal extender placement for migrations, highly selective or translumbar embolization for type II endoleaks, direct surgical or laparoscopic branch vessel ligations, bridging cuffs for component separations, and late open surgical conversions.

Multiple large series have reported that an annual rupture rate of approximately 1% to 1.5% per year after endovascular repair. The EUROSTAR registry reports a rupture rate of 2.3% over 15.4 months in patients with an endoleak, compared with 0.3% in those without. Various causes of late ruptures have been reported in the literature, although presence of a persistent endoleak with aneurysm enlargement remains a common culprit for this complication. It has been shown that even successfully excluded aneurysms can lead to the development of attachment-site leaks and device failure, caused in part by aneurysm remodeling resulting in stent migration or kinking. Mehta and colleagues reported that 63% of delayed AAA ruptures after endovascular repair were caused by type I endoleaks with endograft migration, 11% by type I without migration, 19% by type II, and the rest of unknown type.

Treatment of rupture may be open conversion or endovascular stent graft placement. May and associates reported a mortality rate of 43% in those patients who underwent open conversion. Emergent endovascular repair should be considered in these patients since it is potentially much faster and less likely to cause physiologic stress than open conversion. Several reports have shown that endovascular repair can be performed successfully in patients previously treated with endoluminal prostheses.

### MESENTERIC ARTERY DISEASE

Vascular occlusive disease of the mesenteric vessels is a relatively uncommon but potentially devastating condition that generally presents in patients over 60 years of age, is three times more frequent in women, and has been recognized as an entity since 1936. The incidence of such a disease is low and represents 2% of the revascularization operations for atherosomatous lesions. The most common cause of mesenteric ischemia is atherosclerotic vascular disease. Autopsy studies have demonstrated splanchic atherosclerosis in 35% to 70% of cases. Other etiologies exist and include FMD, panarteritis nodosa, arteritis, and celiac artery compression from a median arcuate ligament, but they are unusual and have an incidence of one in nine compared with that of atherosclerosis.

Chronic mesenteric ischemia is related to a lack of blood supply in the splanchic region and is caused by disease in one or more visceral arteries: the celiac trunk, the superior mesenteric artery, and the IMA. Mesenteric ischemia is thought to occur when two of the three visceral vessels are affected with severe stenosis or occlusion; however, in as many as 9% of cases, only a single vessel is involved (SMA in 5% and celiac trunk in 4% of cases). This disease process may evolve in a chronic fashion, as in the case of progressive luminal obliteration due to atherosclerosis. On the other hand, mesenteric ischemia can occur suddenly, as in the case of thromboembolism. Despite recent progress in perioperative management and better understanding of pathophysiology, mesenteric ischemia is considered one of the most catastrophic vascular disorders with mortality rates ranging from 50% to 75%. Delays in diagnosis and treatment are the main contributing factors in its high mortality. It is estimated that mesenteric ischemia accounts for 1 in every 1000 hospital admissions in this country. The prevalence is rising due to the increased awareness of this disease, the advanced age of the population, and the significant comorbidity of these elderly patients. Early recognition and prompt treatment before the onset of irreversible intestinal ischemia are essential to improve the outcome.

### Anatomy and Pathophysiology

Mesenteric arterial circulation is remarkable for its rich collateral network. Three main mesenteric arteries provide the arterial perfusion to the gastrointestinal system: the celiac artery (CA), the superior mesenteric artery (SMA), and the IMA. In general, the CA provides arterial circulation to the foregut (distal esophagus to duodenum), hepatobiliary system, and spleen; the SMA supplies the midgut (jejunum to mid-colon); and the IMA supplies the hindgut (mid-colon to rectum). The CA and SMA arise from the ventral surface of the infradiaphragmatic suprarenal abdominal aorta, whereas the IMA originates from the left lateral portion of the infrarenal aorta. These anatomic origins in relation to the aorta are important when a mesenteric angiogram is performed to determine the luminal patency. In order to fully visualize the origins of the CA and SMA, it is necessary to perform both an anteroposterior and a lateral projection of the aorta since most arterial occlusive lesions occur in the proximal segments of these mesenteric trunks.
Because of the abundant collateral flow between these mesenteric arteries, progressive diminution of flow in one or even two of the main mesenteric trunks is usually tolerated, provided that uninvolved mesenteric branches can enlarge over time to provide sufficient compensatory collateral flow. In contrast, acute occlusion of a main mesenteric trunk may result in profound ischemia due to lack of sufficient collateral flow. Collateral networks between the CA and the SMA exist primarily through the superior and inferior pancreaticoduodenal arteries. The IMA may provide collateral arterial flow to the SMA through the marginal artery of Drummond, the arc of Riolan, and other unnamed retroperitoneal collateral vessels termed meandering mesenteric arteries (Fig. 23-36). Lastly, collateral visceral vessels may provide important arterial flow to the IMA and the hindgut through the hypogastric arteries and the hemorrhoidal arterial network.

Regulation of mesenteric blood flow is largely modulated by both hormonal and neural stimuli, which characteristically regulate systemic blood flow. In addition, the mesenteric circulation responds to the gastrointestinal contents. Hormonal regulation is mediated by splanchnic vasodilators, such as nitric oxide, glucagon, and vasoactive intestinal peptide. Certain intrinsic vasoconstrictors, such as vasopressin, can diminish the mesenteric blood flow. On the other hand, neural regulation is provided by the extensive visceral autonomic innervation.

Clinical manifestation of mesenteric ischemia is predominantly postprandial abdominal pain, which signifies that the increased oxygen demand of digestion is not met by the gastrointestinal collateral circulation. The postprandial pain frequently occurs in the mid-abdomen, suggesting that the diversion of blood flow from the SMA to supply the stomach impairs perfusion to the small bowel. This leads to transient anaerobic metabolism and acidosis. Persistent or profound mesenteric ischemia will lead to mucosal compromise with release of intracellular contents and by-products of anaerobic metabolism to the splanchnic and systemic circulation. Injured bowel mucosa allows unimpeded influx of toxic substances from the bowel lumen with systemic consequences. If full-thickness necrosis occurs in the bowel wall, intestinal perforation ensues, which will lead to peritonitis. Concomitant atherosclerotic disease in cardiac or systemic circulation frequently compounds the diagnostic and therapeutic complexity of mesenteric ischemia.

**Types of Mesenteric Artery Occlusive Disease**

There are three major mechanisms of visceral ischemia involving the mesenteric arteries: (a) acute mesenteric ischemia, which can be either embolic or thrombotic in origin; (b) chronic mesenteric ischemia; and (c) nonocclusive mesenteric ischemia. Despite the variability of these syndromes, a common anatomic pathology is involved in these processes. The superior mesenteric artery (SMA) is the most commonly involved vessel in acute mesenteric ischemia. Acute thrombosis occurs in patients with underlying mesenteric atherosclerosis, which typically involves the origin of the mesenteric arteries while sparing the collateral branches. In acute embolic mesenteric ischemia, the emboli typically originate from a cardiac source and frequently occur in patients with atrial fibrillation or following myocardial infarction (Figs. 23-37 and 23-38). Nonocclusive mesenteric ischemia is characterized by a low flow state in otherwise normal mesenteric arteries and most frequently occurs in critically ill patients on vasopressors. Finally, chronic mesenteric ischemia is a functional consequence of a long-standing atherosclerotic process that typically involves at least two of the three main mesenteric vessels. The gradual development of the occlusive process allows the development of collateral vessels that prevent the manifestations of acute ischemia, but are not sufficient to meet the high postprandial intestinal oxygen requirements, giving rise to the classical symptoms of postprandial abdominal pain and the resultant food fear.

Several less common syndromes of visceral ischemia involving the mesenteric arteries can also cause serious debilitation. Chronic mesenteric ischemic symptoms can occur due

---

**Figure 23-36.** An aortogram showing a prominent collateral vessel, which is the arc of Riolan (arrow) in a patient with an inferior mesenteric artery (IMA) occlusion. This vessel network provides collateral flow between the superior mesenteric artery and IMA.

**Figure 23-37.** An anteroposterior view of a selective superior mesenteric artery angiogram shows an abrupt cutoff of the middle colic artery, which was caused by emboli (arrow) due to atrial fibrillation.
to extrinsic compression of the celiac artery by the diaphragm, which is termed median arcuate ligament syndrome or celiac artery compression syndrome. Acute visceral ischemia may occur following an aortic operation, due to ligation of the IMA in the absence of adequate collateral vessels. Furthermore, acute visceral ischemia may develop in aortic dissection, which involves the mesenteric arteries, or after coarctation repair. Finally, other unusual causes of ischemia include mesenteric arteritis, radiation arteritis, and cholesterol emboli.

**Clinical Manifestations**

Abdominal pain out of proportion to physical findings is the classic presentation in patients with acute mesenteric ischemia and occurs following an embolic or thrombotic ischemic event of the SMA. Other manifestations include sudden onset of abdominal cramps in patients with underlying cardiac or atherosclerotic disease, often associated with bloody diarrhea, as a result of mucosal sloughing secondary to ischemia. Fever, nausea, vomiting, and abdominal distention are some common but nonspecific manifestations. Diffuse abdominal tenderness, rebound, and rigidity are late signs and usually indicate bowel infarction and necrosis.

Clinical manifestations of chronic mesenteric ischemia are more subtle due to the extensive collateral development. However, when intestinal blood flow is unable to meet the physiologic gastrointestinal demands, mesenteric insufficiency ensues. The classical symptoms include postprandial abdominal pain, food fear, and weight loss. Persistent nausea and occasionally diarrhea may coexist. Diagnosis remains challenging, and most of the patients will undergo an extensive and expensive gastrointestinal tract workup for the above symptoms prior to referral to a vascular service.

The typical patient who develops nonocclusive mesenteric ischemia is an elderly patient who has multiple comorbidities, such as congestive heart failure, acute myocardial infarction with cardiogenic shock, hypovolemic or hemorrhagic shock, sepsis, pancreatitis, and administration of digitalis or vasoconstrictor agents such as epinephrine. Abdominal pain is only present in approximately 70% of these patients. When present, the pain is usually severe but may vary in location, character, and intensity. In the absence of abdominal pain, progressive abdominal distention with acidosis may be an early sign of ischemia and impending bowel infarction.

Abdominal pain due to narrowing of the origin of the CA may occur as a result of extrinsic compression or impingement by the median arcuate ligament (Fig. 23-39). This condition is known as celiac artery compression syndrome or median arcuate ligament syndrome. Angiographically, there is CA compression that augments with deep expiration and poststenotic dilatation. The celiac artery compression syndrome has been implicated in some variants of chronic mesenteric ischemia. Most patients are young females between 20 and 40 years of age. Abdominal symptoms are nonspecific, but the pain is localized in the upper abdomen, which may be precipitated by meals.

**Diagnostic Evaluation**

The differential diagnosis of acute mesenteric ischemia includes other causes of severe abdominal pain of acute onset, such as perforated viscus, intestinal obstruction, pancreatitis, cholecystitis, and nephrolithiasis. Laboratory evaluation is neither sensitive nor specific in distinguishing these various diagnoses.
In the setting of mesenteric ischemia, complete blood count may reveal hemoconcentration and leukocytosis. Metabolic acidosis develops as a result of anaerobic metabolism. Elevated serum amylase may indicate a diagnosis of pancreatitis but is also common in the setting of intestinal infarction. Finally, increased lactate levels, hyperkalemia, and azotemia may occur in the late stages of mesenteric ischemia.

Plain abdominal radiographs may provide helpful information to exclude other causes of abdominal pain such as intestinal obstruction, perforation, or volvulus, which may exhibit symptoms mimicking intestinal ischemia. Pneumoperitoneum, pneumatosis intestinalis, and gas in the portal vein may indicate infarcted bowel. In contrast, radiographic appearance of an dynamic ileus with a gasless abdomen is the most common finding in patients with acute mesenteric ischemia.

Upper endoscopy, colonoscopy, or barium radiography does not provide any useful information when evaluating acute mesenteric ischemia. Moreover, barium enema is contraindicated if the diagnosis of mesenteric ischemia is being considered. The intraluminal barium can obscure accurate visualization of mesenteric circulation during angiography. In addition, intraperitoneal leakage of barium can occur in the setting of intestinal perforation, which can lead to added therapeutic challenges during mesenteric revascularization.

Diagnosis of chronic mesenteric ischemia can be more challenging. Usually prior to the evaluation by a vascular service, the patients have undergone an extensive workup for the symptoms of chronic abdominal pain, weight loss, and anorexia. Rarely, the vascular surgeon is the first to encounter a patient with the above symptoms. In this situation, it is advisable to keep in mind that mesenteric ischemia is a rare entity and that a full diagnostic workup that should include CT scan of the abdomen and evaluation by gastroenterologist should be performed. Mesenteric occlusive disease may coexist with malignancy, and symptoms of mesenteric vessel stenosis may be the result of extrinsic compression by a tumor.

Duplex ultrasonography is a valuable noninvasive means of assessing the patency of the mesenteric vessels. Moneta and associates evaluated the use of duplex ultrasound in the diagnosis of mesenteric occlusive disease in a blinded prospective study. A peak systolic velocity in the SMA > 275 cm/s demonstrated a sensitivity of 92%, specificity of 96%, and overall accuracy of 96% for detecting >70% stenosis. The same authors found sensitivity and specificity of 87% and 82%, respectively, with an accuracy of 82% in predicting >70% celiac trunk stenosis. Duplex has been successfully used for follow-up after open surgical reconstruction or endovascular treatment of the mesenteric vessels to assess recurrence of the disease. Finally, spiral CT with three-dimensional reconstruction (Fig. 23-40) and MRA (Fig. 23-41) have been promising in providing clear radiographic assessment of the mesenteric vessels.

The definitive diagnosis of mesenteric vascular disease is made by biplanar mesenteric arteriography, which should be performed promptly in any patient with suspected mesenteric occlusion. It typically shows occlusion or near-occlusion of the CA and SMA at or near their origins from the aorta. In most cases, the IMA has been previously occluded secondary to diffuse infrarenal aortic atherosclerosis. The differentiation of the different types of mesenteric arterial occlusion may be suggested with biplanar mesenteric arteriogram. Mesenteric emboli typically lodge at the orifice of the middle colic artery, which creates a “meniscus sign” with an abrupt cutoff of a normal proximal SMA several centimeters from its origin on the aorta. Mesenteric thrombosis, in contrast, occurs at the most proximal SMA, which tapers off at 1 to 2 cm from its origin. In the case of chronic mesenteric occlusion, the appearance of collateral circulation is typically present. Nonocclusive mesenteric ischemia produces an arteriographic image of segmental mesenteric vasospasm with a relatively normal-appearing main SMA trunk (Fig. 23-42).

Mesenteric arteriography can also play a therapeutic role. Once the diagnosis of nonocclusive mesenteric ischemia is
SPECIFIC CONSIDERATIONS

Figure 23-42. Mesenteric arteriogram showing nonocclusive mesenteric ischemia as evidenced by diffuse spasm of intestinal arcades with poor filling of intramural vessels.

made on the arteriogram, an infusion catheter can be placed at the SMA orifice, and vasodilating agents, such as papaverine, can be administered intra-arterially. The papaverine infusion may be continued postoperatively to treat persistent vasospasm, a common occurrence following mesenteric reperfusion. Transcatheter thrombolytic therapy has little role in the management of thrombotic mesenteric occlusion. Although thrombolytic agents may transiently recannulate the occluded vessels, the underlying occlusive lesions require definitive treatment. Furthermore, thrombolytic therapy typically requires a prolonged period of time to restore perfusion, during which the intestinal viability will be difficult to assess.

A word of caution would be appropriate here regarding patients with typical history of chronic intestinal angina who present with an acute abdomen and classical findings of peritoneal irritation. Arteriography is the gold standard for the diagnosis of mesenteric occlusive disease; however, it can be a time-consuming diagnostic modality. In this group of patients, immediate exploration for assessment of intestinal viability and vascular reconstruction is the best choice.

Surgical Repair

Acute Embolic Mesenteric Ischemia. Initial management of patients with acute mesenteric ischemia includes fluid resuscitation and systemic anticoagulation with heparin to prevent further thrombus propagation. Significant metabolic acidosis not responding to fluid resuscitation should be corrected with sodium bicarbonate. A central venous catheter, peripheral arterial catheter, and Foley catheter should be placed for hemodynamic status monitoring. Appropriate antibiotics are given prior to surgical exploration. The operative management of acute mesenteric ischemia is dictated by the cause of the occlusion. It is helpful to obtain a preoperative mesenteric arteriogram to confirm the diagnosis and to plan appropriate treatment options. However, the diagnosis of mesenteric ischemia frequently cannot be established prior to surgical exploration, and therefore, patients in a moribund condition with acute abdominal symptoms should undergo immediate surgical exploration, avoiding the delay required to perform an arteriogram.

The primary goal of surgical treatment in embolic mesenteric ischemia is to restore arterial perfusion with removal of the embolus from the vessel. The abdomen is explored through a midline incision, which often reveals variable degrees of intestinal ischemia from the mid-jejunum to the ascending or transverse colon. The transverse colon is lifted superiorly, and the small intestine is reflected toward the right upper quadrant. The SMA is approached at the root of the small bowel mesentery, usually as it emerges from beneath the pancreas to cross over the junction of the third and fourth portions of the duodenum. Alternatively, the SMA can be approached by incising the retroperitoneum lateral to the fourth portion of the duodenum, which is rotated medially to expose the SMA. Once the proximal SMA is identified and controlled with vascular clamps, a transverse arteriotomy is made to extract the embolus, using standard balloon embolectomy catheters. In the event the embolus has lodged more distally, exposure of the distal SMA may be obtained in the root of the small bowel mesentery by isolating individual jejunal and ileal branches to allow a more comprehensive thromboembolectomy. Following the restoration of SMA flow, an assessment of intestinal viability must be made, and nonviable bowel must be resected. Several methods have been described to evaluate the viability of the intestine, which include intraoperative intravenous fluorescein injection and inspection with a Wood’s lamp, and Doppler assessment of antimesenteric intestinal arterial pulsations. A second-look procedure should be considered in many patients and is performed 24 to 48 hours following embolectomy. The goal of the procedure is reassessment of the extent of bowel viability, which may not be obvious immediately following the initial embolectomy. If nonviable intestine is evident in the second-look procedure, additional bowel resections should be performed at that time.

Acute Thrombotic Mesenteric Ischemia. Thrombotic mesenteric ischemia usually involves a severely atherosclerotic vessel, typically the proximal CA and SMA. Therefore, these patients require a reconstructive procedure to the SMA to bypass the proximal occlusive lesion and restore adequate mesenteric flow. The saphenous vein is the graft material of choice, and prosthetic materials should be avoided in patients with nonviable bowel, due to the risk of bacterial contamination if resection of necrotic intestine is performed. The bypass graft may originate from either the aorta or iliac artery. Advantages from using the supraceliac infradiaphragmatic aorta as opposed to the infrarenal aorta as the inflow vessel include a smoother graft configuration with less chance of kinking and the absence of atherosclerotic disease in the supraceliac aortic segment. Exposure of the supraceliac aorta is technically more challenging and time consuming than that of the iliac artery, which unless calcified is an appropriate inflow. Patency rates are similar regardless of inflow vessel choice.92

Chronic Mesenteric Ischemia. The therapeutic goal in patients with chronic mesenteric ischemia is to revascularize mesenteric circulation and prevent the development of bowel infarction. Mesenteric occlusive disease can be treated successfully by either transaortic endarterectomy or mesenteric artery bypass.
Transaortic endarterectomy is indicated for ostial lesions of patient CA and SMA. A left medial rotation is performed, and the aorta and the mesenteric branches are exposed. A lateral aortotomy is performed encompassing both the CA and SMA orifices. The visceral arteries must be adequately mobilized so that the termination site of endarterectomy can be visualized. Otherwise, an intimal flap may develop, which can lead to early thrombosis or distal embolization.

For occlusive lesions located 1 to 2 cm distal to the mesenteric origin, mesenteric artery bypass should be performed. Multiple mesenteric arteries are typically involved in chronic mesenteric ischemia, and both the CA and SMA should be revascularized whenever possible. In general, bypass grafting may be performed either antegrade from the supraceliac aorta or retrograde from either the infrarenal aorta or iliac artery. Both autogenous saphenous vein grafts and prosthetic grafts have been used with satisfactory and equivalent success. An antegrade bypass also can be performed using a small-caliber bifurcated graft from the supraceliac aorta to both the CA and SMA, which yields an excellent long-term result.103

Celiac Artery Compression Syndrome. The decision to intervene in patients with CA compression syndrome should be based on both an appropriate symptom complex and the finding of celiac artery compression in the absence of other findings to explain the symptoms. The treatment goal is to release the ligamentous structure that compresses the proximal CA and to correct any persistent stenosis by bypass grafting. Some surgeons advocate careful celiac plexus sympathectomy in addition to arcuate ligament decompression to ensure good treatment outcome.102 The patient should be cautioned that relief of the celiac compression cannot be guaranteed to relieve the symp- toms. In a number of reports on endovascular management of chronic mesenteric ischemia, the presence of CA compression syndrome has been identified as a major factor of technical failure and recurrence. Therefore, angioplasty and stenting should not be undertaken if extrinsic compression of the CA by the median arcuate ligament is suspected based on preoperative imaging studies. Open surgical treatment should be performed instead. A study that analyzed the outcome of laparoscopic and open median arcuate ligament release cases in the literature showed both approaches to be effective in symptom relief (85%), with no difference in late symptom recurrence rate (6.8% in the open group and 5.7% in the laparoscopic group).104

Endovascular Treatment

Chronic Mesenteric Ischemia. Endovascular treatment of mesenteric artery stenosis or short segment occlusion by balloon dilatation or stent placement represents a less invasive therapeutic alternative to open surgical intervention, particularly in patients whose medical comorbidities place them at a high operative risk category. Endovascular therapy is also suited in patients with recurrent disease or anastomotic steno- sis following previous open mesenteric revascularization. Prophylactic mesenteric revascularization is rarely performed in the asymptomatic patient undergoing an aortic procedure for other indications. However, the natural history of untreated chronic mesenteric ischemia may justify revascularization in some minimally symptomatic or asymptomatic patients if the operative risks are acceptable, since the first clinical presentation may be acute intestinal ischemia in as many as 50% of the patients, with a mortality rate that ranges from 15% to 70%.105

This is particularly true when the SMA is involved. Mesenteric angioplasty and stenting is particularly suitable for this patient subgroup given its low morbidity and mortality. Because of the limited experience with stent use in mesenteric vessels, appropriate indications for primary stent placement have not been clearly defined. Guidelines generally include calcified ostial stenoses, high-grade eccentric stenoses, chronic occlusions, and significant residual stenosis >30% or the presence of dissection after angioplasty. Restenosis after PTA is also an indication for stent placement.106

Acute Mesenteric Ischemia. Catheter-directed thrombol- ytic therapy is a potentially useful treatment modality for acute mesenteric ischemia, which can be initiated with intra-arterial delivery of thrombolytic agent into the mesenteric thrombus at the time of diagnostic angiography. Various thrombolytic medi- cations, including urokinase (Abbokinase; Abbott Laboratory, North Chicago, IL) or recombinant tissue plasminogen activator (Activase; Genentech, South San Francisco, CA), have been reported to be successful in a small series of case reports. Catheter-directed thrombolytic therapy has a higher probability of restoring mesenteric blood flow success when performed within 12 hours of symptom onset. Successful resolution of a mesenteric thrombus will facilitate the identification of the underlying mesenteric occlusive disease process. As a result, subsequent operative mesenteric revascularization or mesenteric balloon angioplasty and stenting may be performed elec- tively to correct the mesenteric stenosis. There are two main drawbacks with regard to thrombolytic therapy in mesenteric ischemia. Percutaneous catheter-directed thrombolysis does not allow the possibility to inspect the potentially ischemic intesti- ne following restoration of the mesenteric flow. Additionally, a prolonged period of time may be necessary in order to achieve successful catheter-directed thrombolysis, due in part to serial angiographic surveillance to document thrombus resolution. An incomplete or unsuccessful thrombolysis may lead to delayed operative revascularization, which may further necessitate bowel resection for irreversible intestinal necrosis. Therefore, catheter-directed thrombolytic therapy for acute mesenteric ischemia should only be considered in selected patients under a closely scrutinized clinical protocol.

Nonocclusive Mesenteric Ischemia. The treatment of nonocclusive mesenteric ischemia is primarily pharmacologic with selective mesenteric arterial catheterization followed by infusion of vasodilatory agents, such as tolazoline or papaverine. Once the diagnosis is made on the mesenteric arteriography (see Fig. 23-42), intra-arterial papaverine is given at a dose of 30 to 60 mg/h. This must be coupled with the cessation of other vasoconstricting agents. Concomitant intravenous heparin should be administered to prevent thrombosis in the cannulated vessels. Treatment strategy thereafter is dependent on the patient’s clinical response to the vasodilator therapy. If abdominal symptoms improve, mesenteric arteriography should be repeated to docu- ment the resolution of vasospasm. The patient’s hemodynamic status must be carefully monitored during papaverine infusion as significant hypotension can develop in the event that the infusion catheter migrates into the aorta, which can lead to systemic circulation of papaverine. Surgical exploration is indicated if the patient develops signs of continued bowel ischemia or infarction as evidenced by rebound tenderness or involuntary guarding. In these circumstances, papaverine infusion should be continued intraoperatively and postoperatively. The operating room
should be kept as warm as possible, and warm irrigation fluid and laparotomy pads should be used to prevent further intestinal vasocstriction during exploration.

**Techniques of Endovascular Interventions.** To perform endovascular mesenteric revascularization, intraluminal access is performed via a femoral or brachial artery approach. Once an introducer sheath is placed in the femoral artery, an anteroposterior and lateral aortogram just below the level of the diaphragm is obtained with a pigtail catheter to identify the origin of the CA and SMA. Initial catheterization of the mesenteric artery can be performed using a variety of selective angled catheters, which include the RDC, Cobra-2, Simmons I (Boston Scientific/ Meditech, Natick, MA), or SOS Omni catheter (Angiodynamics, Queensbury, NY). Once the mesenteric artery is cannulated, systemic heparin (5000 IU) is administered intravenously. A selective mesenteric angiogram is then performed to identify the diseased segment, which is followed by the placement of a 0.035-inch or less traumatic 0.014- to 0.018-inch guidewire to cross the stenotic lesion. Once the guidewire is placed across the stenosis, the catheter is carefully advanced over the guidewire across the lesion. In the event that the mesenteric artery is severely angulated as it arises from the aorta, a second stiffer guidewire (Amplatz or Rosen Guidewire, Boston Scientific) may be exchanged through the catheter to facilitate the placement of a 6-French guiding sheath (Pinnacle, Boston Scientific).

With the image intensifier angled in a lateral position to fully visualize the proximal mesenteric segment, a balloon angioplasty is advanced over the guidewire through the guiding sheath and positioned across the stenosis. The balloon diameter should be chosen based on the vessel size of the adjacent normal mesenteric vessel. Once balloon angioplasty is completed, a postangioplasty angiogram is necessary to document the procedural result. Radiographic evidence of either residual stenosis or mesenteric artery dissection constitutes suboptimal angioplasty results that warrants mesenteric stent placement. Moreover, atherosclerotic involvement of the proximal mesenteric artery or vessel orifice should be treated with balloon-expandable stent placement. These stents can be placed over a low-profile 0.014- or 0.018-inch guidewire system. It is preferable to deliver the balloon-mouted stent through a guiding sheath, which is positioned just proximal to the mesenteric orifice while the balloon-mouted stent is advanced across the stenosis. The stent is next deployed by expanding the angioplasty balloon to its designated inflation pressure. The balloon is then deflated and carefully withdrawn through the guiding sheath.

Completion angiogram is performed by hand injecting a small volume of contrast though the guiding sheath. It is critical to maintain the guidewire access until satisfactory completion angiogram is obtained. If the completion angiogram reveals suboptimal radiographic results, such as residual stenosis or dissection, additional catheter-based intervention can be performed through the same guidewire. These interventions may include repeat balloon angioplasty for residual stenosis or additional stent placement for mesenteric artery dissection. During the procedure, intra-arterial infusion of papaverine or nitroglycerine can be used to decrease vasospasm. Administration of antiplatelet agents is also recommended for at least 6 months or even indefinitely if other risk factors of cardiovascular disease are present.

**Complications of Endovascular Treatment.** Complications are not common and rarely become life threatening. These include access site thrombosis, hematomas, and infection. Dissection can occur during PTA and is managed with placement of a stent. Balloon-mounted stents are preferred over the self-expanding ones because of the higher radial force and the more precise placement. Distal embolization has also been reported, but it never resulted in acute intestinal ischemia, likely due to the rich network of collaterals already developed.

**Clinical Results of Interventions for Mesenteric Ischemia**

The first successful percutaneous angioplasty of the SMA was reported in 1980. Since 1995, multiple series and scattered case reports have reported results from endovascular management of mesenteric occlusive disease. A literature review by AbuRahma and colleagues in 2003 showed that endovascular intervention had an overall technical success rate of 91%, early and late pain relief rates of 84% and 71%, respectively, and 30-day morbidity and mortality rates of 16.4% and 4.3%, respectively. The average patency was 63% during an average 26-month follow-up.

In our review of the literature from published series since 1995, restenosis developed in 22% of patients during 24.5 months of average follow-up. The long-term clinical relief without reintervention was 82%. Among the patients who experienced a technical failure, 15 were ultimately diagnosed with median arcuate ligament syndrome and underwent successful surgical treatment, an observation that emphasizes the need for careful patient selection. Interestingly, the addition of selective stenting after PTA that was started in 1998, while it slightly increases the technical success rate, is not correlated with any substantial overall clinical benefit or improved long-term patency rates.

In contrast to endovascular treatment, open surgical techniques have achieved an immediate clinical success rate that approaches 100%, a surgical mortality rate of 0% to 17%, and an operative morbidity rate that ranges from 19% to 54% in a number of different series. AbuRahma and colleagues reported their experience of endovascular interventions of 22 patients with symptomatic mesenteric ischemia due to either SMA or CA stenosis. They noted an excellent initial technical and clinical success rates, which were 96% (23 of 24 patients) and 95% (21 of 22 patients), respectively, with no perioperative mortality or major morbidity. During a mean follow-up of 26 months (range, 1–54 months), the primary late clinical success rate was 61%, and freedom from recurrent stenosis was 30%. The freedom from recurrent stenosis rates at 1, 2, 3, and 4 years were 65%, 47%, 39%, and 13%, respectively. The authors concluded that mesenteric stenting, which provides excellent early results, is associated with a relative high incidence of late restenosis.

Several studies have attempted to compare the endovascular with the standard open surgical approach. The results of the open surgery appear to be more durable, but it tends to be associated with higher morbidity and mortality rates and an overall longer hospital stay. In one study that compared the clinical outcome of open revascularization with percutaneous stenting for patients with chronic mesenteric ischemia, 28 patients underwent endovascular treatment and 85 patients underwent open mesenteric bypass grafting. With both patient cohorts having similar baseline comorbidities and symptom duration, there was no difference in early in-hospital complication or mortality rates. Moreover, both groups had similar 3-year cumulative recurrent stenosis and mortality rates. However, patients
treated with mesenteric stenting had a significantly higher incidence of recurrent symptoms. The authors concluded that operative mesenteric revascularization should be offered to patients with low surgical risk.

Based on the above results one could argue that mesenteric angioplasty and stenting demonstrate an inferior technical and clinical success rate. Long-term patency rates appear to also be superior with the open technique. There is a general consensus, however, that the endovascular approach is associated with lower morbidity and mortality rates and is therefore more suitable for high-risk patients. One should also keep in mind that practices representing standard of care for stent placement today were absent in the early era of endovascular experience. These include perioperative heparinization and short-term antiplatelet therapy, use of stents with higher radial force, routine use of postoperative surveillance with arterial duplex and early reintervention to prevent a high-grade stenosis from progressing to occlusion, and placement of drug-eluting stents. One such example is a recent nonrandomized study to compare the outcomes of mesenteric angioplasty using covered stents or bare metal stents in patients undergoing primary or reintervention for chronic mesenteric ischemia. The study showed that covered stents are associated with less restenosis (18% vs. 47%), symptom recurrence (18% vs. 50%), and reintervention (9% vs. 44%) at 24 months and better primary patency at 3 years (92% vs. 52%) than bare metal stents in the primary intervention group. Similar results were found in the reintervention group as well.

**RENA L ARTERY DISEASE**

Obstructive lesions of the renal artery can produce hypertension, resulting in a condition known as renovascular hypertension, which is the most common form of hypertension amenable to therapeutic intervention, and affects 5% to 10% of all hypertensive patients in the United States. Patients with renovascular hypertension are at an increased risk for irreversible end-organ dysfunction, including permanent kidney damage, if inadequate pharmacologic therapies are used to control the blood pressure. The majority of patients with renal artery obstructive disease have vascular lesions of either atherosclerotic disease or fibrodysplasia involving the renal arteries. The proximal portion of the renal artery represents the most common location for the development of atherosclerotic disease. It is well established that renal artery intervention, either by surgical or endovascular revascularization, provides an effective treatment for controlling renovascular hypertension as well as preserving renal function. The decision for intervention is complex and needs to consider a variety of anatomic, physiologic, and clinical features, unique for the individual patient.

**Etiology**

Approximately 80% of all renal artery occlusive lesions are caused by atherosclerosis, which typically involves a short segment of the renal artery ostia and represents spillover disease from a severely atheromatous aorta (Fig. 23-43). Atherosclerotic lesions are bilateral in two thirds of patients. Individuals with this disease commonly present during the sixth decade of life. Men are affected twice as frequently as women. Atherosclerotic lesions in other territories such as the coronary, mesenteric, cerebrovascular, and peripheral arterial circulation are common. When a unilateral lesion is present, the disease process equally affects the right and left renal arteries.

![Figure 23-43](image1.png)

**Figure 23-43.** Occlusive disease of the renal artery typically involves the renal ostium (arrow) as a spillover plaque extension from aortic atherosclerosis.

The second most common cause of renal artery stenosis is FMD, which accounts for 20% of cases and is most frequently encountered in young, often multiparous women. FMD of the renal artery represents a heterogeneous group of lesions that can produce histopathologic changes in the intima, media, or adventitia. The most common variety consists of medial fibroplasia, in which thickened fibromuscular ridges alternate with attenuated media producing the classic angiographic “string of beads” appearance (Figs. 23-44 and 23-45). The cause of medial fibroplasia remains unclear. Most common theories involve a modification of arterial smooth muscle cells in response to
estrogenic stimuli during the reproductive years, unusual traction forces on affected vessels, and mural ischemia from impairment of vasa vasorum blood flow. Fibromuscular hyperplasia usually affects the distal two thirds of the main renal artery, and the right renal artery is affected more frequently than the left. Other less common causes of renal artery stenosis include renal artery aneurysm (compressing the adjacent normal renal artery), arteriovenous malformations, neurofibromatosis, renal artery dissections, renal artery trauma, Takayasu’s arteritis, and renal arteriovenous fistula.

Clinical Manifestations
Renovascular hypertension is the most common sequela of renal artery occlusive disease. Its prevalence varies from 2% in patients with diastolic blood pressure greater than 100 mmHg to almost 30% in those with diastolic blood pressure over 125 mmHg. Clinical features that may indicate the presence of renovascular hypertension include the following: (a) systolic and diastolic upper abdominal bruits; (b) diastolic hypertension of greater than 115 mmHg; (c) rapid onset of hypertension after the age of 50 years; (d) a sudden worsening of mild to moderate essential hypertension; (e) hypertension that is difficult to control with three or more antihypertensives; (f) development of renal insufficiency after angiotensin-converting enzyme inhibitors; and (g) development of hypertension during childhood.

All patients with significant hypertension, especially elevated diastolic blood pressure, must be considered as suspect for renovascular disease. Young adults with hypertension have a great deal to gain by avoiding lifelong treatment if renovascular hypertension is diagnosed and corrected. Appropriate diagnostic studies and intervention must be timely instituted to detect the possibility of renovascular hypertension in patients with primary hypertension who present for clinical evaluation.

Diagnostic Evaluation
The diagnostic requisites for renovascular hypertension include both hypertension and renal artery stenosis. Impairment of the renal function may coexist, although the occurrence of renal insufficiency prior to the development of hypertension is uncommon. Nearly all diagnostic studies for renovascular hypertension evaluate either the anatomic stenosis or renal parenchymal dysfunction attributed to the stenosis. The following section provides an overview of the strengths and limitations of the most common tests used in the diagnostic evaluation of the patient with suspected renovascular hypertension prior to intervention.

Captopril renal scanning is a functional study that assesses renal perfusion before and after administration of the angiotensin-converting enzyme inhibitor captopril. Captopril inhibits the secretion of angiotensin II. Through this mechanism, it reduces the efferent arteriolar vasoconstriction and, as a result, the glomerular filtration rate (GFR). The test consists of a baseline renal scan and a second renal scan after captopril administration. A positive result indicates that captopril administration (a) increases the time to peak activity to more than 11 minutes or (b) the GFR ratio between sides increases to greater than 1.5:1 compared to a normal baseline scan. Significant parenchymal disease limits the reliability of this study.

Renal artery duplex ultrasonography is a noninvasive test of assessing renal artery stenosis both by visualization of the vessel and measurement of the effect of stenosis on blood flow velocity and waveforms. The presence of a severe renal artery stenosis correlates with peak systolic velocities of greater than 180 cm/s and the ratio of these velocities to those in the aorta of greater than 3.5 (Table 23-11). Renal artery duplex is a technically demanding exam, requiring a substantial amount of operator expertise. In addition, the presence of bowel gas and obesity may make the exam difficult to perform and interpret. However, in experienced hands and with appropriate patient selection, it can be a high-yield exam and is typically the initial screening test for patients with suspected renal artery occlusive disease.

Selective catheterization of the renal vein via a femoral vein approach for assessing renin activity is a more invasive test of detecting the physiologic sequelae of renal artery stenosis. If unilateral disease is present, the affected kidney should secrete high levels of renin while the contralateral kidney should have low renin production. A ratio between the two kidneys, or the renal vein renin ratio (RVRR), of greater than 1.5 is indicative of functionally important renovascular hypertension, and it also predicts a favorable response from renovascular revascularization. Since this study assesses the ratio between the two kidneys, it is not useful in patients with bilateral disease because both kidneys may secrete abnormally elevated renin levels.

The renal:systemic renin index (RSRI) is calculated by subtracting systemic renin activity from individual renal vein renin activity and dividing the remainder by systemic renin activity. This value represents the contribution of each kidney to renin production. In the absence of renal artery stenosis, the

### Table 23-11

| Renal duplex diagnostic criteria |
|---------------------------------|-----------------|--------|
| **RENA L ARTERY**               | **RENA L ARTERY** | **RAR** |
| **DIAMETER REDUCTION**          | **PSV**      |       |
| Normal                          | <180 cm/s    | <3.5  |
| <60%                            | ≥180 cm/s    | <3.5  |
| ≥60%                            | ≥180 cm/s    | ≥3.5  |
| Oclusion                        | No signal    | No signal |

PSV = peak systolic velocity; RAR = renal-to-aortic ratio.
renal vein renin activity from each kidney is typically 24% or 0.24 higher than the systemic level. As the result, the total of both kidneys’ renin activity is usually 48% greater than the systemic activity, a value that represents a steady state of renal renin activity. The RSRI of the affected kidney in patients with renovascular hypertension is greater than 0.24. In the case of unilateral renal artery stenosis with normal contralateral kidney, the increase in ipsilateral renin release is normally balanced by suppression of the contralateral kidney renin production, which results in a drop in its RSRI to less than 0.24. Bilateral renal artery disease may negate the contralateral compensatory response, and the autonomous release of renin from both diseased kidneys may result in the sum of the individual RSRIs to be considerably greater than 0.48. The prognostic value of RSRI remains limited in that approximately 10% of patients with favorable clinical response following renovascular revascularization do not exhibit contralateral renin suppression. As a result, the use of RSRI must be applied with caution in the management of patients with renovascular hypertension.

MRA with intravenous gadolinium contrast enhancement has been increasingly used for renal artery imaging because of its ability to provide high-resolution images (Figs. 23-46 and 23-47) while using a minimally nephrotoxic agent. Flow void may be inaccurately interpreted as occlusion or stenosis in MRA. Therefore, unless the quality of the image analysis software is superior, MRA should be interpreted with caution and used in conjunction with other modalities prior to making plans for operative or endovascular treatment.

DSA remains the gold standard to assess renal artery occlusive disease. A flush aortogram is performed first so that any accessory renal arteries can be detected and the origins of all the renal arteries are adequately displayed. The presence of collateral vessels circumventing a renal artery stenosis strongly supports the hemodynamic importance of the stenosis. A pressure gradient of 10 mmHg or greater is necessary for collateral vessel development, which is also associated with activation of the renin-angiotensin cascade.

**Treatment Indications**

The therapeutic goals in patients with renovascular disease include: (a) improved blood pressure control, in order to prevent end-organ damage on systems such as the cerebral, coronary, pulmonary, and peripheral circulations; and (b) preservation and possibly improvement of the renal function (Table 23-12).

---

**Table 23-12**

**Indications for renal artery revascularization**

**Angiography Criteria**

- Documented renal artery stenosis (>70% diameter reduction)
- Fibromuscular dysplasia lesion
- Pressure gradient >20 mmHg
- Affected/unaffected kidney renin ratio >1.5 to 1

**Clinical Criteria**

- Refractory or rapidly progressive hypertension
- Hypertension associated with flash pulmonary edema without coronary artery disease
- Rapidly progressive deterioration in renal function
- Intolerance to antihypertensive medications
- Chronic renal insufficiency related to bilateral renal artery occlusive disease or stenosis to a solitary functioning kidney
- Dialysis-dependent renal failure in a patient with renal artery stenosis but without another definite cause of end-stage renal disease
- Recurrent congestive heart failure or flash pulmonary edema not attributable to active coronary ischemia
The indications for endovascular treatment for renal artery occlusive disease include 70% or greater stenosis of one or both renal arteries and at least one of the following clinical criteria:

- Inability to adequately control hypertension despite appropriate antihypertensive regimen.
- Chronic renal insufficiency related to bilateral renal artery occlusive disease or stenosis to a solitary functioning kidney.
- Dialysis-dependent renal failure in a patient with renal artery stenosis but without another definite cause of end-stage renal disease.
- Recurrent congestive heart failure or flash pulmonary edema not attributable to active coronary ischemia.

Prior to 1990, the most common treatment modality in patients with renal artery occlusive disease is surgical revascularization, with either renal artery bypass grafting or renal artery endarterectomy. The advancement of endovascular therapy in the past decade has led to various minimally invasive treatment strategies such as renal artery balloon angioplasty or stenting to control hypertension or to preserve renal function.

**Surgical Reconstruction**

The typical approach for surgical renal artery revascularization involves a midline xiphoid-to-pubis incision. The posterior peritoneum is incised, and the duodenum is mobilized to the right, starting at the ligament of Treitz. The left renal hilum can be exposed by extending the retroperitoneal dissection to the left along the avascular plane along the inferior border of the pancreas. Mobilization of the left renal vein is essential in these cases and can be achieved by dividing the gonadal, ilio-lumbar, and adrenal veins. The proximal portion of the right renal artery can be exposed through the base of the mesentery by retraction of the left renal vein cephalad and the vena cava to the right. Accessing the most distal portion of the right renal artery requires a Kocher maneuver and duodenal mobilization. Another approach useful for treating bilateral renal artery lesions involves mobilization of the entire small bowel and the right colon, with a dissection that starts at the ligament of Treitz and proceeds toward the cecum and then along the line of Todd in the right paracolic gutter. Simultaneous dissection along the inferior border of the pancreas provides additional visualization of the left renal artery. Finally, division of the diaphragmatic crura that encircle the suprarenal aorta may sometimes be necessary to achieve suprarenal clamping.

**Types of Surgical Reconstruction.** Aortorenal bypass is the most frequently performed reconstruction of ostial occlusive renal artery disease. After proximal and distal control is obtained, an elliptical segment of the aorta is excised, and the proximal anastomosis is performed in end-to-side fashion. Autologous vein is the preferred conduit. If the vein is not suitable, then prosthetic material can be used. An end-to-end anastomosis is then performed between the conduit of choice and the renal artery using either a 6-0 or 7-0 polypropylene suture. The length of the arteriotomy needs to be at least three times the diameter of the renal artery to prevent anastomotic restenosis. In the event that the surgeon plans to perform a side-to-side anastomosis between the conduit and the renal artery, this is performed first, and the aortic anastomosis follows.

Endarterectomy, either transrenal or transaortic, is an alternative to bypass for short ostial lesions or in patients with multiple renal arteries. The transrenal endarterectomy is performed with a transverse longitudinal incision on the aorta that extends into the diseased renal artery. After plaque removal, the arteriotomy is closed with a prosthetic patch. Transaortic endarterectomy is well suited for patients with multiple renal arteries and short ostial lesions. The aorta is opened longitudinally and aortic sleeve endarterectomy is performed, followed by ever- sion endarterectomy of the renal arteries. Adequate mobilization of the renal arteries is essential for a safe and complete endarterectomy.

Hepatorenal and splenorenal bypass are alternative options of revascularization for patients who might not tolerate aortic clamping or for those with calcified aorta that precludes adequate control. For hepatorenal bypass, a right subcostal incision is used, and the hepatic artery is exposed with an incision in the lesser omentum. A Kocher maneuver is performed, the right renal vein is identified and mobilized, and the right renal artery is identified and controlled posteriorly to the vein. Greater saphenous vein is the conduit of choice. The anastomosis is performed end-to-side with the common hepatic artery, and end-to-end with the renal artery anterior to the inferior vena cava. The splenorenal bypass is performed via a left subcostal incision. The splenic artery is mobilized from the lesser sac, brought through a retropancreatic plane, and anastomosed end-to-end to the renal artery.

Reimplantation of the renal artery is an attractive option of reconstruction in children or in adults with ostial lesions. A redundant renal artery is a prerequisite for the procedure. After mobilization, the artery is transected and spatulated, eversion endarterectomy is performed if necessary, and an end-to-side anastomosis with the aorta is created.

**Clinical Results of Surgical Repair**

Results reflect the need for performance of renal artery bypass in high-volume and experienced centers. In a review from a large tertiary center, 92% of the patients with nonatherosclerotic vascular disease had improvement in hypertension, but only 43% were completely cured and taken off antihypertensives. Patients younger than age 45 fare better, with a cure rate of 68% and improvement rate of 32%. In patients with atherosclerotic renal artery disease, the cure rate was even smaller (12%), and the overall response to hypertension rate was 85%. The operative mortality rates were 3.1% and 0% in the atherosclerotic and nonatherosclerotic groups, respectively.

Renal function improvement occurs within the first week of the operation in approximately two-thirds of patients. A progressive decrease in the GFR is seen after this initial improvement, but the rate of decrease is less compared with patients who did not respond at all to operative intervention. Up to three-quarters of patients were permanently removed from dialysis in a large series. Favorable response of renal function to revascularization improves overall survival.

**Endovascular Treatment**

Endovascular treatment of renal artery occlusive disease was first introduced by Grünzig who successfully dilated a renal artery stenosis using a balloon catheter technique. This technique requires passage of a guidewire under fluoroscopic control typically from a femoral artery approach to across the stenosis in the renal artery. A balloon dilating catheter is passed over the guidewire and positioned within the area of stenosis and inflated to produce a controlled disruption of the arterial wall. Alternatively, a balloon-mounted expandable stent can be used.
to primarily dilate the renal artery stenosis. Completion angiography is usually performed to assess the immediate results. The technical aspect of an endovascular renal artery revascularization is discussed in the following section.

**Techniques of Renal Artery Angioplasty and Stenting.** Access to the renal artery for endovascular intervention is typically performed via a femoral artery approach, although a brachial artery approach can be considered in the event of severe aortoiliac occlusive disease, aortoiliac aneurysm, or severe caudal renal artery angulation. Once an introducer sheath is placed in the femoral artery, an aortogram is performed with a pigtail catheter placed in the suprarenal aorta. Additional oblique views are frequently necessary to more precisely visualize the orifice of the stenosed renal artery and thoroughly assess the presence of accessory renal arteries. Noniodinated contrast agents, such as carbon dioxide and gadolinium, can be used in endovascular renal intervention in patients with renal dysfunction or history of allergic reaction.

After systemic heparinization, catheterization of the renal artery can be performed using a variety of selective angled catheters, including the RDC, Cobra-2, Simmons I, or SOS Omni catheter. A selective renal angiogram is then performed to confirm position, and the lesion is crossed with either 0.035-inch or a 0.018- to 0.014-inch guidewires. It is important to maintain the distal wire position without movement in the tertiary renal branches during guiding sheath placement to reduce the possibility of parenchymal perforation and spasm. A guiding sheath or a guiding catheter is then advanced at the orifice of the renal artery and provides a secure access for balloon and stent deployment.

Balloon angioplasty is performed with a balloon sized to the diameter of the normal renal artery adjacent to the stenosis. Choosing a balloon with diameter 4 mm is a reasonable first choice. The luminal diameter of the renal artery can be further assessed by comparing it to the fully inflated balloon. Such a comparison may provide a reference guide to determine whether renal artery dilatation with a larger diameter angioplasty balloon is necessary.

Once balloon angioplasty of the renal artery is completed, an angiogram is performed to document the procedural result. Radiographic evidence of either residual stenosis or renal artery dissection constitutes suboptimal angioplasty results, which warrants an immediate renal artery stent placement. Moreover, atherosclerotic involvement of the very proximal renal artery that involves the vessel orifice typically requires stent placement. A balloon-expandable stent is typically used and is positioned in such a way that it protrudes into the aorta by 1 to 2 mm. The size of the stent is determined by the size of the renal artery, taking into account a desirable 10% to 20% oversizing. After the stent deployment, the angiogram is repeated, and upon a satisfactory result, the devices are withdrawn. It is critical to maintain the guidewire access across the renal lesion until satisfactory completion angiogram is obtained. Spasm of the branches of the renal artery will usually respond to nitroglycerin 100 to 200 μg administered through the guiding sheath directly into the renal artery.

While endovascular therapy of renal artery occlusive disease is considerably less invasive than conventional renal artery bypass operation, complications relating to this treatment modality can occur. In a study in which Guzman and colleagues compared the complications following renal artery angioplasty and surgical revascularization, the authors noted that major complication rates following endovascular and surgical treatment were 17% and 31%, respectively. In contrast, significantly greater minor complications were associated with the endovascular cohort, with a minor complication rate of 48% compared with 7% in the surgical group. In a prospective randomized study that compared the clinical outcome of renal artery balloon angioplasty versus stenting for renal ostial ath erosclerotic lesion, comparable complications rates were found in the two groups (39% vs. 43%, respectively). However, the incidence of restenosis at 6 months was significantly higher in the balloon angioplasty cohort than the stenting group (48% vs. 14%, respectively). This study underscores the clinical superiority of renal stenting compared to renal balloon angioplasty alone in patients with ostial stenosis.

Deterioration in renal function, albeit transient, is a common complication following endovascular renal artery intervention. This is most likely the combined result of the use of iodinated contrast and the occurrence of renal parenchymal embolism due to wire and catheter manipulation. In most cases, this is a temporary problem, as supportive care with adequate fluid hydration is sufficient to reverse the renal dysfunction. However, transient hemodialysis may become necessary in approximately 1% of patients. Other complications include vascular access complications (bleeding, hematoma, femoral nerve injury, arteriovenous fistula, and pseudoaneurysm), target vessel dissection, perinephric hematoma, early postoperative renal artery thrombosis, and extremity atheroembolism from thrombus in the aorta or the iliac arteries.

**Clinical Results of Endovascular Interventions**

**Percutaneous Transluminal Balloon Angioplasty.** FMD of the renal artery is the most common treatment indication for percutaneous transluminal balloon angioplasty. Patients with symptomatic FMD such as hypertension or renal insufficiency usually respond well to renal artery balloon angioplasty alone. In contrast, balloon angioplasty generally is not an effective treatment for patients with renal artery stenosis or proximal occlusive disease of the renal artery, due to the high incidence of restenosis with balloon angioplasty alone. In the latter group of patients, primary stent placement is the preferred endovascular treatment. The long-term benefit of renal artery balloon angioplasty in patients with FMD was reported by Surowiec and colleagues. They followed 14 patients who underwent 19 interventions on 18 renal artery segments. The technical success rate of balloon angioplasty for FMD was 95%. Primary patency rates were 81%, 69%, 69%, and 69% at 2, 4, 6, and 8 years, respectively. Assisted primary patency rates were 87%, 87%, 87%, and 87% at 2, 4, 6, and 8 years, respectively. The restenosis rate was 25% at 8 years. Clinical benefit, as defined by either improved or cured hypertension, was found in 79% of patients overall, with two-thirds of patients having maintained this benefit at 8 years. The authors concluded that balloon angioplasty is highly effective in symptomatic FMD with excellent durable functional benefits.

The utility of balloon angioplasty alone in the treatment of renovascular hypertension appears to be limited. van Jaarsveld and associates performed a prospective study in which patients with renal artery stenosis were randomized to either drug therapy or balloon angioplasty treatment. A total of 106 patients with 50% diameter stenosis or greater plus hypertension or renal insufficiency were randomized in the study. At 3 months, there was no difference in the degree to which blood
pressure was controlled between the two groups. However, the
degree and dose of antihypertensive medications were slightly
lowered in the balloon angioplasty group. The above advantage
of the angioplasty group completely disappeared at 12 months,
making the authors conclude that in the treatment of patients
with hypertension and renal artery stenosis, percutaneous trans-
luminal balloon angioplasty alone offers minimal advantage
over antihypertensive drug therapy.

Renal Artery Stenting. Endovascular stent placement is the
treatment of choice for patients with symptomatic or high-grade
renal artery occlusive disease (Fig. 23-48). This is due in part to
the high incidence of restenosis with balloon angioplasty alone,
particularly in the setting of ostial stenosis. Renal artery stenting
is also indicated for renal artery dissection caused by balloon
angioplasty or other catheter-based interventions. Numerous
studies have clearly demonstrated the clinical efficacy of renal
artery stenting when compared to balloon angioplasty alone in
patients with high-grade renal artery stenosis.

White and colleagues conducted a study to evaluate the role
of renal artery stenting in patients with poorly controlled hyper-
tension and renal artery lesions that did not respond well to bal-
loon angioplasty alone. The technical success of the procedure
was 99%. The mean blood pressure values were 173 ± 25/88 ±
17 mmHg prior to stent implantation and 146 ± 20/77 ± 12 mmHg
6 months after renal artery stenting (P <0.01). Angiographic
follow-up with 67 patients (mean 8.7 ± 5 months) demonstrated
that restenosis, as defined by 50% or greater luminal narrowing,
occurred in 15 patients (19%). The study concluded that renal
artery stenting is a highly effective treatment for renovascular
hypertension, with a low angiographic restenosis rate. In another
similar study, Blum and colleagues prospectively performed renal
artery stenting in 68 patients (74 lesions) with ostial renal artery
stenosis and suboptimal balloon angioplasty. Patients were fol-
lowed for a mean of 27 months with measurements of blood pres-
sure and serum creatinine, duplex sonography, and intra-arterial
angiography. Five-year patency was 84.5% (mean follow-up,
27 months). Restenosis occurred in 8 of 74 arteries (11%), but
after reintervention, the secondary 5-year patency rate was
92.4%. Hypertension was cured or improved in 78% of patients.
The authors concluded that primary stent placement is an effec-
tive treatment for renal artery stenosis involving the ostium.

The clinical utility of renal artery stenting in renal function
preservation was analyzed by several studies, which measured
serial serum creatinine levels to determine the response of
renal function following endovascular intervention. In a study
reported by Harden and colleagues who performed 33 renal
artery stenting procedures in 32 patients with renal insuffi-
cency, they noted that renal function improved or stabilized in
22 patients (69%). In a similar study, Watson and associates
evaluated the effect of renal artery stenting on renal function
by comparing the slopes of the regression lines derived from
the reciprocal of serum creatinine versus time. A total of 61
renal stenting procedures were performed in 33 patients, and
the authors found that after stent placement, the slopes of the
reciprocal of the serum creatinine (1/Ser) were positive in 18
patients and less negative in 7 patients. The study concluded that
in patients with chronic renal insufficiency due to obstructive
renal artery stenosis, renal artery stenting is effective in improv-
ing or stabilizing renal function.

The clinical outcome of several large clinical studies of
renal artery stenting in the treatment of renovascular hyperten-
sion or chronic renal insufficiency is shown in Table 23-13.
These studies uniformly demonstrated an excellent technical
success rate with low incidence of restenosis or procedural-
related complications. A similar analysis was reported by Leer-
touwer and colleagues who performed a meta-analysis of 14
studies comparing patients with renal arterial stent placement
to those who underwent balloon angioplasty alone for renal arterial
stenosis. The study found that stent placement proved highly
successful, with an initial technical success of 98%. The overall
cure rate for hypertension was 20%, whereas hypertension was
improved in 49%. Renal function improved in 30% of patients
and stabilized in 38% of patients. The restenosis rate at follow-
up of 6 to 29 months was 17%. Renal stenting resulted in a
higher technical success rate and a lower restenosis rate when
compared to balloon angioplasty alone.

AORTOILIAC OCCLUSIVE DISEASE

The distal abdominal aorta and the iliac arteries are common
sites affected by atherosclerosis. The symptoms and natural
history of the atherosclerotic process affecting the aortoiliac
arterial segment are influenced by the disease distribution and extent. Atherosclerotic plaques may cause clinical symptoms by restricting blood flow due to luminal obstruction or by embolizing atherosclerotic debris to the lower extremity circulation. If the aortoiliac plaques reach sufficient mass and impinge on the arterial lumen, obstruction of blood flow to lower extremities occurs. Various risk factors exist that can lead to the development of aortoiliac occlusive disease. Recognition of these factors and understanding of this disease entity will enable physicians to prescribe the appropriate treatment strategy, which may alleviate symptoms and improve quality of life.

**Diagnostic Evaluation**

On clinical examination patients often have weakened femoral pulses and a reduced ABI. Verification of iliac occlusive disease is usually made by color duplex scanning, which reveals either a peak systolic velocity ratio ≥2.5 at the site of stenosis and or a monophasic waveform. Noninvasive tests such as pulse volume recordings (PVRs) of the lower extremity with estimation of the thigh-brachial pressure index may be suggestive of aortoiliac disease. MRA and multidetector CTA are increasingly being used to determine the extent and type of obstruction. DSA offers the interventionalist the benefit of making a diagnosis and the option of performing an endovascular treatment in a single session. Angiography provides important information regarding distal arterial runoff vessels as well as the patency of the PFA. Presence of pelvic and groin collaterals is important to provide crucial collateral flow in maintaining lower limb viability. It must be emphasized, however, that patients should be subjected to angiography only if their symptoms warrant surgical intervention.

**Differential Diagnosis**

Degenerative hip or spine disease, lumbar disk herniation, spinal stenosis, diabetic neuropathy, and other neuromuscular problems can produce symptoms that may be mistaken for vascular claudication. Such cases can be distinguished from true claudication by the fact that the discomfort from neuromuscular problems is often relieved by sitting or lying down, as opposed to cessation of ambulation. In addition, complaints that are experienced upon standing suggest nonvascular causes. When confusion persists, the use of noninvasive vascular laboratory testing modalities, including treadmill exercise, can help establish the diagnosis.

**Collateral Arterial Network**

The principal collateral pathways in severe aortoiliac artery occlusive disease or chronic aortic occlusion that may provide blood flow distal to the aortoiliac lesion include: (a) the superior mesenteric artery to the distal IMA via its superior hemorrhoidal branch to the middle and inferior hemorrhoidalst to the internal iliac artery; (b) the lumbar arteries to the superior gluteal artery to the internal iliac system; (c) the lumbar arteries to the lateral and deep circumflex arteries to the CFA; and (d) Winslow’s pathway from the subclavian to the superior epigastric artery to the external iliac arteries at the groin (Fig. 23-49). In general, treatment indications for aortoiliac artery occlusive disease include disabling claudication, ischemic rest pain, nonhealing lower extremity tissue wound, and lower extremity microembolization that arises from aortoiliac lesions.

**Disease Classification**

Based on the atherosclerotic disease pattern, aortoiliac occlusive disease can be classified into three types (Fig. 23-50). Type I aortoiliac disease, which occurs in 5% to 10% of patients, is confined to the distal abdominal aorta and common iliac vessels (Fig. 23-51). Due to the localized nature of this type of aortic obstruction and formation of collateral blood flow around the occluded segment, limb-threatening symptoms are rare in the absence of more distal disease (Fig. 23-52). This type of aortoiliac occlusive disease occurs in a relatively younger group of patients (in their mid-50s), compared with patients who have more femoropopliteal disease. Patients with a type I disease pattern have a lower incidence of hypertension and diabetes but a significant frequency of abnormal blood lipid levels, particularly type IV hyperlipoproteinemia. Symptoms typically consist of bilateral thigh or buttock claudication and fatigue. Men report diminished penile tumescence and may have complete loss of erectile function. These symptoms in the absence of femoral
pulses constitute Leriche’s syndrome. Rest pain is unusual with isolated aortoiliac disease unless distal disease coexists. Occasionally patients report a prolonged history of thigh and buttock claudication that recently becomes more severe. It is likely that this group has underlying aortoiliac disease that has progressed to acute occlusion of the terminal aorta. Others may present with “trash foot,” which represents microembolization into the distal vascular bed (Fig. 23-53). Type II aortoiliac disease represents a more diffuse atherosclerotic progression that involves predominately the abdominal aorta with disease extension into the common iliac artery. This disease pattern affects approximately 25% patients with aortoiliac occlusive disease. Type III aortoiliac occlusive disease, which affects approximately 65% of patients with aortoiliac occlusive disease, is widespread disease that is seen above and below the inguinal ligament (Fig. 23-54). Patients with “multilevel” disease are older, more commonly male (with a male-to-female ratio of 6:1), and much more likely to have diabetes, hypertension, and associated atherosclerotic disease involving cerebral, coronary, and visceral arteries. Progression of the occlusive process is more likely in these patients than in those with localized aortoiliac disease. For these reasons, most patients with a type III pattern tend to present with symptoms of advanced ischemia and require revascularization for limb salvage rather than for claudication. These patients have a decreased 10-year life expectancy when compared to patients with localized aortoiliac disease.

The most commonly used classification system of iliac lesions has been set forth by the TransAtlantic Inter-Society Consensus (TASC) group with recommended treatment options. This lesion classification categorizes the extent of atherosclerosis and has suggested a therapeutic approach based on this classification (Table 23-14 and Fig. 23-55).

According to this consensus document, endovascular therapy is the treatment of choice for type A lesions, and surgery is the treatment of choice for type D lesions. Endovascular treatment is the preferred treatment for type B lesions, and surgery is the preferred treatment for good-risk patients with type C lesions. In comparison to the 2000 TASC document, the commission has not only made allowances for treatment of more extensive lesions, but it also takes into account the continuing evolution of endovascular technology and the skills of individual interventionalists when stating that the patient’s comorbidities, fully informed patient preference, and the local operator’s long-term success rates
must be considered when making treatment decisions for type B and type C lesions.1,115

**General Treatment Considerations**

There is no effective medical therapy for the management of aortoiliac disease, but control of risk factors may help slow progression of atherosclerosis. Patients should have hypertension, hyperlipidemia, and diabetes mellitus controlled. They should be advised to stop smoking. Most patients are empirically placed on antiplatelet therapy. A graduated exercise program may improve walking efficiency, endothelial function, and metabolic adaptations in skeletal muscle, but, there is usually minimal improvement in patients with aortoiliac disease who are treated with these measures. Failure to respond to exercise and/or drug therapy should prompt consideration for limb revascularization. Patients with buttock claudication and reduced or absent femoral pulses who fail to respond to exercise and drug therapy should be considered for revascularization because they are less likely than patients with more distal lesions to improve without concomitant surgical or endovascular intervention.
Surgical Reconstruction of Aortoiliac Occlusive Disease

Aortobifemoral Bypass. Surgical options for treatment of aortoiliac occlusive diseases consist of various configurations of aortobifemoral bypass grafting, various types of extra-anatomic bypass grafts, and aortoiliac endarterectomy. The procedure performed is determined by several factors, including anatomic distribution of the disease, clinical condition of the patient, and personal preference of the surgeon.

In most cases, aortobifemoral bypass is performed because patients usually have disease in both iliac systems. Although one side may be more severely affected than the other, progression does occur, and bilateral bypass does not complicate the procedure or add to the physiologic stress of the operation. Aortobifemoral bypass reliably relieves symptoms, has excellent long-term patency (approximately 70–80% at 10 years), and can be completed with a tolerable perioperative mortality (2–3%).

Technical Considerations for Aortobifemoral Bypass. Both femoral arteries are initially exposed to ensure that they are adequate for the distal anastomoses. The abdomen is then opened in the midline, the small intestine is retracted to the right, and the posterior peritoneum overlying the aorta is incised. A retroperitoneal approach may be selected as an alternative in certain situations. This approach involves making a left flank incision and displacing the peritoneum and its contents to the right. Such an approach is contraindicated if the right renal artery is acutely occluded, since visualization from the left flank is very poor. Tunneling of a graft to the right femoral artery is also more difficult from a retroperitoneal approach, but can be achieved. The retroperitoneal approach has been reputed to be better tolerated than midline laparotomy for patients with multiple previous abdominal operations and with severe pulmonary disease. Further proposed advantages of the retroperitoneal approach include less gastrointestinal disturbance, decreased third space

### Table 23-14

TASC classification of aortoiliac occlusive lesions

<table>
<thead>
<tr>
<th>Type A lesions</th>
<th>Type B lesions</th>
<th>Type C lesions</th>
<th>Type D lesions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unilateral or bilateral stenoses of CIA</td>
<td>Short (≤3 cm) stenosis of infrarenal aorta</td>
<td>Unilateral CIA occlusion</td>
<td>Infrafemoral aortoiliac occlusion</td>
</tr>
<tr>
<td>Unilateral or bilateral single short (≤3 cm) stenosis of EIA</td>
<td>Single or multiple stenoses totaling 3–10 cm involving the EIA not extending into the CFA</td>
<td>Unilateral CIA occlusion</td>
<td>Diffuse disease involving the aorta and both iliac arteries requiring treatment</td>
</tr>
<tr>
<td>Unilateral or bilateral stenoses 3–10 cm long not extending into the CFA</td>
<td>Unilateral CIA stenosis extending into the CFA</td>
<td>Unilateral EIA occlusion involving the origins of internal iliac artery and/or CFA</td>
<td>Diffuse multiple stenoses involving the unilateral CIA, EIA, and CFA</td>
</tr>
<tr>
<td>Unilateral EIA occlusion not involving the origins of internal iliac artery or CFA</td>
<td>Unilateral EIA stenosis involving the origins of internal iliac artery and/or CFA</td>
<td>Heavily calcified unilateral EIA occlusion with or without involvement of origins of internal iliac artery and/or CFA</td>
<td>Unilateral occlusions of both CIA and EIA</td>
</tr>
<tr>
<td>Unilateral occlusions of EIA</td>
<td>Iliac stenoses in patients with AAA requiring treatment</td>
<td>Bilateral occlusions of EIA</td>
<td>Bilateral CIA occlusions</td>
</tr>
</tbody>
</table>
| Iliac stenoses in patients with AAA requiring treatment and not amenable to endograft placement or other lesions requiring open aortic or iliac surgery | AAA = abdominal aortic aneurysm; CFA = common femoral artery; CIA = common iliac artery; EIA = external iliac artery.

**Figure 23-55.** Schematic depiction of the TransAtlantic Inter-Society Consensus classification of aortoiliac occlusive lesions.
fluid losses, and ease with which the pararenal aorta can be accessed. There are randomized reports, however, that support and refute the superiority of this approach. A collagen-impregnated, knitted Dacron graft is used to perform the proximal aortic anastomosis, which can then be made in either an end-to-end or end-to-side fashion using 3-0 polypropylene suture. The proximal anastomosis should be made as close as possible to the renal arteries to decrease the incidence of restenosis from progression of the atherosclerotic occlusive process in the future.

An end-to-end proximal aortic anastomosis is necessary in patients with an aortic aneurysm or complete aortic occlusion extending up to the renal arteries (Fig. 23-56). Although in theory the end-to-end configuration allows for less turbulence and less chance of competitive flow with still patent host iliac vessels, there have not been consistent results to substantiate differences in patency between end-to-end and end-to-side grafts. Relative indications for an end-to-side proximal aortic anastomosis include the presence of large aberrant renal arteries, an unusually large IMA with poor back-bleeding suggesting inadequate collateralization, and/or occlusive disease involving bilateral external iliac arteries. Under such circumstances, end-to-end bypass from the proximal aorta to the femoral level devascularizes the pelvic region because there is no antegrade or retrograde flow in the occluded external iliac arteries to supply the hypogastric arteries. As a result of the pelvic devascularization, there is an increased incidence of impotence, postoperative colon ischemia, buttock ischemia, and paraplegia secondary to spinal cord ischemia despite the presence of excellent femoral and distal pulses.

An end-to-side proximal aortic anastomosis can be associated with certain disadvantages, which include the potential for distal embolization when applying a partially occlusive aortic clamp (Fig. 23-57). Furthermore, the distal aorta often proceeds to total occlusion after an end-to-side anastomosis. There may also be a higher incidence of aortoenteric fistula following construction of end-to-side proximal anastomoses because the anterior projection makes subsequent tissue coverage and reperitonealization of the graft more difficult. The limbs of the graft are tunneled through the retroperitoneum to the groin, where an end-to-side anastomosis is fashioned between the graft and the bifurcation of the CFA using 5-0 polypropylene suture. Endarterectomy or patch angioplasty of the profunda femoris may be required concurrently. Once the anastomoses have been fashioned and the graft thoroughly flushed, the clamps are removed and the surgeon carefully controls the degree of aortic occlusion until full flow is reestablished. During this period, the patient must be carefully monitored for hypotension. Declamping hypotension is a complication of sudden restoration of aortic flow, particularly following prolonged occlusion. Once flow has been reestablished, the peritoneum is carefully reapproximated over the prosthesis to prevent fistulization into the intestine.

Despite the presence of multilevel disease in most patients, a properly performed aortobifemoral operation can provide arterial inflow and alleviate claudication symptoms in 70% to 80% of patients; however, 10% to 15% of patients will require simultaneous outflow reconstruction to address distal ischemia and facilitate limb salvage. The advantage of concomitant distal revascularization is avoidance of reoperation in a scarred groin. As a rule, if the profunda femoris can accept a 4-mm probe and if a No. 3 Fogarty embolectomy catheter can be passed distally for 20 cm or more, the PFA will be sufficient for outflow, and concomitant distal revascularization is not necessary.

**Aortic Endarterectomy.** Aortoiliac endarterectomy is rarely performed because it is associated with greater blood loss and greater sexual dysfunction and is more difficult to perform. Long-term patency is comparable with aortobifemoral grafting, and thus it remains a reasonable option in cases in which the risk of infection of a graft is excessive because it involves no prosthetic tissue. Aortoiliac endarterectomy was useful when disease was localized to either the aorta or common iliac arteries; however, at present, aortoiliac PTA, stents, and other catheter-based therapies have become first-line treatment in this scenario. Endarterectomy should not be performed if the aorta is aneurysmal because of continued aneurysmal degeneration of the endarterectomized segment. If there is total occlusion of the aorta to the level of the renal arteries, aortic transection several centimeters below the renal arteries with thrombectomy of the aortic cuff followed by graft insertion is easier and more expeditious when compared to endarterectomy. Involvement of the external iliac artery makes aortic endarterectomy more difficult to complete because of decreased vessel diameter, increased length, and exposure issues. The ability to establish an appropriate

---

**Figure 23-56.** In an end-to-end proximal aortic anastomosis, the aorta is divided in half. The proximal end of the aorta is anastomosed to the end of a prosthetic graft, while the distal divided aortic stump is oversewn.

**Figure 23-57.** In an end-to-side aortic anastomosis, the end of a prosthetic graft is connected to the side of an aortic incision.
endarterectomy plane is compromised due to the muscular and inherently adherent nature of the media in this location. There is a higher incidence of early thrombosis and late failure with extended aortoiliofemoral endarterectomy when compared to bypass grafting as a result of recurrent stenosis.

**Axillofemoral Bypass.** An axillofemoral bypass is an extra-anatomic reconstruction that derives arterial inflow from the axillary artery to the femoral artery. This is a treatment option for patients with medical comorbidities that prohibit an abdominal vascular reconstruction. It may be performed under local anesthesia and is used for limb salvage. Extra-anatomic bypasses have lower patency when compared to aortobifemoral and, therefore, are seldom recommended for claudication. Before performing this operation, the surgeon should check pulses and blood pressure in both arms to ensure that there is no obvious disease affecting flow through the axillary system. Angiography of the axillosubclavian vasculature is not necessary, but can be helpful if performed at the time of aortography. The axillary artery is exposed below the clavicle, and a 6- to 8-mm externally reinforced PTFE graft is tunneled subcutaneously down the lateral chest wall and lateral abdomen to the groin. It is anastomosed ipsilaterally at the CFA bifurcation into the SFA and PFA. A femorofemoral crossover graft using a 6- to 8-mm externally reinforced PTFE graft is then used to revascularize the opposite extremity if necessary. Reported patency rates over 5 years vary from 30% to 80%.\(^{116}\) Paradoxically, although it is a less complex procedure than aortofemoral grafting, the mortality rate is higher (10%), reflecting the compromised medical status of these patients.

**Iliofemoral Bypass.** One option for patients with unilateral occlusion of the distal common iliac or external iliac arteries is iliufemoral grafting (Fig. 23-58). Long-term patency is comparable to aortounifemoral bypass, and because the procedure can be performed using a retroperitoneal approach without clamping the aorta, the perioperative mortality is less.\(^{116}\)

**Femorofemoral Bypass.** A femorofemoral bypass is another option for patients with unilateral stenosis or occlusion of the common or external iliac artery who have rest pain, tissue loss, or intractable claudication. The primary (assisted) patency at 5 years is reported to be 60% to 70%, and although this is inferior when compared to aortofemoral bypass, there are physiologic benefits, especially for patients with multiple comorbidities because it is not necessary to cross-clamp the aorta.\(^{117}\) There are no studies supporting the superiority of unsupported or externally supported PTFE over Dacron for choice of conduit. The fear of the recipient extremity stealing blood from the extremity ipsilateral to the donor limb is not realized unless the donor iliac artery and donor outflow arteries are diseased. Depending on the skills of the interventionalist or surgeon, many iliac lesions classified as TASC B, C, or D can now be addressed using an endovascular approach, thus obviating the need to perform a femorofemoral bypass. Additionally, femorofemoral bypass can be used as an adjuvant procedure after iliac inflow has been optimized with endovascular methods.

**Obturator Bypass.** An obturator bypass is used to reconstruct arterial anatomy in patients with groin sepsis resulting from prior prosthetic grafting, intraarterial drug abuse, groin neoplasm, or damage from prior groin irradiation. This bypass can originate from the common iliac artery, external iliac artery, or uninvolved limb of an aortobifemoral bypass. A conduit of Dacron, PTFE, or autologous vein is tunneled through the anteromedial portion of the obturator membrane to the distal superficial femoral artery or popliteal artery. The obturator membrane must be divided sharply so as avoid injury to adjacent structures, and care must be taken to identify the obturator artery and nerve that pass posterolaterally. After the bypass is completed and the wounds isolated, the infected area is entered, the involved arteries are debrided to healthy tissue, and vascularized muscle flaps are mobilized to cover the ligated ends. There have been varied results in terms of patency and limb salvage for obturator bypass. Some authors have reported 57% 5-year patency and 77% 5-year limb salvage rates, whereas others have shown a high rate of reinfecption and low patency requiring reintervention.\(^{118,119}\)

**Thoracofemoral Bypass.** The indications for thoracofemoral bypass are (a) multiple prior surgeries with a failed infrarenal aortic reconstruction and (b) infected aortic prosthesis. This procedure is more physiologically demanding than other extra-anatomic reconstructions because the patient must not tolerate clamping the descending thoracic aorta but also performance of a left thoracotomy. The graft is tunneled to the left CFA from the left thorax posterior to the left kidney in the anterior axillary line using a small incision in the periphery of the diaphragm and an incision in the left inguinal ligament to gain access to the extraperitoneal space from below. The right limb is tunneled in the space of Retzius in an attempt to decrease kinking that is more likely to occur with subcutaneous, suprapubic tunneling. Thoracofemoral bypass has long-term patency comparable to aortofemoral bypass.

**Complications of Surgical Aortoiliac Reconstruction**

With current surgical techniques and conduits, early postoperative hemorrhage is unusual and occurs in 1% to 2%, which is usually the result of technical oversight or coagulation abnormality.\(^{120}\) Acute limb ischemia occurring after aortoiliac surgery may be the result of acute thrombosis or distal thromboembolism. The surgeon can prevent thromboembolic events by (a) avoiding excessive manipulation of the aorta, (b) ensuring

---

**Figure 23-58.** A. Skin markings showing the incisions of an iliofemoral bypass. B. A prosthetic bypass graft is used for an iliofemoral artery bypass in which the proximal anastomosis is connected to the common iliac artery (long arrow) while the distal anastomosis is connected to the common femoral artery (short arrow).
adequate systemic heparinization, (c) judicious placement of vascular clamps, and (d) thorough flushing prior to restoring blood flow. Acute thrombosis of an aortofemoral graft limb in the early perioperative period occurs in 1% to 3% of patients. Thrombectomy of the graft limb is performed through a transverse opening in the hood of the graft at the femoral anastomosis. With this approach, it is possible to inspect the interior of the anastomosis and pass embolectomy catheters distally to clear the superficial femoral and profunda arteries. Various complications may be encountered following aortoiliac or aortobifemoral reconstruction (Table 23-15).

Intestinal ischemia following aortic reconstruction occurs in approximately 2% of cases; however, with colonoscopy mucosal ischemia, which is a milder form, is seen more frequently. The surgeon can identify patients who require concomitant revascularization of the IMA, hypogastric arteries, or mesenteric arteries by examining the preoperative arteriogram for the presence of associated occlusive lesions in the celiac axis, the superior mesenteric arteries, or both. Likewise, patients with a patent and enlarged IMA or a history of prior colonic resections will benefit from IMA reimplantation.

In a comprehensive review of 747 patients who had aortoiliac operations for occlusive disease, secondary operations for late complications such as reocclusion, pseudoaneurysms, and infection were necessary in 21% over a 22-year period. The most frequent late complication is graft thrombosis. Limb occlusion occurs in 5% to 10% of patients within 5 years of the index operation and in 15% to 30% of patients 10 years after the index operation. Anastomotic pseudoaneurysms occur in 1% and 5% of femoral anastomoses in patients with aortofemoral grafts. Predisposing factors to pseudoaneurysm formation include progression of degenerative changes within the host artery, excessive tension at the anastomosis, and infection. Due to the associated risks of thrombosis, distal embolization, infection, and rupture, anastomotic aneurysms should be repaired expeditiously.

Infection following aortoiliac reconstruction is a devastating complication that occurs in 1% of cases. Femoral anastomoses of aortofemoral reconstructions and axillofemoral bypasses are prone to infection. Use of prophylactic antibiotics and meticulous surgical technique are vital in preventing contamination of the graft at the time of implantation. If infection appears localized to a single groin, graft preservation and local measures such as antibiotic irrigation, aggressive debridement, and soft tissue coverage with rotational muscle flaps may prove successful. Most patients with infected aortoiliacofemoral reconstructions usually require graft excision and revascularization via remote uncontaminated routes or the use of in situ replacement to clear the infective process and maintain limb viability. Aortoenteric fistula and associated gastrointestinal hemorrhage are devastating complications, with a 50% incidence of death or limb loss. The incidence of aortoenteric fistula formation appears to be higher after an end-to-side proximal anastomosis because it is more difficult to cover the prosthesis with viable tissue and avoid contact with the gastrointestinal tract with this configuration. Treatment of aortoenteric fistula requires resection of all prosthetic material, closure of the infrarenal abdominal aorta, repair of the gastrointestinal tract, and revascularization by means of an extra-anatomic graft.

### Endovascular Treatment for Aortic Disease

Although aortofemoral bypass surgery has excellent long-term patency and can be performed with low mortality rates, there are patients who are unable to withstand the physiologic stress of longer open procedures performed under general anesthesia, which require aortic cross-clamping and which are associated with greater blood loss. These patients are more suited to endovascular interventions despite the decreased durability and requirement for more frequent reinterventions.

**Focal Aortic Stenosis.** The endovascular technique used to treat infrarenal aortic stenoses is similar to that used for iliac artery disease. Bilateral CFA access is established followed by insertion of a 6-French sheath. The lesion is crossed using a hydrophilic wire and a supporting selective catheter and then changed for a stiffer guidewire. A self-expanding nitinol stent or a balloon-expandable stent mounted on a larger-caliber angioplasty balloon is implanted followed by adequate postdilation. At the physician’s discretion, “kissing” stents, simultaneous bilateral proximal iliac stents, are deployed if the lesion is in the distal aorta in the proximity of the aortic bifurcation. The role of covered stents such as cuffs made for endoluminal AAA repair has not been rigorously studied. The aortic diameter should be sized with a calibrated catheter during the angiography or by preintervention CT scanning to avoid undersizing. Balloon size will range from 12 to 18 mm in most cases. A single stent is generally sufficient in most cases. Concentric aortic stenosis may encroach upon the IMA, and coverage of this vessel may be unavoidable. Care should be taken to use low inflation pressures (5 mmHg) to minimize the risk of aortic rupture. Patient complaints of back or abdominal pain during balloon inflation should be taken seriously as they may suggest impending rupture. In case of a calcified small-caliber, hypoplastic aorta (≤12 mm, typically in female patients), it is recommended to

---

### Table 23-15

<table>
<thead>
<tr>
<th>Perioperative complications of aortobifemoral bypass grafting</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Medical Complications</strong></td>
</tr>
<tr>
<td>• Perioperative myocardial infarction</td>
</tr>
<tr>
<td>• Respiratory failure</td>
</tr>
<tr>
<td>• Ischemia-induced renal failure</td>
</tr>
<tr>
<td>• Bleeding from intravenous heparinization</td>
</tr>
<tr>
<td>• Stroke</td>
</tr>
<tr>
<td><strong>Procedure-Related Complications</strong></td>
</tr>
<tr>
<td><em>Early</em></td>
</tr>
<tr>
<td>• Declamping shock</td>
</tr>
<tr>
<td>• Graft thrombosis</td>
</tr>
<tr>
<td>• Retroperitoneal bleeding</td>
</tr>
<tr>
<td>• Groin hematoma</td>
</tr>
<tr>
<td>• Bowel ischemia/infarction</td>
</tr>
<tr>
<td>• Peripheral embolization</td>
</tr>
<tr>
<td>• Erectile dysfunction</td>
</tr>
<tr>
<td>• Lymphatic leak</td>
</tr>
<tr>
<td>• Chylous ascites</td>
</tr>
<tr>
<td>• Paraplegia</td>
</tr>
<tr>
<td><em>Late</em></td>
</tr>
<tr>
<td>• Graft infection</td>
</tr>
<tr>
<td>• Anastomotic pseudoaneurysm</td>
</tr>
<tr>
<td>• Aortoenteric fistula</td>
</tr>
<tr>
<td>• Aortourinary fistula</td>
</tr>
<tr>
<td>• Graft thrombosis</td>
</tr>
</tbody>
</table>
use smaller diameter stents. To achieve clinical improvement, these patients can be recanalized to an aortic diameter of 8 or 9 mm. Distal embolization is one of the potential complications of endovascular treatment for aortic stenoses. Full heparinization, meticulous technique during wire and catheter manipulations, and primary stenting reduce the risk of this complication. Since calcified aortic stenoses are prone to rupture during dilatation, it is recommended to be cognizant of the extent of the calcification with preoperative CT scans. In case of aortic rupture, as long as wire access has been maintained, an occlusion balloon can be inflated proximal to the disrupted segment to achieve hemostasis, and the rupture can be covered with a stent graft or repaired with open surgery.

**Occlusive Lesions of the Aortic Bifurcation.** Occlusive lesions are treated with the kissing balloon technique to avoid dislodging aortic plaque. Two angioplasty balloons of equal size are positioned across the ostia of the common iliac arteries, using a retrograde approach, and inflated. Simultaneous balloon dilatation at the origins of both common iliac arteries is advocated, even in the presence of unilateral lesion, to protect the contralateral common iliac artery from dissection or plaque embolization. Calcified lesions that typically occur at the aortic bifurcation are not amenable to balloon dilatation and frequently require that a distal aortic reconstruction be performed using “kissing stents.” Fears that the proximal ends of the stents that extend into the distal aorta will become a nidus for thrombus formation or cause hemolysis have not been realized. The results are difficult to interpret because these bifurcation lesions are usually included in studies with iliac artery lesions. Patency rates for aortic bifurcation PTA range from 76% to 92% at 3 years. The largest series reported to date includes 79 patients with aortic bifurcation lesions. The cumulative clinical success rate at a mean of 4 years was 93%.122 Stents have also been used to reconstruct the aortic bifurcation with high success rates. The kissing stent technique is well suited for orificial lesions. Technical success with kissing stents at the aortic bifurcation has been reported to be 95% to 100%.123 In the largest series reported, the primary patency at 3 years was 79%.124

**Endovascular Treatment for Iliac Artery Disease**

**Percutaneous Transluminal Angioplasty.** PTA is most useful in the treatment of isolated iliac stenoses of less than 4 cm in length. When used for stenoses rather than occlusion, a 2-year patency of 86% can be achieved.125 The complication rate is approximately 2%, consisting of distal embolization, medial dissection, and acute thrombosis.

**Technical Considerations for Iliac Interventions** Crossing a high-grade stenosis or occlusion can be challenging in the iliac arteries. It is vital to image the lesion well because multiple views and use of the image intensifier will frequently uncover the anatomic reason for the difficulty. Frequently, the difficulty is the result of vessel tortuosity that cannot be appreciated on the original view. Use of an angled hydrophilic guidewire and an angled catheter can provide steering and add extra support for the wire trying to cross the lesion. Patience, persistence, and periodic reimaging will facilitate the crossing of a lesion in the great majority of cases. Guidewire traversal must be achieved for performance of endovascular iliac intervention. Over 90% of iliac occlusions can be passed with simple guidewire techniques. The preferred approach for recanalizing a common iliac artery occlusion is retrograde passage of devices from an ipsilateral CFA puncture because, in this manner, distance to the lesion is short and access is straighter. A stenosis is normally crossed using a combination of a soft-tip 0.035-inch guidewire (i.e., Benton-type wire) or hydrophilic wire and a 5-French straight or selective catheter. One of the hazards of retrograde recanalization is that the guidewire stays in a subintimal location and cannot be redirected into the true lumen at the aortic bifurcation. There are several approaches that can be used to achieve reentry of total chronic occlusions. Specialized catheters allow passage of a needle and guidewire across the intima distal to the occlusion. Intravascular ultrasound can be used for true lumen reentry under fluoroscopic guidance. Another method of achieving true lumen reentry involves performing the recanalization from an antegrade contralateral CFA approach. A 4-French Berenstein catheter (Coridis Corp., Miami Lakes, FL) is used to probe the occlusion. The lesion can be crossed in most instances (5–20% failure rate) with a hydrophilic guidewire or occasionally with its stiffer back end. As soon as the guidewire has crossed the obstruction and lies within the ipsilateral external iliac artery lumen, it is snared and partially pulled out of the ipsilateral CFA. A short catheter is then inserted in a retrograde fashion over the wire end into the abdominal aorta proximal to the lesion. The hydrophilic guidewire is then exchanged for a stiffer Amplatz (Boston Scientific, Natick, MA) guidewire to facilitate iliac stenting.

Obtaining arterial access when there are absent femoral pulsations is aided by the use of ultrasound guidance and “roadmap” imaging software, which is available on modern angiographic equipment. When the lesion is successfully crossed, balloons of an appropriate size and length are selected for the angioplasty. Most common iliac arteries will accommodate 8- to 10-mm diameter balloons, whereas most external iliac arteries will accommodate 6- to 8-mm diameter balloons. Inflation is performed with caution, especially if there is heavy calcification, and should be guided by patient discomfort, pressure gauge readings, and changes in balloon outline.

If guidewire traversal is straightforward, consideration should be given to the presence of an acute thrombosis that may benefit from catheter-directed thrombolysis. If guidewire traversal is challenging, it is unlikely that catheter-directed thrombolysis will be beneficial. Stents should be placed after inadequate angioplasty. Stents are warranted when there is a greater than 30% residual stenosis, when there is a flow-limiting dissection, or when there is a pressure gradient of ≥2 mmHg across the treated segment. Placement of stents can precipitate distal embolization in up to 10%, especially if lesions are friable and vulnerable to manipulation. Routine primary stent placement is not recommended because it has not been found to be superior to selective stenting in terms of outcomes or cost.

**Primary Stenting Versus Selective Stenting in Iliac Arteries.** Primary stenting rather than selective stenting should be considered for longer iliac lesions and for all TASC C and D lesions. The primary patency rates at 1, 2, and 3 years were 96%, 90%, and 72%, respectively, for longer lesions (>5 cm) that were primarily stented versus 46%, 46%, and 28%, respectively, with selective stenting.126 Primary stenting is generally advocated for chronic iliac artery occlusions, recurrent stenosis after previous iliac PTA, and complex stenoses with eccentric, calcified, ulcerated plaques or plaques with spontaneous dissection. All of these lesions are prone to distal
embolization during manipulation of wires and angioplasty balloons. Distal embolization with isolated PTA is not common for uncomplicated lesions, but can occur in up to 24% of cases, when treating ulcerated plaques, aortic/iliac bifurcation lesions, or iliac occlusions. It is believed that direct stent placement without predilatation significantly reduces the risk of distal embolization by trapping potentially embolic material between the arterial wall and the stent mesh. While PTA has demonstrated excellent results in focal stenoses of the abdominal aorta and iliacs, primary stenting in these locations is safe, improves patency rates, reduces the degree of restenosis when compared with PTA alone, and decreases the risk of distal embolization. Additional potential advantages of direct stenting include shorter procedural time and less radiation exposure. The Dutch Iliac Stent Trial has provided evidence that refutes the superiority of primary stenting over angioplasty alone.151 Most interventionists continue to perform angioplasty first and stent selectively for inadequate results. The approach to aortoiliac stenting is intuitive. Individual judgment and experience are important in the decision-making process, and there are lesions with unstable morphology such as long occlusions, ulceration, and dissection that warrant primary stenting.

Stent Graft Placement for Aortoiliac Interventions. Stent grafts have been used to treat complex iliac lesions in an attempt to exclude these sources of embolization. A recent report suggested that the use of stent grafts was beneficial for TASC C and D lesions.127 Bosiers and colleagues published a series of 91 limbs with diseased iliacs that they treated with 107 stent grafts. They reported successful deployment in all patients without distal embolization or vessel rupture and a primary patency rate of 91.1% at 1 year.128 The authors commented on their concerns of causing embolization during placement of the stent grafts and recommended that once an occlusion was traversed with the guidewire, to gently predilate with a 5-mm balloon, followed by smooth stent graft insertion into the newly created channel. The role of stent grafts in aortoiliac occlusive disease has not been fully elucidated yet.

Complications of Endovascular Aortoiliac Interventions

Iliac artery angioplasty is associated with a 2% to 4% major complication rate and 4% to 15% minor complication rate. Many of these minor complications are related to the arterial puncture site. The most frequent complications relate to access site cannulation. Hemorrhage can range from the more common access site hematoma to the rarer retroperitoneal and intraperitoneal hemorrhage. Distal embolization occurs in 2% to 10% of iliac PTA and stenting procedures.129 Percutaneous catheter aspiration should be the initial treatment for calf vessel embolization, but, for larger emboli, such as those that lodge in the profunda femoris or common femoral arteries, surgical embolectomy may be required because the embolic material contains atherosclerotic plaque, which is not amenable to transcatheter aspiration or catheter-directed thrombolysis. The incidence of pseudoaneurysm formation at the puncture site is 0.5%. The treatment of choice for pseudoaneurysms >2 cm in diameter is percutaneous thrombin injection under ultrasound guidance. Arterial rupture may complicate the procedure in 0.3% of cases. Tamponade of the ruptured artery with an occlusion balloon should be performed, and a covered stent should be placed. In case of failure, surgical treatment is required.

Clinical Results Comparing Surgical and Endovascular Treatment of Aortoiliac Disease

The mortality risk of aortobifemoral bypass in patients with isolated, localized aortoiliac disease is relatively low, whereas for patients with concomitant atherosclerosis in coronary, carotid, and visceral vessels, mortality and morbidity are higher. For this reason, the cumulative long-term survival rate for patients receiving aortoiliac reconstruction remains 10 to 15 years less than anticipated for a normal age- and sex-matched population. Twenty-five percent to 30% of patients with concomitant atherosclerosis in other vascular distributions are dead within 5 years, and 50% to 60% will have died by 10 years.129

Compared with conventional aortobifemoral bypass, common iliac angioplasty was shown to have a 10% to 20% lower overall patency rate. It should be noted that these results were reported in early trials that used older generations of endovascular equipment. With continued progress and newer angioplasty balloons and stenting practices, more comparable outcomes are being reported. Review of the literature confirms that there is an 85% to 90% graft patency rate at 5 years and a 70% to 75% graft patency rate at 10 years after aortobifemoral reconstruction.132 Due in part to factors including continued refinements in anesthetic management, intraoperative monitoring, and postoperative intensive care, low perioperative mortality rates for aortobifemoral bypass can be achieved commonly in today’s clinical practice. The most recent systematic review and meta-analysis of 5358 patients who underwent direct open bypass or endovascular treatment for aortoiliac occlusive disease demonstrated superior durability for open bypass, although with longer length of stay and increased risk for complications and mortality, when compared to the endovascular approach.130 In this study, poor preoperative runoff was greater in the open bypass group (50.0% vs. 24.6%). Mean length of hospital stay was 13 days for open bypass versus 4 days for endovascular treatment procedures. The open bypass group experienced more complications (18.0% vs. 13.4%) and greater 30-day mortality (2.6% vs. 0.7%). At 1, 3, and 5 years, pooled primary patency rates were greater in the open bypass group (94.8% vs. 86.0%, 86.0% vs. 80.0%, and 82.7% vs. 71.4%, respectively); the same was true for secondary patency (95.7% vs. 90.0%, 91.5% vs. 86.5%, and 91.0% vs. 82.5%, respectively).

Despite its lower long-term success, common iliac angioplasty is a useful procedure in patients with focal disease and mild symptoms in whom a major surgical revascularization is not justified. Angioplasty of the iliac vessels can be a useful adjunct to distal surgical bypass as well, increasing the success of distal revascularization and eliminating the risks associated with aortoiliac bypass. Thus, with long-term patency less than, but comparable to, open surgical bypass, and with more favorable morbidity rates, iliac angioplasty has become a well-accepted modality of treatment for iliac occlusive disease. Ideal iliac angioplasty lesions are nonocclusive and short. Patency after intervention is better when lesions occur in larger diameter vessels, when stenoses rather than occlusions are treated, when runoff vessels are patent, and when the indication for intervention is lifestyle-limiting claudication rather than critical limb ischemia.

Becker and colleagues estimated a 5-year patency rate of 72% in an analysis of 2697 cases of iliac angioplasty and noted a better patency (79%) in claudicants.131 Less favorable results are obtained with long stenoses, external iliac stenoses, and tandem lesions. The reported technical and initial clinical success of balloon angioplasty in iliac artery stenoses exceeds 90% in
most series, and the 5-year patency rates range from 54% to 92%. The reported technical and initial clinical success of balloon angioplasty in iliac artery occlusions ranges from 78% to 98%, and the 3-year patency rates range from 48% to 85%.131,132

Factors reported to affect the patency of aortoiliac endovascular interventions adversely include quality of runoff vessels, severity of ischemia, and length of diseased segments treated. Likewise, as vessel diameter and flow rates change, so do success rates after angioplasty. It was reported in the literature that location of the lesion at the external iliac artery adversely affects both primary and assisted-primary patency. Following angioplasty of the common iliac artery, patency rates were 81% and 52% at 1 and 6 years, respectively; whereas, after external iliac artery angioplasty, they were 74% and 48% at 1 and 4 years, respectively.133 Although some literature supports location of the lesion in the external iliac artery as a factor that adversely affects both primary and assisted-primary patency, this has not been a universal finding. Female patients are also reported to have lower patency rates than males following iliac PTA, with or without stent placement in the external iliac artery.134

Stenting of the iliac arteries provides a durable and curative treatment, with a 3-year patency rate of 41% to 92% for stenosis and a 3-year patency rate of 64% to 85% and 4-year patency rate of 54% to 78% for occlusions.132 A meta-analysis of 2116 patients by Bosch and Hunink showed that aortoiliac stenting resulted in a 39% improvement in long-term patency compared to balloon angioplasty, despite the fact that complication rates and 30-day mortality rates did not differ significantly.135 Park and colleagues presented long-term follow-up results in a cohort of patients with all four TASC types of iliac lesions. The authors presented primary patency rates of 87%, 83%, 61%, and 49% at 3, 5, 7, and 10 years, respectively, after the index intervention.136 Levine and colleagues achieved primary and secondary patency rates of 76% and 90%, respectively, after 3 years, in a cohort of patients who received stents for iliac occlusions.137 The authors postulated that endovascular treatment for iliac occlusive disease should be extended to type C and D lesions, because they observed no detectable differences between the four TASC classifications in terms of primary and secondary patency rates.137 They concluded that presence of TASC C and D lesions should not preclude endovascular treatment and believe that endovascular attempts should be exhausted before open surgical repair of iliac occlusions is attempted because of the decreased perioperative morbidity and good midterm durability.

Not all results have been in favor of stenting, and at present, universal primary stenting cannot be recommended. Although stents are often used to improve the outcome of PTA, there is no general consensus that stenting should be mandatory in all iliac lesions. Complex, ulcerated iliac lesions with high embolicogenic potential or recanalized chronic iliac occlusions may be an exception. In the Dutch Iliac Stent Trial, primary stenting did not prove to be superior to iliac angioplasty and selective stenting.138 The researchers in this prospective randomized multicenter study concluded that balloon angioplasty with selective stenting had comparable 2-year patency rates with primary stenting (77% and 78%, respectively). It must be noted, however, that it was necessary to stent 43% of the patients in the PTA treatment group due to unsatisfactory angioplasty results. The 5-year outcomes between the two groups were also similar, with 82% and 80% of the treated iliac segments remaining free of the need for new revascularization procedures after a mean follow-up of 5.6 ± 1.3 years.138

### Lower Extremity Arterial Occlusive Disease

The symptoms of lower extremity occlusive disease are classified into two large categories: acute limb ischemia (ALI) and chronic limb ischemia (CLI). Ninety percent of acute ischemia cases are either thrombotic or embolic. Frequently, sudden onset of limb-threatening ischemia may be the result of acute exacerbation of the preexisting atherosclerotic disease. Chronic ischemia is largely due to atherosclerotic changes of the lower extremity that manifest from asymptomatic to limb-threatening gangrene. As the population ages, the prevalence of chronic occlusive disease of the lower extremity is increasing, and it significantly influences lifestyle, morbidity, and mortality. In addition, multiple comorbid conditions increase risks of surgical procedures. Endovascular interventions become an important alternative in treating lower extremity occlusive disease. However, despite rapidly evolving endovascular technology, lower extremity endovascular intervention continues to be one of the most controversial areas of endovascular therapy.

### Epidemiology

In a detailed review of the literature, McDaniel and Cronenwett concluded that claudication occurred in 1.8% of patients under 60 years of age, 3.7% of patients between 60 and 70 years of age, and 5.2% of patients over 70 years of age.139 Leng and his colleagues scanned 784 subjects using ultrasound in a random sample of men and women age 56 to 77 years. Of the subjects who were scanned, 64% demonstrated atherosclerotic plaque.140 However, a large number of patients had occlusive disease without significant symptoms. In a study by Schroll and Munck, only 19% of patients with peripheral vascular disease were symptomatic.141 Using ABIs, Stoffers and colleagues scanned 3171 individuals between the ages of 45 and 75 and identified that 6.9% of patients had ABIs <0.95, only 22% of whom had symptoms.142 In addition, they demonstrated that concomitant cardiovascular and cerebrovascular diseases were three to four times higher among the group with asymptomatic peripheral vascular diseases than those without peripheral vascular disease. Furthermore, they confirmed that 68% of all peripheral arterial obstructive diseases were unknown to the primary care physician, and this group mainly represented less advanced cases of atherosclerosis. However, among patients with an ABI ratio <0.75, 42% were unknown to the primary physicians.

### Diagnostic Evaluation

The diagnosis of lower extremity occlusive disease is often made based on a focused history and physical examination and confirmed by the imaging studies. A well-performed physical examination often reveals the site of lesions by detecting changes in pulses, temperature, and appearances. The bedside ABIs using blood pressure cuff also aid in diagnosis. Various clinical signs and symptoms are useful to differentiate conditions of viable, threatened, and irreversible limb ischemia caused by arterial insufficiency (Table 23-16).

Noninvasive studies are important in documenting the severity of occlusive disease objectively. Ultrasound Dopplers measuring ABIs and segmental pressures are widely used in North America and Europe. Normal ABI is greater than 1.0. In patients with claudication, ABIs decrease to 0.5 to 0.9 and to even lower levels in patients with rest pain or tissue loss. Segmental pressures are helpful in identifying the level of involvement. Decrease in segmental pressure between two segments indicates significant disease. Ultrasound duplex scans are used
to identify the site of lesion by revealing flow disturbance and velocity changes. A meta-analysis of 71 studies by Koelmay and associates confirmed that duplex scanning is accurate for assessing arterial occlusive disease in patients suffering from claudication or critical ischemia with an accumulative sensitivity of 80% and specificity of over 95%.\textsuperscript{143} Adding an ultrasound contrast agent further increases the sensitivity and specificity of ultrasound technology. Other noninvasive imaging technologies, such as MRA and CTA, are rapidly evolving and gaining popularity in the diagnosis of lower extremity occlusive disease (Figs. 23-59 and 23-60).

Contrast angiography remains the gold standard imaging study. Using contrast angiography, interventionists can locate and size the anatomic significant lesions and measure the pressure gradient across the lesion, as well as plan for potential intervention. Angiography is, however, semi-invasive and should be confined to patients for whom surgical or percutaneous intervention is contemplated. Patients with borderline renal function may need to have alternate contrast agents, such as gadolinium or carbon dioxide, to avoid contrast-induced nephrotoxicity.

**Differential Diagnosis**

Arterial insufficiency frequently leads to muscle ischemic pain involving the lower extremity muscles, particularly during exercise. Intermittent claudication is pain affecting the calf and, less commonly, the thigh and buttock that is induced by exercise and relieved by rest. Symptom severity varies from mild to severe. Intermittent claudication occurs as a result of muscle ischemia during exercise caused by obstruction to arterial flow. Regarding the differential diagnosis of intermittent claudication, there are a variety of neurologic, musculoskeletal, and venous conditions that may produce symptoms of calf pain (Table 23-17). Additionally, various nonatherosclerotic conditions can also cause symptoms consistent with intermittent lower extremity claudication (Table 23-18). Nocturnal calf muscle spasms or night cramps are not indicative of arterial disease. They are common but are difficult to diagnose with certainty. Foot ulceration is not always the result of arterial insufficiency. Ischemic ulcers occur on the toes or lateral side of the foot and are painful. By comparison, venous ulcers, which are also common, occur above the medial malleolus, usually in an area with the skin changes of lipodermatosclerosis, and cause mild discomfort. Neuropathic ulcers are usually found on weight-bearing surfaces, have thick calluses, and are pain free. Ulcers may be the result of more than one etiology. Rest pain must be distinguished from peripheral neuropathy, which is prevalent in diabetic patients. Patients with diabetic neuropathy tend to have decreased vibration and position sense and decreased reflexes. Spinal stenosis causes pain that is exacerbated with standing and back extension.

**Lower Extremity Occlusive Disease Classification**

Lower extremity occlusive disease may range from exhibiting no symptoms to limb-threatening gangrene. There are two major classifications developed based on the clinical presentations.

**Table 23-16**

<table>
<thead>
<tr>
<th>DESCRIPTION</th>
<th>VIABLE</th>
<th>THREATENED</th>
<th>IRREVERSIBLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical description</td>
<td>Not immediately threatened</td>
<td>Salvageable if promptly treated</td>
<td>Major tissue loss, amputation unavoidable</td>
</tr>
<tr>
<td>Capillary return</td>
<td>Intact</td>
<td>Intact, slow</td>
<td>Absent (marbling)</td>
</tr>
<tr>
<td>Muscle weakness</td>
<td>None</td>
<td>Mild, partial</td>
<td>Profound, paralysis (rigor)</td>
</tr>
<tr>
<td>Sensory loss</td>
<td>None</td>
<td>Mild, incomplete</td>
<td>Profound anesthetic</td>
</tr>
<tr>
<td>Arteriovenous Doppler finding</td>
<td>Audible</td>
<td>Inaudible or audible</td>
<td>Inaudible</td>
</tr>
</tbody>
</table>

![Figure 23-59](image_url) High-resolution computed tomography angiography of a patient with normal right lower extremity arterial circulation. Distal occlusive disease is noted in the left tibial arteries (arrow).
The Fontaine classification uses four stages: Fontaine I is the stage when patients are asymptomatic; Fontaine II is when they have mild (IIa) or severe (IIb) claudication; Fontaine III is when they have ischemic rest pain; and Fontaine IV is when patients suffer tissue loss, such as ulceration or gangrene (Table 23-19).

The Rutherford classification has four grades (0–III) and seven categories (0–6). Asymptomatic patients are classified into category 0; claudicants are stratified into grade I and divided into three categories based on the severity of the symptoms; patients with rest pain belong to grade II and category 4; and patients with tissue loss are classified into grade III and categories 5 and 6 based on the significance of the tissue loss.

These clinical classifications help to establish uniform standards in evaluating and reporting the results of diagnostic measurements and therapeutic interventions (Table 23-19).

The most clinically useful classification of lower extremity atherosclerotic disease should be based on the morphologic character of the lesions. The TASC taskforce published a guideline separating lower extremity arterial diseases into femoropopliteal and infrapopliteal lesions (Table 23-20). This guideline is particularly useful in determining intervention strategies based on the disease classifications. Based on the guideline, femoropopliteal lesions are divided into four types: A, B, C, and D. Type A lesions are single focal lesions less than 3 cm in length and do not involve the origins of the SFA or the distal popliteal artery. Type B lesions are single lesions 3 to 5 cm in length not involving the distal popliteal artery or multiple or heavily calcified lesions less than 3 cm in length. Type C lesions are multiple stenoses or occlusions greater than 15 cm in length or recurrent stenoses or occlusions that need treatment after two endovascular interventions. Type D lesions are those with complete occlusion of CFA, SFA, or popliteal artery.

In a similar fashion, infrapopliteal arterial diseases are classified into four types based on TASC guideline (Fig. 23-61). Type A lesions are single lesions less than 1 cm in length not involving the trifurcation. Type B lesions are multiple lesions less than 1 cm in length or single lesions shorter than 1 cm involving the trifurcation. Type C lesions are lesions that extensively involve trifurcation or 1- to 4-cm stenotic or 1- to 2-cm occlusive lesions. Type D lesions are occlusions longer than 2 cm or diffuse lesions.

**Etiology of Acute Limb Ischemia**

ALI is defined as sudden loss of limb perfusion, and the term is applicable up to 2 weeks after an initiating event. While the instances of acute leg ischemia caused by emboli have decreased due to more effective treatment of rheumatic fever and atrial fibrillation, the incidence of thrombotic acute leg ischemia has increased. Even with the extensive use of newer endovascular techniques including thrombolysis, most published series report a 10% to 30% 30-day amputation rate. The short-term mortality of patients presenting with acute ischemia is 15% to 20%. The most common etiologies of ALI include embolism, native vessel thrombosis, reconstruction thrombosis, trauma, and complications of peripheral aneurysm. Most cases of lower extremity ALI are the result of thrombosis of a prosthetic conduit. This stems from increased use of prosthetic conduits to address CLI.

Presenting symptoms in ALI are pain and loss of sensory or motor function. The abruptness and time of onset of the pain, its location and intensity, and change in severity over time should all be taken into consideration. The duration and intensity of the pain and presence of motor or sensory changes are very important in clinical decision making and urgency of revascularization. Thrombolysis may be less effective for thrombosis of ≥2 weeks in duration compared with acute thrombosis.

**Arterial Embolism.** The heart is the most common source of distal emboli, which accounts for more than 90% of peripheral arterial embolic events. Atrial fibrillation is the most common source. Sudden cardiovension results in the dilated noncontractile atrial appendage regaining contractile activity, which can dislodge the contained thrombus. Other cardiac sources include mural thrombus overlying a myocardial infarction or thrombus forming within a dilated left ventricular aneurysm. Mural thrombi can also develop within a dilated ventricle dilated by
### Differential diagnosis of intermittent claudication

<table>
<thead>
<tr>
<th>CONDITION</th>
<th>LOCATION OF PAIN OR DISCOMFORT</th>
<th>CHARACTERISTIC DISCOMFORT</th>
<th>ONSET RELATIVE TO EXERCISE</th>
<th>EFFECT OF REST</th>
<th>EFFECT OF BODY POSITION</th>
<th>OTHER CHARACTERISTICS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermittent claudication (calf)</td>
<td>Calf muscles</td>
<td>Cramping pain</td>
<td>After same degree of exercise</td>
<td>Quickly relieved</td>
<td>None</td>
<td>Reproducible</td>
</tr>
<tr>
<td>Intermittent claudication (e.g., herniated disk)</td>
<td>Radiates down leg, usually posteriorly</td>
<td>Sharp lancinating pain</td>
<td>Soon, if not immediately after onset</td>
<td>Not quickly relieved (also often present at rest)</td>
<td>Relief may be aided by adjusting back position</td>
<td>History of back problems, provocative increase in intra-abdominal pressure</td>
</tr>
<tr>
<td>Chronic compartment syndrome</td>
<td>Entire leg, but usually worse in thigh and groin</td>
<td>Tight, bursting pain</td>
<td>After walking</td>
<td>Subsides slowly</td>
<td>Relief speeded by elevation</td>
<td>History of iliofemoral venous thrombosis, venous congestion</td>
</tr>
<tr>
<td>Venous claudication</td>
<td>Entire leg, but usually worse in thigh and groin</td>
<td>Tight, bursting pain</td>
<td>After walking</td>
<td>Subsides slowly</td>
<td>Relief speeded by elevation</td>
<td>History of iliofemoral venous thrombosis, venous congestion</td>
</tr>
<tr>
<td>Nerve root compression (e.g., herniated disk)</td>
<td>Radiates down leg, usually posteriorly</td>
<td>Sharp lancinating pain</td>
<td>Soon, if not immediately after onset</td>
<td>Not quickly relieved (also often present at rest)</td>
<td>Relief may be aided by adjusting back position</td>
<td>History of back problems, provocative increase in intra-abdominal pressure</td>
</tr>
<tr>
<td>Baker’s cyst</td>
<td>Behind knee, down calf</td>
<td>Swelling, soreness, tenderness</td>
<td>With exercise</td>
<td>Present at rest</td>
<td>None</td>
<td>Not intermittent</td>
</tr>
<tr>
<td>Hip, thigh, buttocks</td>
<td>Aching discomfort, weakness</td>
<td>After same degree of exercise</td>
<td>Quickly relieved</td>
<td>None</td>
<td>Reproducible</td>
<td></td>
</tr>
<tr>
<td>Hip, thigh, buttocks</td>
<td>Aching discomfort</td>
<td>After variable degree of exercise</td>
<td>Not quickly relieved (and may be present at rest)</td>
<td>More comfortable sitting, weight taken off legs</td>
<td>Variable, may relate to activity level, weather changes</td>
<td></td>
</tr>
<tr>
<td>Hip, thigh, buttocks (follows dermatome)</td>
<td>Weakness more than pain</td>
<td>After walking or standing for same length of time</td>
<td>Relieved by stopping only if position changed</td>
<td>Relief by lumbar spine flexion (sitting or stooping forward) pressure</td>
<td>Frequent history of problems, provocation by increased intra-abdominal pressure</td>
<td></td>
</tr>
<tr>
<td>Foot, arch</td>
<td>Severe deep pain and numbness</td>
<td>After same degree of exercise</td>
<td>Quickly relieved</td>
<td>None</td>
<td>Reproducible</td>
<td></td>
</tr>
<tr>
<td>Foot, arch</td>
<td>Aching pain</td>
<td>After variable degree of exercise</td>
<td>Not quickly relieved (and may be present at rest)</td>
<td>May be relieved by not bearing weight</td>
<td>Variable, may relate to activity level</td>
<td></td>
</tr>
</tbody>
</table>
cardiomyopathy. Emboli that arise from a ventricular aneurysm or from a dilated cardiomyopathy can be very large and can lodge at the aortic bifurcation (saddle embolus), thus rendering both legs ischemic. Diseased valves are another source of distal embolization. Historically, this occurred as a result of rheumatic heart disease. Currently, subacute endocarditis and acute bacterial endocarditis are the more common causes. Infected emboli can seed the recipient vessel wall, creating mycotic aneurysms.

An electrocardiogram (ECG) will diagnose atrial fibrillation. A transthoracic or transesophageal echocardiogram should be performed looking for a cardiac source. It is important to seek other sources of the embolus using CT scanning of the descending thoracic and abdominal aorta. More unusual sources include mural thrombus from an aortic aneurysm, and occasionally, idiopathic arterial-to-arterial thrombus occurs, usually from thrombus that has formed in an atherosclerotic aortic arch or descending thoracic aorta. The presence of mobile plaque on transesophageal echocardiography is suggestive of this source.

Paradoxical embolus occurs when a patient has a patent foramen ovale and an embolus from a deep venous thrombosis crosses through the atrial defect into the left side of the heart and passes into the peripheral circulation. This is diagnosed using a bubble echocardiography, in which air bubbles introduced into the venous circulation can be seen traversing the septal defect.

**Arterial Thrombosis.** Thrombosis can occur in native arteries and in arterial reconstructions. Patients with thrombosed arterial segments often have an underlying atherosclerotic lesion at the site of thrombosis or aneurysmal degeneration with mural thrombosis. It is important to obtain a history, determine risk factors for atherosclerosis and hypercoagulable status, and examine the contralateral extremity for circulatory problems. Patients with thrombosis or prior arterial reconstructions have limb incisions from previous surgery, and graft occlusion can be confirmed with duplex imaging.

### Clinical Manifestations of Acute Limb Ischemia

Acute lower extremity ischemia manifests with the “five Ps”: pain, pallor, paresthesias, paralysis, and pulselessness, to which some add a sixth “P”—poikilothermia or “perishing cold.” Pain is the usual symptom that causes a patient to present to the emergency room. The most common location for an embolus to lodge in the leg is at the common femoral bifurcation. Typically, a patient will complain of foot and calf pain. Pulses are absent, and there may be diminution of sensation. Inability to move the affected muscle group is a sign of very severe ischemia and necessitates urgent revascularization. During evaluation of the affected extremity, it is important to compare findings with the contralateral limb. Clinical evaluation is extremely important in determining the etiology and location of the obstruction. One of the most important pieces of information to obtain is whether the patient has had prior vascular procedures or if there is a history of lower extremity claudication. Either of these features suggests preexisting vascular disease, renders revascularization more complicated, and usually mandates angiography to permit surgical planning. On the contrary, in a patient with no history suggestive of prior vascular disease, the etiology is most likely embolic, and simple thrombectomy is more likely to be successful.

Absent bilateral femoral pulses in a patient with bilateral lower extremity ischemia is most likely due to saddle embolus to the aortic bifurcation. A palpable femoral pulse and absent popliteal and distal pulses may either be due to distal common femoral embolus (the pulse being palpable above the level of occlusion) or embolus to the superficial femoral or popliteal arteries. Typically, emboli lodge at arterial bifurcations where they are trapped due to sudden reductions in arterial diameter. A popliteal trifurcation embolus will present with calf ischemia and absent pedal pulses, possibly with a popliteal pulse present. The finding of palpable contralateral pulses and the absence of ipsilateral pulses in the acutely ischemic leg are suggestive of an embolus, irrespective of presence of Doppler signals.

<table>
<thead>
<tr>
<th>Table 23-18</th>
<th>Nonatherosclerotic causes of intermittent claudication</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Aortic coarctation</td>
<td>• Arterial thrombosis</td>
</tr>
<tr>
<td>• Arterial fibrodysplasia</td>
<td>• Primary vascular tumors</td>
</tr>
<tr>
<td>• Iliac syndrome of the cyclist</td>
<td>• Pseudoaxanthoma elasticum</td>
</tr>
<tr>
<td>• Peripheral emboli</td>
<td>• Remote trauma or radiation injury</td>
</tr>
<tr>
<td>• Persistent sciatic artery</td>
<td>• Takayasu’s disease</td>
</tr>
<tr>
<td>• Popliteal aneurysm</td>
<td>• Thromboangiitis obliterans</td>
</tr>
</tbody>
</table>

### Table 23-19

<table>
<thead>
<tr>
<th>Classification of peripheral arterial disease based on the Fontaine and Rutherford classifications</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FONTAINE CLASSIFICATION</strong></td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td>I</td>
</tr>
<tr>
<td>Ia</td>
</tr>
<tr>
<td>Iib</td>
</tr>
<tr>
<td>III</td>
</tr>
<tr>
<td>IV</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Arteriography is not mandatory in patients without antecedent history suggestive of vascular disease; nevertheless, all patients should be positioned on the operating room table in such a way that fluoroscopic access to the entire inflow and outflow tract is possible if necessary.

The main question to be answered by the history and physical examination is the severity of the ALI, which is the major consideration in early management decisions. Patients with ALI should be evaluated in a fashion that considers the severity and duration of ischemia at the time of presentation. Ideally, all patients with acute ischemia should be investigated with imaging, especially if there is an antecedent vascular reconstruction; however, the clinical condition and access to resources must guide further investigations. Unnecessary delays can result in amputation. Arteriography, if it can be performed in a timely fashion, is an excellent modality for localizing obstructions and deciding which type of intervention (endovascular, embolectomy, or bypass) patients will benefit more from. One of the goals of treatment for ALI is to prevent thrombus propagation; therefore, expedient anticoagulation with heparin is indicated as soon as the diagnosis is suspected.

### Treatment Considerations for Acute Limb Ischemia

In the absence of any significant contraindication, the patient with an ischemic lower extremity should be immediately anticoagulated. This will prevent propagation of the clot into unaffected vascular beds. Intravenous fluid should be started and a Foley catheter inserted to monitor urine output. Baseline labs should be obtained and creatinine levels noted. A hypercoagulable workup should be performed prior to initiation of heparin if there is sufficient suspicion. According to results from randomized trials, there is no clear superiority for thrombolysis over surgery in terms of 30-day limb salvage or mortality. Access to each treatment option is a major issue in the decision-making process, as time is often critical. National registry data from the United States reveal that surgery is used three- to five-fold more frequently than thrombolysis. Three randomized studies have investigated the role of catheter-directed thrombolytic therapy in the treatment of ALI.145

### Endovascular Treatment

The potential to reduce mortality and morbidity while achieving limb salvage is the impetus that makes thrombolysis preferable to open surgery as first-line treatment in patients with ALI (classes I and IIa). Advantages of thrombolytic therapy over balloon embolectomy include the reduced endothelial trauma and potential for more gradual and complete clot lysis in branch vessels usually too small to access by embolectomy balloons. It is hoped that the more gradual clot dissolution with thrombolysis may decrease the incidence of reperfusion injury that is encountered after open surgical procedures where rapid return of blood flow may precipitate compartment syndrome. Skeletal muscle tissue appears to be most vulnerable to ischemia. Pathophysiologic studies reveal that irreversible damage to muscle tissue starts after 3 hours of ischemia and is nearly complete at 6 hours. Progressive microvascular damage appears to follow rather than precede skeletal muscle tissue damage. The more severe the cellular damage, the greater are the microvascular changes. When the musculature and microvasculature are severely damaged, amputation rather than attempts at revascularization may be the most prudent course to prevent wash-out of toxic by-product from the ischemic limb into the systemic circulation. The mortality rate associated with reperfusion syndrome is high because of the development of concomitant adult respiratory distress syndrome, shock, disseminated intravascular coagulation, and renal failure.

Patients with small-vessel occlusion are poor candidates for surgery because they lack distal target vessels to use for bypass. These patients should be offered a trial of thrombolysis, unless they have contraindications to thrombolysis or their

---

**Table 23-20** TransAtlantic Inter-Society Consensus classification of femoral popliteal occlusive lesions

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Single stenosis ≤10 cm in length&lt;br&gt;Single occlusion ≤5 cm in length</td>
</tr>
<tr>
<td>B</td>
<td>Multiple lesions (stenoses or occlusions), each ≤5 cm&lt;br&gt;Single stenosis or occlusion ≤15 cm not involving the infrageniculate popliteal artery&lt;br&gt;Single or multiple lesions in the absence of continuous tibial vessels to improve inflow for a distal bypass&lt;br&gt;Heavily calcified occlusion ≤5 cm in length&lt;br&gt;Single popliteal stenosis</td>
</tr>
<tr>
<td>C</td>
<td>Multiple stenoses or occlusions totaling &gt;15 cm with or without heavy calcification&lt;br&gt;Recurrent stenoses or occlusions that need treatment after two endovascular interventions</td>
</tr>
<tr>
<td>D</td>
<td>Chronic total occlusions of CFA or SFA (&gt;20 cm, involving the popliteal artery)&lt;br&gt;Chronic total occlusion of popliteal artery and proximal trifurcation vessels</td>
</tr>
</tbody>
</table>

CFA = common femoral artery; SFA = superficial femoral artery.

---

**Figure 23-61.** Schematic depiction of TransAtlantic Inter-Society Consensus classification of femoral popliteal occlusive lesions.
ischemia is so severe that the time needed to achieve adequate lysis is considered too long. The major contraindications of thrombolysis are recent stroke, intracranial primary malignancy, brain metastases, or intracranial surgical intervention. Relative contraindications for performance of thrombolysis include renal insufficiency, allergy to contrast material, cardiac thrombus, diabetic retinopathy, coagulopathy, and recent arterial puncture or surgery (Table 23-21).

Advances in clot removal techniques with percutaneous mechanical thrombectomy and thromboaspiration may extend the applicability of this intervention to patients with more advanced degrees of ALI (class IIb) and contraindications to thrombolysis. Several thrombectomy devices have received FDA approval for acute lower extremity arterial thrombosis. The utility of these thrombectomy devices is that they can be used as a standalone therapy when there are contraindications for thrombolytic therapy. Additionally, these thrombectomy devices can be used in conjunction with thrombolytic agents, for pharmacomechanical thrombectomy, to enhance clot lysis and to limit the doses and time required for thrombolysis.145

**Table 23-21**

<table>
<thead>
<tr>
<th>Absolute contraindications</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Established cerebrovascular events (including transient ischemic attack) within last 2 months</td>
<td></td>
</tr>
<tr>
<td>Active bleeding diathesis</td>
<td></td>
</tr>
<tr>
<td>Recent (&lt;10 days) gastrointestinal bleeding</td>
<td></td>
</tr>
<tr>
<td>Neurosurgery (intracranial or spinal) within last 3 months</td>
<td></td>
</tr>
<tr>
<td>Intracranial trauma within last 3 months</td>
<td></td>
</tr>
<tr>
<td>Intracranial malignancy or metastasis</td>
<td></td>
</tr>
<tr>
<td><strong>Relative major contraindications</strong></td>
<td></td>
</tr>
<tr>
<td>Cardiopulmonary resuscitation within last 10 days</td>
<td></td>
</tr>
<tr>
<td>Major nonvascular surgery or trauma within last 10 days</td>
<td></td>
</tr>
<tr>
<td>Uncontrolled hypertension (&gt;180 mmHg systolic or &gt;110 mmHg diastolic)</td>
<td></td>
</tr>
<tr>
<td>Puncture of noncompressible vessel</td>
<td></td>
</tr>
<tr>
<td>Intracranial tumor</td>
<td></td>
</tr>
<tr>
<td>Recent eye surgery</td>
<td></td>
</tr>
<tr>
<td><strong>Minor contraindications</strong></td>
<td></td>
</tr>
<tr>
<td>Hepatic failure, particularly with coagulopathy</td>
<td></td>
</tr>
<tr>
<td>Bacterial endocarditis</td>
<td></td>
</tr>
<tr>
<td>Pregnancy</td>
<td></td>
</tr>
<tr>
<td>Diabetic hemorrhagic retinopathy</td>
<td></td>
</tr>
</tbody>
</table>

Surgical Treatment

**Embolectomy.** When a decision is made to proceed with open surgical intervention, the abdomen, contralateral groin, and entire lower extremity are prepped in the field. The groin is opened through a vertical incision, exposing the CFA and its bifurcation. Frequently, the location of the embolus at the femoral bifurcation is readily apparent by the presence of a palpable proximal femoral pulse, which disappears distally. The artery is clamped and opened transversely over the bifurcation. Thrombus is extracted by passing a Fogarty balloon embolectomy catheter. Good back-bleeding and antegrade bleeding suggest that the entire clot has been removed. Embolic material often forms a cast of the vessel and is sent for culture and histologic examination. Completion angiography is advisable to ascertain the adequacy of clot removal. The artery is then closed and the patient fully anticoagulated.

When an embolus lodges in the popliteal artery, in most cases it can be extracted via a femoral incision using the techniques previously described. A femoral approach is preferred because the larger diameter of the femoral artery results in decreased likelihood of arterial compromise when the arteriotomy is closed. The disadvantage with using the femoral approach for embolectomy is the greater difficulty involved in directing the embolectomy catheter into each of the infrapopliteal arteries. Use of fluoroscopic imaging and an over-the-wire thrombectomy catheter can overcome this problem. Alternatively, use of a separate incision to expose the popliteal bifurcation may be necessary to achieve a complete thrombectomy.

A more complex situation arises when a patient has antecedent peripheral vascular disease and in situ thrombosis develops on top of preexisting atheroma because, frequently, embolectomy catheters will not pass through these occlusions. Similarly, when a bypass graft fails, it is usually due to progression of atheroma proximal or distal to the graft anastomoses or to intrinsic stenoses that develop within a vein graft. In these scenarios, expeditious angiography is useful to determine the extent of the occlusion, to search for inflow and distal outflow vessels, and to decide whether thrombolysis or surgery will be the better intervention. Although the surgeon’s preference tends to dictate the approach selected, the decision is based on the presence or absence of good target vessels and availability of a suitable bypass conduit. If there are good distal vessels and the saphenous vein is suitable, surgical bypass is recommended because it is fast, durable, and reliable. In the absence of a good distal target and saphenous vein, or in a patient at high risk for surgery, lysis is recommended.

**Bypass Graft Thrombectomy.** Bypass thrombectomy is more likely to succeed with prosthetic bypasses. Bypass graft revision or replacement is more appropriate for acute vein graft failures because they are less likely to respond to thrombolysis and require some type of revision, such as valve lysis, interposition, or extension. Thrombectomy of autogenous grafts is prone to failure unless an anatomic cause for failure such as a retained valve or unligated side branch is found and corrected. The performance of a fasciotomy to circumvent reperfusion injury/compartment syndrome is an important consideration.

**Complications Related to Treatment for Acute Limb Ischemia**

Adverse events related to catheter-directed thrombolysis are primarily related to bleeding complications. The overall risk of hemorrhagic stroke from a thrombolysis procedure has been reported to be 1% to 2.3%, with 50% of hemorrhagic complications occurring during the thrombolytic procedure.146 Hematoma at the vascular puncture site has been reported in 12% to 17% of cases. Gastrointestinal bleeding is reported in 5% to 10% of cases. Hematuria following thrombolysis is uncommon and should prompt a search for urinary tumors. Hemorrhage requiring transfusion can occur in approximately 25% of patients undergoing thrombolysis. Lytic agents are absolutely contraindicated in patients with intracranial surgery, intracranial hemorrhage within the last 3 months, or any active bleeding. Most bleeding complications occur at the arterial puncture sites, but concealed retroperitoneal bleeding is possible. The most feared complication that patients can sustain is intracerebral hemorrhage. Older patients may be more susceptible
Complications of arterial revascularization

<table>
<thead>
<tr>
<th>Complication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compartment syndrome</td>
</tr>
<tr>
<td>Ischemic neuropathy</td>
</tr>
<tr>
<td>Muscle necrosis</td>
</tr>
<tr>
<td>Recurrent thrombosis</td>
</tr>
<tr>
<td>Lower leg swelling</td>
</tr>
<tr>
<td>Reperfusion syndrome</td>
</tr>
<tr>
<td>Hypotension</td>
</tr>
<tr>
<td>Hyperkalemia</td>
</tr>
<tr>
<td>Myoglobinuria</td>
</tr>
<tr>
<td>Renal failure</td>
</tr>
</tbody>
</table>

to this complication, and thus many interventionalists are extremely reticent to use thrombolysis in patients older than 80 years of age.

Patients who are treated for acute ischemia are susceptible to two major complications following revascularization: reperfusion and compartment syndromes. Other procedure-related complications include arterial rethrombosis, recurrent embolization, and arterial injuries secondary to the balloon catheter manipulations.

Reperfusion of the ischemic limb is variable in its physiologic effects and directly relates to the severity and extent of the ischemia. Patients with a saddle embolus of the aortic bifurcation and severely ischemic limbs may develop the full-blown “reperfusion syndrome,” whereas patients with minimal muscle ischemia who are reperfused in a timely fashion essentially develop no effects. Many patients with ALI have severe underlying cardiac disease and are unable to tolerate even short ischemic periods. Complications occurring after revascularization of the lower extremity and causes of recurrent thrombosis are listed in Table 23-22.

Compartment syndrome occurs after prolonged ischemia is followed by reperfusion. The capillaries leak fluid into the interstitial space in the muscles, which are enclosed within a nondistensible fascial envelope. When the pressure inside the compartment exceeds the capillary perfusion pressure, nutrient flow ceases and progressive ischemia occurs, even in the presence of peripheral pulses. Consequently, every patient who has sustained an ischemic event and is reperfused is monitored for compartment syndrome, which is characterized by excessive pain in the compartment, pain on passive stretching of the compartment, and sensory loss due to nerve compression of the nerves coursing through the compartment (Table 23-23 and Fig. 23-62). The most commonly affected compartment is the anterior compartment in the leg. Numbness in the web space between the first and second toes is diagnostic due to compression of the deep peroneal nerve. Compartment pressure is measured by inserting an arterial line into the compartment and recording the pressure. Although controversial, pressures greater than 20 mmHg are an indication for fasciotomy. Compartment pressures are relieved in the leg by medial and lateral incisions. Through the medial incision, long openings are then made in the fascia of the superficial and deep posterior compartments. Through the lateral incision, the anterior and peroneal compartments are opened. Both skin and fascial incisions should be of adequate length to ensure full compartment decompression. Laboratory evidence of rhabdomyolysis is seen in 20% of cases. The myoglobin from damaged muscle precipitates in kidney tubules and causes acute tubular necrosis. Alkalization of urine increases the solubility of myoglobin, thus preventing

Table 23-23

<table>
<thead>
<tr>
<th>ANTERIOR COMPARTMENT</th>
<th>LATERAL COMPARTMENT</th>
<th>SUPERFICIAL POSTERIOR COMPARTMENT</th>
<th>DEEP POSTERIOR COMPARTMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muscles</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tibialis anterior</td>
<td>Peroneus longus</td>
<td>Gastrocnemius</td>
<td>Tibialis posterior</td>
</tr>
<tr>
<td>Extensor digitorum</td>
<td>Peroneus brevis</td>
<td>Plantaris</td>
<td>Flexor digitorum longus</td>
</tr>
<tr>
<td>longus</td>
<td></td>
<td>Soleus</td>
<td>Flexor hallucis longus</td>
</tr>
<tr>
<td>Peroneus tertius</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extensor hallucis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>longus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extensor digitorum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>brevis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extensor hallucis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>brevis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Artery</td>
<td>Anterior tibial artery</td>
<td>Anterior and posterior tibial</td>
<td>Posterior tibial artery</td>
</tr>
<tr>
<td></td>
<td></td>
<td>branches of the popliteal artery</td>
<td>Peroneal artery</td>
</tr>
<tr>
<td>Nerve</td>
<td>Deep peroneal nerve</td>
<td>Superficial peroneal nerve</td>
<td>Tibial nerve</td>
</tr>
</tbody>
</table>
it from crystallizing in the tubules. In addition to alkalinization, therapy consists of forced saline diuresis and removal of the source of dead muscle that is releasing the myoglobin.

**Clinical Manifestations of Chronic Limb Ischemia**

The term CLI is reserved for patients with objectively proven arterial occlusive disease and symptoms lasting for more than 2 weeks. Symptoms include rest pain and tissue loss, such as ulceration or gangrene (Table 23-24). The diagnosis should be corroborated with noninvasive diagnostic tests, such as the ABI, toe pressures, and transcutaneous oxygen measurements. Ischemic rest pain most commonly occurs below an ankle pressure of 50 mmHg or a toe pressure less than 30 mmHg. Ulcers are not always of an ischemic etiology (Table 23-25). In many instances, there are other etiologic factors (traumatic, venous, or neuropathic) that are contributory, but it is underlying peripheral arterial disease that may be responsible for delayed or absent healing (Fig. 23-63). Healing of ulcers requires an inflammatory response and greater perfusion than is required to support intact skin and underlying tissues. As a result, the ankle and toe pressure levels needed for healing are higher than the pressures seen with ischemic rest pain. For patients with ulcers or gangrene, the presence of CLI is suggested by an ankle pressure less than 70 mmHg or a toe systolic pressure less than 50 mmHg. It is important to understand that there is no definite consensus regarding the vascular hemodynamic parameters required to make the diagnosis of CLI.

One of the most common sites for occlusive disease is in the distal SFA as it passes deep through the adductor canal. It may be that the entrapment by the adductor hiatus prevents the compensatory dilation that occurs in atherosclerotic vessels. Stenoses, which develop here, progress to occlusion of the distal third of the SFA (Fig. 23-64). When distal SFA occlusion develops slowly, it may be totally asymptomatic because of development of collaterals from the proximal SFA, or the PFA can bypass the occlusion and reconstitute the popliteal artery. Symptom development is a function of the extent of occlusion, adequacy of collaterals, and the activity level of the patients. Presenting symptoms of femoropopliteal occlusive disease are broadly classified into two types: limb-threatening and non–limb-threatening ischemia. Claudication is non–limb-threatening, while rest pain, ulceration, and gangrene are limb-threatening and warrant urgent intervention. Occlusive disease of the femoral artery may be isolated or occur in conjunction with multilevel disease that involves both the aortoiliac segment and the tibial vessels. Symptoms in patients with

---

**Table 23-24**

<table>
<thead>
<tr>
<th>GRADE</th>
<th>CATEGORY</th>
<th>CLINICAL DESCRIPTION</th>
<th>OBJECTIVE CRITERIA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>Asymptomatic—no hemodynamically significant occlusive disease</td>
<td>Normal treadmill or reactive hyperemia test</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Mild claudication</td>
<td>Able to complete treadmill exercise; AP after exercise &gt;50 mmHg but at least 20 mmHg lower than resting value</td>
</tr>
<tr>
<td>I</td>
<td>2</td>
<td>Moderate claudication</td>
<td>Between categories 1 and 3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Severe claudication</td>
<td>Cannot complete standard treadmill exercise and AP after exercise &lt;50 mmHg</td>
</tr>
<tr>
<td>II</td>
<td>4</td>
<td>Ischemic rest pain</td>
<td>Resting AP &lt;40 mmHg, flat or barely pulsatile ankle or metatarsal PVR; TP &lt;30 mmHg</td>
</tr>
<tr>
<td>III</td>
<td>5</td>
<td>Minor tissue loss—nonhealing ulcer, focal gangrene with diffuse pedal ischemia</td>
<td>Resting AP &lt;60 mmHg, ankle or metatarsal PVR flat or barely pulsatile; TP &lt;40 mmHg</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Major tissue loss—extending above TM level, functional foot no longer salvageable</td>
<td>Same as category 5</td>
</tr>
</tbody>
</table>

1Five minutes at 2 miles per hour on a 12% incline of treadmill exercise.
2Grades II and III, categories 4, 5, and 6, are encompassed by the term chronic critical ischemia.
AP = ankle pressure; PVR = pulse volume recording; TM = transmetatarsal; TP = toe pressure.

---

**Table 23-25**

<table>
<thead>
<tr>
<th>NEUROPATHIC ULCER</th>
<th>ISCHEMIC ULCER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Painless</td>
<td>Painful</td>
</tr>
<tr>
<td>Normal pulses</td>
<td>Absent pulses</td>
</tr>
<tr>
<td>Regular margins, typically punched-out appearance</td>
<td>Irregular margin</td>
</tr>
<tr>
<td>Often located on plantar surface of foot</td>
<td>Commonly located on toes, glabrous margins</td>
</tr>
<tr>
<td>Presence of calluses</td>
<td>Calluses absent or infrequent</td>
</tr>
<tr>
<td>Loss of sensation, reflexes, and vibration</td>
<td>Variable sensory findings</td>
</tr>
<tr>
<td>Increased in blood flow (arteriovenous shunting)</td>
<td>Decreased in blood flow</td>
</tr>
<tr>
<td>Dilated veins</td>
<td>Collapsed veins</td>
</tr>
<tr>
<td>Dry, warm foot</td>
<td>Cold foot</td>
</tr>
<tr>
<td>Bony deformities</td>
<td>No bony deformities</td>
</tr>
<tr>
<td>Red or hyperemic in appearance</td>
<td>Pale and cyanotic in appearance</td>
</tr>
</tbody>
</table>
multilevel disease are more severe than in those with single-level disease. Pain from isolated SFA and popliteal occlusion typically manifests as calf claudication. Cramping pain develops in the calf on ambulation, occurs at a reproducible distance, and is relieved by rest. Activities such as climbing stairs or going uphill also exacerbate the pain. Many patients report worsening symptoms during cold weather. It is important to evaluate whether the symptoms are progressive or static. In greater than 70% of patients, the disease is stable, particularly with risk factor modification.

Progression of the underlying atherosclerotic process is more likely to occur in patients with diabetes, those who continue to smoke, and those who fail to modify their atherosclerotic risk factors. In comparison, rest pain is constant, and usually occurs in the forefoot across the metatarsophalangeal joint. It is worse at night and requires placing the foot in a dependent position to improve symptoms. Patients may report that they either sleep in a chair or hang the foot off the side of the bed. The pain is severe and relentless, even with narcotics. Ischemic ulceration most commonly involves the toes. Any toe can be affected. Occasionally ulcers develop on the dorsum of the foot. Ulceration can occur in atypical positions in an ischemic foot from trauma such as friction from poorly fitting shoes. Injury to a foot with borderline ischemia can convert an otherwise stable situation into one that is limb-threatening. The initial development of gangrene commonly involves the digits. As with all vascular patients, it is important to evaluate their risk factors, intercurrent cardiac diseases, and any prior vascular interventions.

**Treatment Considerations for Chronic Limb Ischemia**

Patients with vascular diseases frequently have complicated medical comorbidities. Careful patient evaluation and selection should be performed for any peripheral arterial vascular procedure. The fundamental principle is to assess not only the surgical risk from the peripheral arterial system but also the global nature of the atherosclerotic process. Full cardiac evaluations are often necessary due to the high incidence of concomitant atherosclerotic coronary artery disease, resulting in a high risk for ischemic events. Hertzer and associates reviewed coronary angiographies on 1000 patients undergoing elective vascular procedure and identified 25% of concomitant correctable coronary artery disease, including 21% in patients undergoing elective peripheral vascular intervention. Conte and associates analyzed their 20-year experience in 1642 open lower extremity reconstructive surgeries and concluded that patients requiring lower extremity reconstruction presented an increasingly complex medical and surgical challenge compared with the previous decade in a tertiary practice setting. With aging of the population, a growing number of vascular patients have prohibitive medical comorbidities and are deemed high-risk for open surgical repair. Endovascular intervention provides an attractive alternative.

As for open surgical repair, the clinical indications for endovascular intervention of lower extremity peripheral arterial diseases include lifestyle-limiting claudication, ischemic rest pain, and tissue loss or gangrene. Importantly, endovascular procedures should be performed by a competent vascular interventionist who understands the vascular disease process and is familiar with a variety of endovascular techniques. In addition, certain lesions may not be amenable

---

**Figure 23-63.** A. A neuropathic ulcer is characterized by a punched-out appearance with loss of sensation in the surrounding skin. The foot may be warm to touch, and pulses may be present in the distal pedal arteries. B. An ischemic ulcer is characterized by a gangrenous skin change in the foot or toes. The foot is usually cold to touch with absent pedal pulses. The foot is painful to touch with decreased distal capillary refills.

---

**Figure 23-64.** Computed tomography angiogram of a patient with an occluded left superficial femoral artery (*single long arrow*) with reconstituted superficial femoral artery at the level of mid-thigh. Diffuse arterial calcifications (*double small arrows*) are noted in the mid and distal left superficial femoral arteries.
to endovascular treatment or may be associated with poor outcomes, such as long segment occlusion, heavily calcified lesion, orifice lesion, or lesions that cannot be traversed by a guidewire. Proper selection of patients and techniques is critical in achieving good long-term outcome.

Endovascular intervention for lower extremity occlusive disease is continuously evolving. Success and patency rates of endovascular intervention are closely related to the anatomic and morphologic characteristics of the treated lesions. The TASC work group made recommendations on the intervention strategies of lower extremity arterial diseases based on the morphologic characteristics. Based on TASC guidelines, endovascular treatment is recommended for type A lesions, open surgery is recommended for type D lesions, and no recommendations were made for types B and C lesions. However, with rapid advancement in endovascular technologies, there are increased numbers of lesions amendable to endovascular interventions.

There is less literature support for infrapopliteal endovascular intervention due to higher complication and lower success rates. The treatment is restricted for patients with limb-threatening ischemia who lack surgical alternatives. However, with further advancement of endovascular technology and the development of new devices, endovascular intervention is becoming an integral part of treatment (Table 23-26). By itself or combined with open technique, percutaneous intervention plays an important role in therapeutic options for lower extremity occlusive disease. As described by TASC guidelines, four criteria should be measured to evaluate the clinical success of the treatment: improvement in walking distance, symptomatic improvement, quality of life, and overall graft patency. These criteria should all be carefully weighed and evaluated for each individual prior to endovascular therapy.

### Endovascular Treatment

**Technical Considerations.** A sterile field is required in either an operating room or an angiography suite with image capability. The most common and safest access site is CFA via either a retrograde or an antegrade approach. For diagnostic angiography, arterial access should be contralateral to the symptomatic sides. For therapeutic procedures, location of the lesion and the anatomic structures of the arterial tree determine the puncture site. To avoid puncturing the iliac artery or SFA, the femoral head is located under the fluoroscopy and used as the guide for the level of needle entry. In addition, there are several useful techniques to help access a pulseless CFA including ultrasound-guided puncture, using a micropuncture kit, and targeting calcification in a calcified vessel. The antegrade approach may be challenging, particularly in obese patients. Meticulous technique is crucial in preventing complications, and a bony landmark can be used as guidance to ensure CFA puncture.

Traversing the lesion with a wire is the most critical part of the procedure. Typically, 0.035-inch guidewires are used for femoropopliteal lesions, and 0.014- or 0.018-inch guidewires are used for infrapopliteal access. Hydrophilic-coated wires, such as Glidewires, are useful in navigating through tight stenosis or occlusion. An angled-tip wire with a torque device may be helpful in crossing an eccentric lesion, and a shaped selective catheter is frequently used to help manipulate the wire across the lesion. The soft and floppy end of the wire is carefully advanced crossing the lesion under fluoroscopy, and gentle force is applied while manipulating the wire. Once the lesion is traversed, one needs to pay particular attention on the tip of the wire to ensure a secure wire access and avoid vessel wall perforation or dissection.

<table>
<thead>
<tr>
<th>Table 23-26</th>
<th>Summary of endovascular treatment strategies using device-based infrapopliteal intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTERVENTION</td>
<td>ADVANTAGES</td>
</tr>
<tr>
<td>Angioplasty</td>
<td>• Easy to use</td>
</tr>
<tr>
<td></td>
<td>• Broad range of applications</td>
</tr>
<tr>
<td>Balloon-expandable stent</td>
<td>• Overcomes arterial recoil from angioplasty</td>
</tr>
<tr>
<td></td>
<td>• Useful in treatment of flow-limiting dissection</td>
</tr>
<tr>
<td>Self-expanding stent</td>
<td>• Vessel conformability and wall apposition prevent kinking and crushing of stent</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Bioabsorbable stent</td>
<td>• Overcomes arterial recoil from angioplasty</td>
</tr>
<tr>
<td></td>
<td>• Absorbed long term to prevent risk of stent thrombosis</td>
</tr>
<tr>
<td>Cryoplasty</td>
<td>• Reduces the risk of flow-limiting dissection, therefore reducing the need for stent implantation</td>
</tr>
<tr>
<td>Cutting balloon</td>
<td>• Useful in anastomotic segments of bypass grafts and in-stent restenosis where “watermelon seeding” can prevent adequate expansion of plaque</td>
</tr>
<tr>
<td>Mechanical atherectomy</td>
<td>• Allows for debulking of plaque without the need for stent implantation in most cases</td>
</tr>
<tr>
<td></td>
<td>• Allows for removal of plaque for histologic analysis</td>
</tr>
<tr>
<td>Laser</td>
<td>• Useful in acute thrombotic and chronic total occlusions</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Once the access to the diseased vessel is secured and the wire has successfully traversed the lesion, several treatment modalities can be used either alone or in conjunction with others, including angioplasty, stent or stent graft placement, and atherectomy. The available angioplasty techniques are balloon angioplasty, cryoplasty, subintimal angioplasty, and cutting balloon; the most commonly used atherectomy techniques include percutaneous atherectomy catheter and laser atherectomy device.

Systemic anticoagulation should be maintained routinely during lower extremity arterial interventions to minimize the risk of pericatheter thrombosis. Unfractionated heparin is the most commonly used agent, administered using a weight-based formula. We typically use 80 to 100 mg/kg initial bolus for therapeutic procedure to achieve an activated clotting time above 250 seconds upon catheter insertion and administer a subsequent 1000 units for each additional hour of the procedure. Newer agents, such as low molecular weight heparin, platelet IIb/IIIa inhibitors, direct thrombin inhibitors, or recombinant hirudin, have been available and can be used either alone or in conjunction with heparin, particularly in patients who are sensitive to unfractionated heparin. After the procedure, all patients are placed on antiplatelet therapy, such as aspirin. Additional antiplatelet agents, such as clopidogrel (Plavix), are given to selected patients with stent placement for at least 6 weeks after lower extremity interventions unless otherwise contraindicated.

**Percutaneous Transluminal Balloon Angioplasty.** After the lesion is crossed with a wire, an appropriated balloon angioplasty catheter is selected and tracked along the wire to traverse the lesion. The length of the selected catheter should be slightly longer than the lesion, and the diameter should be equal to the adjacent normal vessel. The balloon tends to be approximately 10% to 20% oversized. The radiopaque markers of the balloon catheter are placed so that they will straddle the lesion. Then, the balloon is inflated with saline and contrast mixture to allow visualization of the insufflation process under the fluoroscopy (Fig. 23-65). The patient may experience mild pain, which is not uncommon. However, severe pain can be indicative of vessel rupture, dissection, or other complications. An angiography is crucial in confirming the intraluminal location of the catheter and absence of contrast extravasation. The inflation is continued until the waist of the atherosclerotic lesion is disappeared and the balloon is at the full profile. Frequently, several inflations are required to achieve a full profile of the balloon (Fig. 23-66). Occasionally, a lower profile balloon is needed to predilate the tight stenosis so that the selected balloon catheter can cross the lesion.

Besides length and diameter, the operators need to be familiar with several balloon characters. Noncompliant and low-compliant balloons tend to be inflated to their preset diameter and offer greater dilating force at the site of stenosis. Low-compliant balloons are the mainstay for peripheral intervention. A balloon with a low profile is used to minimize complications at the entry site and for crossing the tight lesions. Upon inflation, most balloons do not rewrap to their preinflation diameter and assume larger profiles. Furthermore, trackability, pushability, and crossability of the balloon should be considered when choosing a particular type of balloon. Lastly, shoulder length is an important characteristic when performing PTA to avoid injury to the adjacent arterial segments. After PTA, a completion angiogram is performed while the wire is still in place. Leaving the wire in place provides access for repeating the procedure if the result is unsatisfactory.

PTA is an established and effective therapy for select patients with lower extremity occlusive diseases. Studies have shown that PTA of femoropopliteal segment achieved over 90% technical success rate and 38% to 58% 5-year primary patency rates. However, efficacy of PTA is highly dependent on anatomic selection and patient condition. PTA of lesions longer than 7 to 10 cm offers limited patency, whereas PTA of shorter lesions, such as those less than 3 cm, has fairly good results. Lofberg and associates performed 127 femoropopliteal PTA

---

**Figure 23-65.** A. Angiogram demonstrating a focal stenosis in the superficial femoral artery (arrow). B. This lesion was treated with a balloon angioplasty catheter that inflated a dilating balloon and expanded the flow lumen. C. Completion angiogram demonstrating satisfactory radiographic result.
procedures and reported a primary 5-year success rate of 12% in limbs with occlusion longer than 5 cm versus 32% in limbs with occlusion less than 5 cm in length.\textsuperscript{131} Occlusive lesions have much worse initial technical success rates than stenotic lesions. Concentric lesions respond better to PTA than eccentric lesions, and heavy calcifications have a negative impact on success rates. A meta-analysis by Hunink and associates showed that adjusted 5-year primary patencies after angioplasty of femoropopliteal lesions varied from 12% to 68%, with the best results being for patients with claudication and stenotic lesions.\textsuperscript{132} Distal runoff is another powerful predictor of long-term success. Johnston analyzed 254 consecutive patients who underwent femoral and popliteal PTA and reported a 5-year patency rate of 53% for stenotic lesions and 36% for occlusive lesions in patients with good runoff versus a 5-year patency rate of 31% for stenotic lesions and 16% for occlusive lesions in patients with poor runoff.\textsuperscript{149} Literature reviews showed that 5-year patency rates varied from 27% to 67% based on runoff status.\textsuperscript{152}

Due to limited success with infrapopliteal PTA, the indication for infrapopliteal PTA is stringent and reserved for limb salvage. Current patency rates from infrapopliteal PTA can be improved further by proper patient selection, ensuring straight-line flow to the foot in at least one tibial vessel, and close patient surveillance for early reintervention. Possible future advances, including the use of drug-eluting stents, cutting balloons, and atherectomy devices, are being investigated to improve clinical outcomes following endovascular interventions on the tibial arteries. Varty and associates reported a 1-year limb salvage rate of 77% in patients with critical ischemia who underwent infrapopliteal PTA.\textsuperscript{153} In patients with favorable anatomies, the 2-year limb salvage rate after infrapopliteal PTA is expected to exceed 80%.

**Subintimal Angioplasty.** The technique of subintimal angioplasty was first described in 1987 when successful establishment of flow was made by accidental creation of a subintimal channel during treatment of a long popliteal artery occlusion. Subintimal angioplasty is recommended for chronic occlusion, long segment of lesion, and heavily calcified lesions. In addition, this technique is applicable for vessels with diffuse disease and for vessels that had previously failed an intraluminal approach, when it is difficult to negotiate the wire across the entire diseased segment without dissection.

The principle of this technique is to bypass the occlusion by deliberately creating a subintimal dissection plan commencing proximal to the lesion and continuing in the subintimal space before retreating into the true lumen distal to the lesion. The occluded lumen is recanalized through the subintimal plan. Subintimal angioplasty can be performed through either an ipsilateral antegrade or contralateral retrograde approach using the CFA approach. If selecting contralateral CFA puncture, a long guiding sheath is placed across the aortic bifurcation to provide access for the femoropopliteal and infrapopliteal vessels. The subintimal dissection is initiated at the origin of an occlusion by directing the tip of an angled guide wire, usually an angled hydrophilic wire, such as a Glidewire. A supporting catheter is used to guide the tip of the guidewire away from the important collaterals. When the wire is advanced, a loop is naturally formed at the tip of the guidewire. Once the subintimal plan is entered, the wire tends to move freely in the dissection space. Subintimal location of the wire and the catheter can be confirmed by injecting a small amount of diluted contrast. At this point, the wire and the catheter are then advanced along the subintimal plan until the occlusion segment is passed. A loss of resistance is often encountered as the guidewire reenters the true lumen distal to the occlusion. Recanalization is confirmed by advancing the catheter over the guidewire beyond the point of reentry and obtaining an angiogram. This is followed by a balloon angioplasty. To confirm the patency following balloon dilatation, a completion angiogram is performed prior to withdrawing the catheter and wire. If flow is impaired, repeat balloon

\[ \text{Figure 23-66. A. Angiogram demonstrating a segmental occlusion in the distal superficial femoral artery (single arrow). B. This lesion was treated with cryoplasty, which lowered the balloon catheter temperature to a temporary freezing state during the balloon angioplasty procedure (double arrows). C. Completion angiogram demonstrated satisfactory result with no evidence of vessel dissection.} \]
dilatation may be necessary. Frequently, a stent is required to maintain a patent lumen and treat residual stenosis if more than 30% luminal reduction is confirmed on completion angiogram.

Multiple studies have demonstrated the efficacy of subintimal angioplasty. Bolia and colleagues reported their extensive experiences on subintimal angioplasty for treating long-segment occlusions of infrainguinal vessels.154 They achieved a technical success rate of over 80% for both femoropopliteal and tibial arteries. One-year patency rates varied from 53% for infrapopliteal vessels to 71% for femoropopliteal segments. Limb salvage rates reached over 80% at 12 months. They also reported that the factors influencing patency are smoking, number of runoff vessels, and occlusion length. Studies by other groups showed similar results.155 Treiman and colleagues treated 25 patients with 6- to 18-cm femoropopliteal occlusion and achieved a technical success rate of 92% and a 13-month primary patency rate of 92%,156 whereas Lipsitz and associates reported a technical success rate of 87% in 39 treated patients and a 12-month cumulative patency rate of 74%.155 In addition, Ingle and associates reported a technical success rate of 87% in 67 patients with femoropopliteal lesions and a 36-month limb salvage rate of 94%.157 As demonstrated herein, although technical success rates are similar in most series, the patency rates vary widely in different studies. Patient selection, anatomic character, and lesion locations may account for the wide range of outcomes.

Stent Placement. Although suggested by Dotter during the late 1960s, the use of an endoluminal stent was not pursued until the limitations of PTA were widely recognized. There are several situations where stent placement is appealing. The primary indication is the potential salvage of an unacceptable angioplasty result. Stent placement is typically used when residual stenosis after PTA is 30% or greater. An endoluminal stent is also used for dissection, perforation, and other PTA complications. Primary stent placement has become a viable alternative for treating ulcerative lesions that may potentially be the source for embolization. Primary stent is also used to treat occlusive lesions that have a tendency for reocclusion and distal embolization after PTA. In addition, an endoluminal stent is potentially beneficial for early restenosis after PTA. Drug-eluting stents are currently under investigation in the United States and may be promising in decreasing restenosis rates.

Although technical success rates are high, published series on femoropopliteal artery stents show that patency rates are comparable to PTA alone, with primary patency rates varying from 18% to 72% at 3 years.158 Gray and associates stented 58 limbs after suboptimal PTA for long SFA lesions and demonstrated a 1-year primary patency rate of 22%.159 However, Mewissen treated 137 limbs using self-expanding SMART nitinol stents in patients with TASC A, B, and C femoropopliteal lesions and reported a 1-year primary patency of 76% and a 24-month primary patency rate of 60%.160 Appropriate patient selection and the anatomic characteristics of the lesions are crucial in the success of treatment outcomes. Additionally, stent characteristics may contribute to the patency rate.

Several clinical studies have demonstrated the significant improvements of the new generation of nitinol stents for the SFA lesions: the German Multicenter Experience, the Mewissen trial, the BLASTER Trial, and the SIROCCO trial.161 The German Multicenter Experience was a retrospective review of 111 SFA stenting procedures and predicted that the 6-month patency rate for SMART stents was 82% versus 37% for the Wallstent. The BLASTER (Bilateral Lower Arterial Stenting Employing Reopro) Trial evaluated the feasibility of using nitinol stents with and without intravenous abciximab for the treatment of femoral artery disease, and the preliminary results showed a 1-year clinical patency rate of 83%.162

Furthermore, the drug-eluting stent, which proved effective in decreasing restenosis in coronary intervention, may offer another promising alternative in lower extremity diseases. The drug released over a period of time interferes with smooth muscle cell proliferation, the main cellular element and source of extracellular matrix—producing restenosis. The first drug-eluting stent clinical trial used Cordis Cypher SMART stents coated with sirolimus (SIROCCO trial).163 The SIROCCO results showed binary in-lesion restenosis rates of 0% in the sirolimus-eluting group versus 23.5% in the noneluting group at 6-month follow-up angiography. The ParADISE (Preventing Amputations Using Drug-Eluting Stents) Trial investigated the efficacy and safety of using balloon-expandable drug-eluting stents to prevent amputations in patients with below-the-knee critical limb ischemia.164 One hundred six patients (118 limbs) were treated with drug-eluting stents in this prospective, nonrandomized trial. There were 228 drug-eluting stents implanted (83% Cypher [Cordis, Johnson & Johnson, Warren, NJ], 17% Taxus [Boston Scientific, Maple Grove, MN]). The average length treated was 60 mm. The 3-year cumulative incidence of amputation was 6%, the survival rate was 71%, and the amputation-free survival rate was 68%. Only 12% of patients who died had a preceding major amputation. Rutherford category, age, creatinine level, and dialysis were predictors of death but not amputation. Target limb revascularization occurred in 15% of patients.

Stent Graft. The concept of endoluminal bypass using stent graft in treating atherosclerotic SFA disease has been entertained. A stent graft is placed percutaneously across a long segment or multiple segments of lesions and is used to create a femoropopliteal bypass. Theoretically, endobypass has the potential of being as successful as surgical bypass graft by relining the vessel wall in its anatomic position without the negative impact of anastomosis. Stent grafts can be divided into two categories: unsupported and fully supported. The unsupported grafts consist of segments of bypass graft, such as PTFE, with an expandable stent at one or both ends. The unsupported grafts are flexible with a low profile, but prone to external compression. The supported stent grafts consist of a metallic skeleton covered with graft fabric. The presence of a dense metal skeleton promotes an extensive inflammatory response and increases the risk of thrombosis. There is no FDA-approved stent graft for peripheral intervention. However, Viabahn (WL Gore & Associates, Flagstaff, AZ) is the most commonly used device in the United States and is composed of an ultra-thin PTFE graft externally supported by self-expanding nitinol meshwork. The Viabahn device has a specific delivery mechanism by pulling back the attached string, which results in proximal-to-distal delivery of the endoprosthesis.

Although it is an intriguing concept, data on endobypass results are limited, and the graft thrombosis rate is high. Additionally, covering major collateral vessels can potentially jeopardize the viability of the limb if stent graft occlusion occurs. Bauermeister treated 35 patients with Hemobahn and reported a 28.6% occlusion rate at an average 7-month follow-up.165 Kodera and colleagues recently conducted a prospective, randomized study comparing covered PTFE/nitinol self-expanding stent grafts with prosthetic above-the-knee femoropopliteal bypass. Fifty limbs were randomized into each group. Primary patency
at 1 year was approximately 74% for both cohorts, with a mean follow-up of 18 months. The covered nitinol/PTFE stent graft in the SFA had a 1-year patency comparable to surgical bypass, with a significantly shorter hospital stay (0.9 vs. 3.1 days). A recent randomized prospective study comparing the treatment of SFA occlusive disease percutaneously with an expanded PTFE (ePTFE)/nitinol self-expanding stent graft (stent graft) versus surgical femoral to above-knee popliteal artery bypass with synthetic graft material showed no difference between the two groups with respect to primary or secondary patency rates at 48 months. Mean total lesion length of the treated arterial segment in the stent graft group was 25.6 cm. The stent graft group demonstrated a primary patency of 72%, 63%, 63%, and 59% with a secondary patency of 83%, 74%, 74%, and 74% at 12, 24, 36, and 48 months, respectively. The surgical femoral-popliteal group demonstrated a primary patency of 76%, 63%, 63%, and 58% with a secondary patency of 86%, 76%, 76%, and 71% at 12, 24, 36, and 48 months, respectively. The authors concluded that ePTFE/nitinol self-expanding stent graft placement can be offered as an alternative to treatment of the SFA segment for revascularization when prosthetic bypass is being considered or when autologous conduit is unavailable.

Atherectomy. The basic principle of atherectomy is to remove the atheroma from obstructed arterial vessels. The currently available atherectomy devices can be generally categorized into directional, nondirectional, orbital, and rotational types based on their mechanism. A few examples of FDA-approved atherectomy devices are Simpson AtheroCath (DVI, Redwood City, CA), Transluminal Extraction Catheter (Interventional Technologies, San Diego, CA), Thoratec recanalization arterial catheter (Thoratec, Pleasanton, CA), Auth Rotablator (Heart Technologies, Redmond, WA), SilverHawk system (Fox Hollow Technologies, Redwood City, CA), Jetstream atherectomy system (Bayer, Indianola, PA), Diamondback 360° orbital atherectomy device (Cardiovascular Systems, Inc, St. Paul, MN), and Rotablator system (Boston Scientific Corporation, Natick, MA). These devices either cut and remove or pulverize the atheroma plaques.

The Simpson AtheroCath has a directional cutting element that is exposed to one-third of the circumference of the arterial wall. The atheroma protruding into the window is excised and pushed into the collection chamber. The Transluminal Extraction Catheter has an over-the-wire nondirectional cutter mounted on the distal end of a torque tube. The excised atheroma is simultaneously removed by aspiration through the torque tube. The Thoratec recanalization arterial catheter is a nondirectional, noncoaxial, atheroablative device. The rotating cannula tip pulverizes the atheromatous lesion into minute particles. The Auth Rotablator is a nondirectional, coaxial, atheroablative device with a metal burr embedded with fine diamond chips. SilverHawk device is a monorail catheter designed to overcome the drawbacks of a directional atherectomy catheter. The working end consists of a hinged housing unit containing a carbide cutting blade. The blade is activated from the motor drive unit, and the catheter is then advanced through the length of the lesion. Once each pass is completed, the cutter then packs the tissue into the distal end of the nosecone to maximize collection capacity. The SilverHawk can then either be removed or torqued to treat a different quadrant in the same lesion or other lesions. Jetstream atherectomy system is a rotating, aspirating catheter with tip sizes of 1.6 and 1.8 mm for tibial arteries, and an expandable catheter with a tip size ranging from 2.1 to 3.4 mm for active removal of atherosclerotic debris and thrombus. The Diamondback 360° orbital atherectomy device uses a drive shaft with an eccentrically mounted, diamond-coated crown to create an orbital spin. As the speed of the crown increases from centrifugal force, it sands wider spaces, thereby providing variability in its working range. It can create a lumen that is >1.75 times the crossing profile depending on the size of the grit and the eccentricity of the offset. The greater the speed of the crown, the larger is the arc of debulking and, ultimately, the resultant lumen size. A constant flow of saline solution is delivered by a roller pump that lubricates the device and helps to flush the debris. The Rotablator system high-speed rotational device uses calcium ablation to achieve larger lumens. It has been used for more than 20 years to treat challenging, calcified coronary artery disease. The diamond-coated burr is designed to preferentially engage calcium and modify lesion compliance.

Despite the promising early technical and clinical success, the mid- and long-term results have been disappointing due to high incidence of restenosis. A multicenter clinical registry of plaque atherectomy in patients with femoropopliteal occlusive disease showed potential clinical efficacy of this technology, as the 6- and 12-month rates of survival free of target lesion revascularization were 90% and 80%, respectively. Importantly, nearly three-quarters (73%) of patients treated with plaque excision modality did not require adjunctive endovascular therapy as infrainguinal stenting was necessary in only 6.3% of lesions. Results from the TALON registry support the role of plaque excision in selected patients with lower extremity arterial disease.

Recent technologic advances have made it possible to increase the spectrum of treatable peripheral arterial lesions with high acute procedure success rates. Recently presented data from multiple registries have shown some promising results in terms of short-term primary patency rates and freedom from unplanned major amputation. Randomized clinical trials, which may provide conclusions on the effectiveness of these procedures, are expected.

Laser Atherectomy. Since laser atherectomy was reported in the 1960s, a variety of innovative approaches have been developed trying to overcome the limitation of laser angioplasty. Recent developments in Excimer laser technology have led to increased optimism regarding the ability to safely deliver laser energy. Excimer laser atherectomy approved by the FDA for peripheral artery intervention employs precision laser energy control (shallow tissue penetration) and safer wavelengths (ultraviolet as opposed to the infrared spectra in older laser technology), which decrease perforation and thermal injury to the treated vessels.

A laser atherectomy catheter, with diameters varying from 0.9 to 2.5 mm, is tracked over the guidewire to the desired target. Once activated, the Excimer laser uses ultraviolet energy to ablate the lesion and create a nonthrombogenic arterial lumen. This lumen is further diluted by an angioplasty balloon. Because the Excimer laser can potentially reduce the rate of distal embolization by evaporating the lesion, it may be used as an adjunct tool for ostial lesions and lesions that can be traversed by a wire but not an angioplasty balloon catheter.

Several studies regarding the use of Excimer laser atherectomy combined with balloon angioplasty on lower extremity occlusive disease have shown promising clinical outcomes. The Peripheral Excimer Laser Angioplasty (PELA) trial involved 318 patients with chronic SFA occlusion
and achieved a technical success rate of 83.2%, a 1-year primary patency rate of 33.6%, and an assisted-primary patency rate of 65%. Steinkamp and colleagues treated 127 patients with long-segment popliteal artery occlusion using laseratherectomy followed by balloon angioplasty and reported a 3-year primary patency rate of 22%. A multicenter clinical trial, the Laser Angioplasty for Critical Limb Ischemia (LACI) trial, supports the efficacy of this treatment modality in selected patients, with 6-month primary patency and clinical improvement rates of 33% and 89%, respectively. The technology and devices continue to evolve. With the Turbo-Booster and Turbo-Tandem technologies (Spectranetics Corporation, Colorado Springs, CO), the efficacy of plaque reduction was reported to be significantly improved in the ClariPath Excimer Laser System to Enlarge Lumen Openings (CELLO) study. The CELLO study was a single-arm, prospective registry trial conducted at 17 investigational sites in the United States to evaluate the safety and efficacy of a modified laser catheter designed for the endovascular treatment of peripheral artery disease affecting the SFA and proximal popliteal artery. Laser ablation reduced percent diameter stenosis from 77% to 21% after adjunctive therapy with balloon angioplasty or balloon angioplasty with stenting; 12.3% patients did not receive postlaser adjunctive therapy. Patency rates were 59% and 54% at 6 and 12 months, respectively. Target lesion revascularization was not required in 76.9% of CELLO participants within the 1-year follow-up.

Complications of Endovascular Interventions

Angioplasty-Related Complications. Complications related to PTA vary widely and include dissection, rupture, embolization, pseudoaneurysms, restenosis, hematoma, and acute occlusion secondary to thrombosis, vasospasm, or intimal injury. Clark and associates analyzed the data from 205 patients in the SCVIR Transluminal Angioplasty and Revascularization (STAR) registry and reported a complication rate of 7.3% for patients undergoing femoropopliteal angioplasty. Minor complications accounted for 75% of the cases, including distal emboli (41.7%), puncture site hematomas (41.7%), contained vessel rupture (8.3%), and vagal reactions (8.3%). In another study, Axisa and colleagues reported an overall rate of significant complications for patients undergoing PTA of the lower extremities of 4.2%, including retroperitoneal bleeding (0.2%), false aneurysm (0.2%), ALI (1.5%), and vessel perforation (1.7%).

Complications limiting the application of subintimal angioplasty are parallel to those of PTA. A study investigating the use of subintimal angioplasty in 65 patients with SFA occlusion found that complications developed in 15% of patients. These complications included significant stenosis (44%), SFA rupture (6%), distal embolization (3%), retroperitoneal hemorrhage (1.5%), and pseudoaneurysm (1.5%). Additional complications reported included perforation, thrombosis, dissection, and extensions beyond the planned reentry site. Importantly, damage to significant collateral vessels may occur in 1% to 1.5% of patients who undergo subintimal angioplasty. If a successful channel is not achieved in this situation, the patient may have a compromised distal circulation that necessitates distal bypass. Cryoplasty is a modified form of angioplasty, and long-term results on lower extremity intervention are not yet available. Fava and associates treated 15 patients with femoropopliteal disease and had a 13% complication rate involving guidewire dissection and PTA-induced dissection of a tandem lesion remote to the cryoplasty zone.

Endoluminal Stent— and Stent Graft—Related Complications. In addition to the aforementioned complications with angioplasty, endoluminal stent is associated with the risk of stent fracture and deformity. The adductor canal has nonlaminar flow dynamics, especially with walking. The forces exerted on the SFA include torsion, compression, extension, and flexion. These forces exert significant stress on the SFA and stents. In addition, the lower extremity is subject to external trauma, which further increases the risk of stent deformity and fracture (Fig. 23-67). The SIROCCO study showed that stent fracture, although not associated with clinical symptoms, occurs in 18.2% of the procedures involving both drug-eluting stents and control stents.

Stent grafts may present the additional complication of covering important collaterals, which results in compromised distal circulation. A prospective study evaluating Hembahn stent grafts in the treatment of femoropopliteal arterial

Figure 23-67. Due to various geometric forces, including torsion, compression, extension, and flexion, exerted on the superficial femoral artery (SFA), stent fracture (arrows) is a known complication following SFA stent placement.
occlusions demonstrated a 23% immediate complication rate including distal embolization (7.7%), groin hematoma (13.5%), and arteriovenous fistula (1.9%).

**Atherectomy-Related Complications.** Overall complication rates associated with atherectomy range from 15.4% to 42.8%, including spasm, thrombosis, dissection, perforation, distal emboli, no reflow, and hematoma. Jahnke and associates conducted a prospective study evaluating high-speed rotational atherectomy in 15 patients with infrapopliteal occlusive disease. They yielded a 94% technical success rate, but success was complicated by vessel rupture (5%), distal embolization (5%), and arterial spasm (5%). Although Excimer laser atherectomy reduces embolic events by evaporating the lesion, embolization still remains a problematic complication. Studies show that distal embolic events occur in 3% to 4% of procedures and perforation occurs in 2.2% to 4.3% of cases. Other complications associated with laser atherectomy therapy include acute thrombosis, vasospasm, direct vessel injury, and dissection.

**Surgical Treatment for Chronic Limb Ischemia due to Femoropopliteal Disease**

**Endarterectomy.** Endarterectomy has a limited, albeit important role in lower extremity occlusive disease. It is most frequently used when there is disease in the CFA or involving the PFA. In this procedure, the surgeon opens the diseased segment longitudinally and develops a cleavage plane within the media that is developed proximally and distally. This permits the inner layer containing the atheroma to be excised. Great care must be taken at the distal end of the endarterectomy to either ensure a smooth transition or tack down the distal endpoint to prevent the flow from elevating a potentially occlusive atheromatous flap. Currently, there is essentially no role for long open endarterectomy in the treatment of SFA stenoses or occlusions. The high incidence of restenosis is what limits utility of endarterectomy in this location. Short-segment stenoses are more appropriately treated with balloon angioplasty. Endarterectomy using a catheter-based approach (e.g., Moll endarterectomy device) supplemented with stent grafting or stenting across the endpoint of the endarterectomy is currently being reevaluated; however, no long-term data are available.

**Bypass Grafting.** Bypass grafting remains the primary intervention for lower extremity occlusive disease. The type of bypass and the type of conduit are important variables to consider. Patients with occlusive disease limited to the SFA, who have at least 4 cm (ideally 10 cm) of normal popliteal artery reconstituted above the knee joint, and with at least one continuous vessel to the foot can be treated with an above-knee femoropopliteal bypass graft. Despite the fact that in this above-knee location, the differential patencies between prosthetic (PTFE) and vein graft are comparable, undoubtedly, it remains ideal to use a saphenous vein as the bypass conduit if possible. Saving the vein for future coronary artery bypass or distal leg bypass grafting has been shown to be a flawed argument. One must also consider that the consequences to the vascular outflow after a thrombosed prosthetic are worse than after a thrombosed vein graft.

When the disease extends to involve the popliteal artery or the tibial vessels, the surgeon must select an appropriate outflow vessel to perform a bypass. Suitable outflow vessels are defined as uninterrupted flow channels beyond the anastomosis into the foot. Listed in order of descending preference, they are as follows: above-knee popliteal artery, below-knee popliteal artery, posterior tibial artery, anterior tibial artery, and peroneal artery. In patients with diabetes, it is frequently the peroneal artery that is spared. Although it has no direct flow into the foot, collateralization to the posterior tibial and anterior tibial arteries makes it an appropriate outflow vessel. There is no objective evidence to preferentially select tibial over peroneal arteries if they are vessels of equal caliber and quality. The dorsalis pedis, which is the continuation of the anterior tibial in the foot, is frequently spared from atherosclerotic disease and can be used as a target for distal bypasses. Patency is affected by the length of the bypass (longer bypasses have reduced patency), quality of the recipient artery, extent of runoff to the foot, and quality of the conduit (saphenous vein/graft). Five-year assisted patency rate for infrapopliteal venous bypasses is 60%. Venous conduits have also been shown to be suitable for bypasses to plantar arteries. In this location, venous conduits have a 3-year limb salvage rate of 84% and a 3-year secondary patency rate of 74%. A meta-analysis suggests unsatisfactory results when PTFE-coated grafts are used to bypass to infrapopliteal arteries. In this location, prosthetic grafts have a 5-year primary patency rate of 30.5%. Additionally, due to distal embolization and compromise of outflow vessels, prosthetic graft occlusion may have more severe consequences than vein graft occlusion.

Two techniques are used for distal bypass grafting: reversed saphenous vein grafting and in situ saphenous vein grafting. There is no difference in outcomes (patency or limb salvage) between these techniques. In the former, the vein is excised in its entirety from the leg using open or endoscopic vein harvest, reversed to render the valves nonfunctional, and tunneled from the CFA inflow to the distal target vessels. End-to-side anastomoses are then created.

Several adjunctive techniques have been used to try to improve the patency of bypass grafts to tibial arteries. Creation of an arteriovenous fistula at the distal anastomosis is one option, but it has not been shown to improve patency. Another method involves creating varying configurations of vein cuffs or patches at the distal anastomosis in an attempt to streamline the flow and to reduce the likelihood of neoimal hyperplasia. Results with this approach are more promising, especially when done to improve patency of a below-the-knee prosthetic; however, there are no definitive comparative trials that support the superiority of one configuration over another.

**Amputation.** Primary amputation is defined as an amputation that is performed without a prior attempt at surgical or endovascular revascularization. It is rarely necessary in patients who, as a result of neglect, present with class III CLI. Primary amputation may play a role in patients with critical limb ischemia who are deemed nonambulatory because of knee contractures, debilitating strokes, or dementia.

**Complications of Surgical Reconstruction**

**Vein Graft Stenoses.** Fifteen percent of vein grafts will develop intrinsic stenoses within the first 18 months following implantation. Consequently, patients with vein grafts were entered into duplex surveillance protocols (scans every 3 months) to detect elevated (>300 cm/s) or abnormally low (<45 cm/s) graft velocities early. Stenoses greater than 50%, especially if associated with changes in ABI, should be repaired to prevent graft thrombosis. Repair usually entails patch angioplasty or short-segment venous interposition, but PTA/stenting
is an option for short, focal lesions. Grafts with stenoses that are identified and repaired prior to thrombosis have assisted-primary patency identical to primary patency, whereas a thrombosed autogenous bypass has limited longevity resulting from ischemic injury to the vein wall. Secondary patency is markedly inferior to primary assisted patency. The recommendation for routine duplex ultrasound surveillance of autogenous infrainguinal bypasses was recently brought into question by a randomized controlled trial that demonstrated no cost benefit or quality-of-life improvement in patients with femoropopliteal venous bypasses after 18 months. Many surgeons continue with programs of vein graft surveillance, as has been suggested in older trials, awaiting further confirmation of the findings from the more recent study. When intervening in a failing infrainguinal bypass, the original indication for surgery is an important consideration. Limb salvage rates for occluded grafts are better if the indication for the original bypass was claudication rather than rest pain or tissue loss. An acutely occluded infrainguinal graft (≤30 postoperative days) has a 25% limb salvage rate.

Limb Swelling. Limb swelling is common following revascularization and usually returns to baseline within 2 to 3 months. The etiology is multifactorial with lymphatic interruption, interstitial edema, and disruption of venous drainage all contributing. Limb swelling tends to worsen with repeat revascularization (see Table 23-22).

Wound Infection. Since the most common inflow vessel for distal bypass is the CFA, groin infection is common and occurs in 7% of cases. When an autogenous conduit such as the saphenous vein is used, most infections can be managed with local wound care because the infection involves the subcutaneous tissue or skin rather than infection of the actual vein. When a prosthetic graft has been used, management of graft infection is a major undertaking. Infection of a lower extremity prosthetic bypass graft is associated with a significant amputation rate because of the tendency for graft thrombosis and anastomotic disruption. Prosthetic graft infections cannot be eradicated with antibiotics and mandate graft excision and complex revascularization using a vein if available.

Choice of Conduit for Infrainguinal Bypass Grafting

Autogenous Vein. The autogenous vein is superior to prosthetic conduits for all infrainguinal bypasses, even in the above-knee position. This preference is applicable not only for the initial bypass but also for reoperative cases. For long bypasses, the ipsilateral great saphenous vein, contralateral great saphenous vein, small saphenous vein, arm vein, and spliced vein are used in decreasing order of preference. If only a short segment of vein is missing, the SFA can be endarterectomized and the proximal anastomosis performed distally to decrease the length of the conduit and to avoid harvesting and splicing additional vein. When the great saphenous vein is not available and a relatively short bypass is necessary, the arm vein or small saphenous vein is effective. The small saphenous vein is of particular utility when a posterior approach is used. If a longer bypass with vein is necessary, the arm vein is preferable because it is less awkward to harvest. Another conduit alternative is to harvest the upper arm basilic, median cubital, and cephalic veins in continuity, while incising valves in the basilic segment and using the cephalic segment in reversed configuration to provide a relatively long, unspliced autogenous conduit.

Cryopreserved Grafts. Cryopreserved grafts are usually cadaveric arteries or veins that have been subjected to rate-controlled freezing with dimethyl sulfoxide (DMSO) and other cryopreservatives. Cryopreserved vein grafts are more expensive than prosthetic grafts and are more prone to failure. The endothelial lining is lost as part of the freezing process, making these grafts prone to early thrombosis. Cryopreserved grafts are also prone to aneurysmal degeneration. Despite the fact that these grafts have not performed as well as prosthetic bypasses and autogenous vein bypasses in clinical practice, they can still play a role when revascularization is required following removal of infected prosthetic bypass grafts, especially when the autogenous vein is unavailable to create a new bypass through clean tissue planes.

Human Umbilical Vein. Human umbilical vein (HUV) is less commonly used than PTFE because it is thicker and more cumbersome to handle and because of concerns about aneurysmal degeneration. HUV allografts are stabilized with glutaraldehyde and do not have viable cells or antigenic reactivity. These grafts have poor handling characteristics and require extra care when suturing because of an outer Dacron mesh wrapping that is used to decrease aneurysmal degeneration. Dardik and colleagues have reported favorable results after using HUV and an adjunctive distal arteriovenous fistula. One trial comparing HUV with PTFE and saphenous vein showed that HUV was better than PTFE but worse than saphenous vein in terms of 5-year patency in the above-knee location. In a systematic review, HUV appears to perform better than cryopreserved veins in terms of 1-year graft patency in infrainguinal revascularization.

Prosthetic Conduits and Adjunctive Modifications. If a vein is truly unavailable, PTFE or Dacron is the best option for above-knee bypass. The addition of rings to PTFE did not confer benefit in a single prospective, randomized clinical trial. For infrageniculate prosthetic bypasses, use of a vein patch, cuff, or other venous anastomotic modification can improve patency (52% patency at 2 years for PTFE with vein cuff vs. 29% for PTFE with no cuff) and also improve limb salvage (84% vs. 62%).

Although prosthetic grafts are quickly available, easy to handle, and do not require extensive dissection to harvest, their propensity to undergo thrombosis and develop neointimal hyperplasia makes them a less favorable alternative when compared to vein. In a recent review of vein and prosthetic above-knee femoropopliteal bypasses, the 5-year primary patency rates were reported to be 74% and 39%, respectively. Outcomes were even worse for below-knee prosthetic bypasses. Unfortunately, the use of autologous venous conduits is not possible in as many as 30% of patients. The great saphenous vein may be unsuitable because of small size and poor quality or unavailable due to prior harvest.

Methods to improve prosthetic graft performance have consisted of altering the geometry at the distal anastomosis to get the benefit obtained with vein cuffs (Distalflo; Bard Peripheral Vascular, Tempe, AZ) and covalently bonding agents onto the luminal surface with anticoagulant, anti-inflammatory, and anti-proliferative characteristics (Propanen; Gore, Flagstaff, AZ). One randomized trial that compared precuffed PTFE versus PTFE with a vein cuff enrolled 104 patients at 10 centers. Of 89 patients, 47 were randomized to precuffed PTFE bypasses and 44 were randomized to bypasses with a vein cuff. At 1 and
2 years, primary patency rates were 52% and 49% in the precuffed group and 62% and 44% in the vein cuffed group, respectively. At 1 and 2 years, the limb salvage rates were 72% and 65% in the precuffed group and 75% and 62% in the vein cuffed group, respectively. Although numbers are small and follow-up short, the midterm analysis revealed that Distal occluded grafts and PTFE grafts with vein cuff had similar results. The authors concluded that a precuffed graft was a reasonable alternative for infragenicular reconstruction in the absence of saphenous vein.

Other authors have been less optimistic and question whether there is any benefit derived from geometrically altering prosthetic conduits.

Another approach for improving outcomes when using prosthetic for bypass grafts involves bonding anticoagulants to the conduit. The Gore Prophath graft has harpin bonded onto the luminal surface of the PTFE graft using Carmida BioActive Surface (CABS) technology, which immobilizes the harpin molecule with a single covalent bond that does not alter its anticoagulant properties. The harpin binding does not alter the microstructure and handling characteristics of the PTFE. A prospective, randomized trial by Devine and colleagues suggested that harpin-bonded Dacron or PTFE was superior to plain PTFE for above-knee popliteal bypasses. The 3-year primary patency rate for the harpin-bonded grafts was 55% compared with 42% for PTFE (P < 0.044). Both of these patency rates are inferior to great saphenous vein grafts; however, if the improved results with harpin bonding continue to be substantiated, then harpin-bonded prosthetic grafts will become the preferred conduit for above-knee bypass in the absence of suitable vein.

A recent review of available studies with this graft showed an 80% 1-year patency rate for below-knee bypasses. Randomized controlled clinical trials with more patients and longer follow-up are necessary to validate whether the Prophath vascular graft is superior to other prosthetics and whether it is comparable to autogenous vein for below-knee interventions.

**Clinical Results of Surgical and Endovascular Interventions for Femoropopliteal Occlusive Disease**

Balloon angioplasty of the femoropopliteal vessels has not enjoyed the degree of success seen with iliac angioplasty. Patency in this region is dependent on whether the patient presents with claudication versus limb-threatening ischemia, the status of the distal runoff vessels, and lesion morphology. Initial technical success for femoropopliteal angioplasty is seen in 80% to 90% of cases, with failure to cross a lesion occurring in 7% of stenoses and 18% of occlusive lesions. Studies have shown that PTA of the femoropopliteal segment achieved a greater than 90% technical success rate and had a 38% to 58% 5-year primary patency rate. PTA of lesions longer than 7 to 10 cm results in compromised patency, whereas PTA of shorter lesions (<3 cm) gives fairly good results. Lofberg and colleagues performed 127 femoropopliteal PTA procedures and reported a primary patency rate at 5-year follow-up of 12% in limbs with occlusion longer than 5 cm versus 32% in limbs with occlusion less than 5 cm in length.

Occlusive lesions have much worse initial technical success rates than stenotic lesions. Concentric lesions respond better to PTA than eccentric lesions, and heavy calcifications have a negative impact on success rates. Distal runoff is another powerful predictor of long-term success.

Johnston analyzed 254 consecutive patients who underwent femoropopliteal PTA and reported a 5-year patency rate of 53% for stenotic lesions and 36% for occlusive lesions in patients with good runoff versus a 5-year patency rate of 31% for stenotic lesions and 16% for occlusive lesions in patients with poor runoff. A meta-analysis by Humink and colleagues showed that adjusted 5-year primary patencies after angioplasty of femoropopliteal lesions varied from 12% to 68%, with the best results occurring in patients with claudication and stenotic lesions. Although the initial technical success is better for stenoses than occlusions, long-term patency rates for stenoses and short occlusions have been variable, and there have been conflicting results regarding the efficacy of stent use. Early published series that examined efficacy of femoropopliteal artery stents showed patency rates that were comparable to standalone PTA, with primary patency rates varying from 18% to 72% at 3 years. Patient selection and the anatomic character of the lesions may play important roles in the outcomes. Additionally, stent characteristics may contribute to the patency rate. Several recent clinical studies have demonstrated significant improvements in patency when the newer generations of nitinol stents are used to treat SFA lesions.

Mewissen treated 137 lower limbs in 122 patients with CLI, secondary to TASC A (n = 12) or TASC B or C (n = 125) lesions in the SFA. Patients were treated with Cordis SMART self-expanding nitinol stents. Binary restenosis (>50%) was measured by standard duplex velocity criteria at various postintervention intervals. Primary stent patency, defined as absence of binary restenosis in this study, was calculated by life-table methods from the time of intervention. The mean lesion length was 12.2 cm (range, 4–28 cm). The technical success was 98%. Mean follow-up was 302 days. The primary stent patency rates were 92%, 76%, 66%, and 60% at 6, 12, 18, and 24 months, respectively. Ferreira and colleagues treated 59 patients who had 74 femoropopliteal lesions (60% TASC D) with Zilver nitinol self-expanding stents (Cook, Bloomington, IN). Mean recanalization length was 19 cm (range, 3–53 cm). Mean follow-up time was 2.4 years (range, 3 days–4.8 years). Kaplan-Meier estimates for primary patency rates were 90%, 78%, 74%, 69%, and 69% at 1, 2, 3, 4, and 4.8 years, respectively.

There is general agreement that for suboptimal PTA of an SFA lesion, stent placement is indicated, but a recent randomized trial by Schillinger and associates suggests that primary stenting results in lower restenosis rates than PTA and selective stenting. Restenosis rates at 2 years were 45.7% versus 69.2% in favor of primary stenting compared with PTA and optional secondary stenting using an intent-to-treat analysis (P = 0.031). Consistently, stenting, both primary and selective, was superior to standalone PTA with respect to the occurrence of restenosis (49.2% vs. 74.3%; P = 0.028) by a treatment-received analysis.

Nitinol bare metal stents that are designed specifically for below-knee interventions are showing very encouraging results. Bosiers and colleagues reported their 12-month results using the commercially available non–drug-eluting Xpert (Abbott Vascular, Santa Clara, CA) nitinol stent system in below-knee arterial interventions.

They had a 12-month primary patency rate of 76.3% and a limb salvage rate of 95.9%. They followed patients for 12 months and performed angiography with quantitative vessel analysis on the 73% of patients available. Angiography revealed a binary restenosis rate (>50%) of only 20.5%, which is comparable to well-accepted coronary drug-eluting stent study outcomes. The authors attributed this optimal performance to the maintenance of flow dynamics because
the stent was specifically designed for use in small vessels. Kickuth and colleagues also have obtained good results using the Xpert stent. After stent placement, the primary cumulative patency rate at 6 months for the study group of 35 patients was 82%. The sustained clinical improvement rate as evidenced by improved ABI was 80%, and freedom from major amputation was 100% at the 6-month follow-up. The rate of major complications was 17%.204

Wolf and colleagues published a multicenter, prospective randomized trial comparing PTA with bypass in 263 men who had iliac, femoral, or popliteal artery obstruction.205 In 56 patients, cumulative 1-year primary patency rate was 43% after PTA and 82% after bypass surgery, demonstrating that for long SFA stenoses or occlusions, surgery is better than PTA. Another recent randomized study (BASIL trial) of 452 patients with CLI demonstrated no difference in amputation-free survival at 6 months between surgery and PTA/stenting.206 The authors commented that surgery was somewhat more expensive and recommended that endovascular intervention should be used as first-line therapy especially in medically unfit patients. They did conclude that at the 2-year follow-up, healthy patients without medical comorbidities derived greater benefit from surgery because it was associated with decreased need for reintervention and had a decreased hazard ratio in terms of all-cause mortality. The recently published randomized prospective study comparing the treatment of SFA occlusive disease percutaneously with an ePTFE/nitinol self-expanding stent graft versus surgical femoral to above-knee popliteal artery bypass with synthetic graft material showed no difference between the two groups with respect to primary or secondary patency rate at four years.167 This finding suggests that ePTFE/nitinol self-expanding stent graft placement can be offered as an alternative to treatment of the SFA segment for revascularization when prosthetic bypass is being considered or when autologous conduit is unavailable. Using the 2000 TASC definitions and a Markov state transition model decision analysis, Nolan and colleagues showed that PTA/stenting surpasses bypass efficacy for TASC C lesions if PTA/stenting primary patency is >32% at 5 years, patient age is >80 years, and/or greater saphenous vein bypass operative mortality is >6%.207

NONATHEROSCLEROTIC DISORDERS OF BLOOD VESSELS

The majority of cases of peripheral vascular disease that are seen by vascular surgeons are attributable to underlying atherosclerosis. Nonatherosclerotic disease states that result in arterial pathology are less commonly encountered, but are nonetheless important, as they are potentially treatable lesions that may mimic atherosclerotic lesions and result in vascular insufficiency (see Table 23-18). A thorough knowledge of these rare disease states is important for the practicing vascular surgeon in order to make medical recommendations and provide appropriate surgical treatment.

Giant Cell Arteritis (Temporal Arteritis)

Giant cell arteritis is also known as temporal arteritis, which is a systemic chronic inflammatory vascular disease with many characteristics similar to those of Takayasu’s disease. The histologic and pathologic changes and laboratory findings are similar. Patients tend to be white women over the age of 50 years, with a high incidence in Scandinavia and women of Northern European descent. Genetic factors may play a role in disease pathogenesis, with a human leukocyte antigen (HLA) variant having been identified. Differences exist between Takayasu’s and giant cell arteritis in terms of presentation, disease location, and therapeutic efficacy. The inflammatory process typically involves the aorta and its extracranial branches, of which the superficial temporal artery is specifically affected.

The clinical syndrome begins with a prodromal phase of constitutional symptoms, including headache, fever, malaise, and myalgias. The patients may be initially diagnosed with coexisting polymyalgia rheumatica; an HLA-related association may exist between the two diseases. As a result of vascular narrowing and end-organ ischemia, complications may occur such as visual alterations, including blindness and minal weakness, resulting in acute aortic dissection that may be devastating. Ischemic optic neuritis resulting in partial or complete blindness occurs in up to 40% of patients and is considered a medical emergency. Cerebral symptoms occur when the disease process extends to the carotid arteries. Jaw claudication and temporal artery tenderness may be experienced. Aortic lesions are usually asymptomatic until later stages and consist of thoracic aneurysms and aortic dissections.

The diagnostic gold standard is a temporal artery biopsy, which will show the classic histologic findings of multinucleated giant cells with a dense perivascular inflammatory infiltrate. Treatment regimens are centered on corticosteroids, and giant cell arteritis tends to rapidly respond. Remission rates are high, and treatment tends to have a beneficial and preventative effect on the development of subsequent vascular complications.

Takayasu’s Arteritis

Takayasu’s arteritis is a rare but well-recognized chronic inflammatory arteritis affecting large vessels, predominantly the aorta and its main branches (Table 23-27). Chronic vessel inflammation leads to wall thickening, fibrosis, stenosis, and thrombus formation. Symptoms are related to end-organ ischemia. The acute inflammation can destroy the arterial media and lead to aneurysm formation. This rare autoimmune disease occurs predominantly in women between the ages of 10 and 40 years who are of Asian descent. Genetic studies have demonstrated a high frequency of HLA haplotypes in patients from Japan and Mexico, suggesting increased susceptibility to developing the disease in patients with certain alleles. However, these associations have not been seen in North America. Vascular inflammation leads

<table>
<thead>
<tr>
<th>Table 23-27</th>
<th>Angiographic classification of Takayasu’s arteritis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TYPE</strong></td>
<td><strong>VESSEL INVOLVEMENT</strong></td>
</tr>
<tr>
<td>Type I</td>
<td>Branches from the aortic arch</td>
</tr>
<tr>
<td>Type IIa</td>
<td>Ascending aorta, aortic arch and its branches</td>
</tr>
<tr>
<td>Type IIb</td>
<td>Ascending aorta, aortic arch and its branches, thoracic descending aorta</td>
</tr>
<tr>
<td>Type III</td>
<td>Thoracic descending aorta, abdominal aorta, and/ or renal arteries</td>
</tr>
<tr>
<td>Type IV</td>
<td>Abdominal aorta and/or renal arteries</td>
</tr>
<tr>
<td>Type V</td>
<td>Combined features of types IIb and IV</td>
</tr>
</tbody>
</table>

Involvement of the coronary or pulmonary arteries is designated as C (+) or P (+), respectively.
to arterial wall thickening, stenosis, and eventually, fibrosis and thrombus formation. The pathologic changes produce stenosis, dilation, aneurysm formation, and/or occlusion.

The clinical course of Takayasu’s arteritis begins with a “prepulseless” phase in which the patient demonstrates constitutional symptoms. These include fever, anorexia, weight loss, general malaise, arthralgias, and malnutrition. As the inflammation progresses and stenoses develop, more characteristic features of the disease become evident. During the chronic phase, the disease is inactive or “burned out.” It is during this latter stage that patients most frequently present with bruits and vascular insufficiency according to the arterial bed involved. Laboratory data may show elevations in erythrocyte sedimentation rate, C-reactive protein, and white blood cell count, or conversely, anemia may predominate. Characteristic clinical features during the second phase vary according to the involved vascular bed and include hypertension reflecting renal artery stenosis, retinopathy, aortic regurgitation, cardiac output disturbances, cardiovascular symptoms, angina and congestive heart failure, abdominal pain or gastrointestinal bleeding, pulmonary hypertension, or extremity claudication.

The gold standard for diagnosis remains angiography showing narrowing or occlusion of the entire aorta or its primary branches, or focal or segmental changes in large arteries in the upper or lower extremities. Six types of Takayasu’s arteritis exist and are graded in terms of severity: type I, affecting the aorta and arch vessels; type IIa, affecting the ascending aorta, aortic arch, and branches; type IIb, affecting the ascending aorta, aortic arch and branches, and thoracic descending aorta; type III, affecting the thoracic descending aorta, abdominal aorta, and/or renal arteries; type IV, affecting the abdominal aorta and/or renal arteries; and type V, with combined features of types IIb and IV.²⁰⁸

Treatment consists of steroid therapy initially, with cytotoxic agents used in patients who do not achieve remission. Surgical treatment is performed only in advanced stages, and bypass needs to be delayed during active phases of inflammation. There is no role for endarterectomy, and synthetic or autogenous bypass grafts need to be placed onto disease-free segments of vessels. For focal lesions, there have been reports of success with angioplasty.

**Ehlers-Danlos Syndrome**

Ehlers-Danlos syndrome is one of the more significant inheritable disorders affecting the connective tissue, along with Marfan’s syndrome. This syndrome represents a heterogeneous group of connective tissue disorders (types I through IV) that were first described in 1682 by van Meekeren.²³º It is an autosomal dominant disorder affecting approximately 1 in 5000 persons that is characterized by skin elasticity, joint hypermobility, tissue fragility, multiple ecchymoses, and subcutaneous pseudotumors. Ehlers-Danlos syndrome is a disorder of fibrillar collagen metabolism with identifiable, specific defects that have been found in the collagen biosynthetic pathway that produce clinically distinct forms of this disease. Ten different phenotypes have been described, each with variable modes of inheritance and biochemical defects. Of the four basic types of collagen found in the body, the predominant type in blood vessels is type III. Within the vessel wall, type III collagen contributes to structural integrity and tensile strength and plays a role in platelet aggregation and thrombus formation.

Of the three types of Ehlers-Danlos syndrome that have arterial complications, type IV represents 5% of cases and is the one most likely to be seen by a vascular surgeon. These patients synthesize abnormal type III collagen (mutation COL3A1) and represent 5% of all cases.²³⁰ Affected individuals do not show the typical skin and joint manifestations, and thus typically present for diagnosis when a major vascular catastrophe occurs. In a review of 36 patients with this disorder, Cikrit and colleagues reported a 44% mortality rate from major hemorrhage prior to any surgical intervention.²³¹ In the 20 patients who underwent 29 vascular procedures, there was a 29% mortality rate. Arterial rupture, aneurysm formation, and acute aortic dissection may occur in any major artery, with the most frequent site of rupture being the abdominal cavity. Repair is problematic because the vessel wall is soft and sutures pull through the fragile tissue. Ligation may be the only option in many circumstances.

**Marfan’s Syndrome**

Another heterogeneous inheritable disorder of connective tissue, Marfan’s syndrome is characterized by abnormal musculoskeletal, ocular, and cardiovascular features first described by Antoine Marfan in 1896.²³² The inborn error of metabolism in this syndrome has been localized to the long arm of chromosome 15 (15q21.3). Defects occur in fibrillin, a basic protein in the microfibrillar apparatus that serves as a backbone for elastin, which is one of the main extracellular structural proteins in blood vessels. This is an autosomal dominant gene with high penetrance; however, approximately 15% to 20% of cases are secondary to new spontaneous mutations.

Classic recognizable features of Marfan’s syndrome include tall stature, long limbs (dolichostenomelia), long fingers (arachnodactyly), joint hyperextensibility, chest wall deformities, and scoliosis. Ocular manifestations are flattened corneas, lens subluxation, and myopia. Ninety-five percent of patients have cardiovascular involvement, which may include ascending aortic dilatation, mitral valve prolapse, valvular regurgitation, and aortic dissection. Skin, central nervous system, and pulmonary features may be present as well. Aortic root dilatation will generally occur in all patients. This may not be evident on standard chest radiograph until dilatation has resulted in an ascending aortic aneurysm, aortic valve regurgitation, or dissection. Left untreated, the cardiovascular complications are devastating and reduce the life expectancy to about 40 years for men and slightly higher for women. Death is usually attributable to life-threatening complications of aortic regurgitation, dissection, and rupture after the ascending aorta has dilated to 6 cm or more.

Aggressive medical management with β-adrenergic blocking agents and other blood pressure–lowering regimens is crucial to treatment. Surgical intervention entails replacement of the aortic root with a composite valve graft (e.g., Bentall procedure).²³³ Prophylactic operative repair is indicated for an aneurysm greater than 5.5 cm, with an acceptable perioperative mortality of less than 5%.

**Pseudoxanthoma Elasticum**

Pseudoxanthoma elasticum is a rare inherited disorder of connective tissue that is characterized by an unbalanced elastic fiber metabolism and synthesis, resulting in fragmentation and calcification of the fibers. Clinical manifestations occur in the skin, ocular, gastrointestinal, and cardiovascular systems. Characteristic skin lesions are seen in the axilla, antecubital and popliteal fossae, and groin. The yellow, xanthoma-like papules occur in redundant folds of skin and are said to resemble plucked chicken skin. The inheritance pattern includes both autosomal dominant
and recessive types and has a prevalence of 1 in 160,000 individuals. The ATP-binding cassette subfamily C member 6 (ABCC6) gene has been demonstrated to be responsible, and 43 mutations have been identified, all of which lead to calcification of the internal elastic laminae of medium-sized vessel walls.

Cardiovascular features are common and include premature coronary artery disease, cerebrovascular disease, renovascular hypertension, diminished peripheral pulses, and restrictive cardiomyopathy. Symptom onset typically occurs in the second decade of life, with onset at an average age of 13 years. Patients should be counseled to reduce potential contributing factors for atherosclerosis such as tobacco use and high cholesterol levels. Calcium intake should be restricted in adolescents, as a positive correlation has been found between disease severity and calcium intake. Surgical management involves standard vascular techniques, with the exception that arterial conduits should not be employed in cardiac bypass.

Kawasaki’s Disease
Kawasaki’s disease was first described in 1967, as a mucocutaneous lymph node syndrome occurring in young children. In most studies, more than half the patients are younger than 2 years of age, with a higher prevalence in boys. Although originally described in Japan, the disease is found worldwide. An infectious agent may be causative; however, no specific agent has been identified. Immune activation with the contribution of cytokines, elastases, growth factors, and metalloproteinases is believed to be a mechanism for inflammation and aneurysm formation. Coronary artery aneurysms, the hallmark of the disease, histologically demonstrate a panarteritis with fibrinoid necrosis. Coronary arteriography may show occlusions, recanalization, and localized stenosis, in addition to multiple aneurysms. A variety of constitutional symptoms and signs resulting from systemic vasculitis are present in the acute phase of the illness.

Medical therapy for Kawasaki’s disease clearly decreases the manifestations of coronary artery involvement. Intravenous gamma globulin and aspirin therapy are most successful if begun within the first 10 days of illness. Up to 20% of untreated patients will develop coronary arterial lesions. A long-term, low-dose aspirin therapy regimen is usually recommended.

Inflammatory Arteritis and Vasculitis
Chronic inflammatory arteritis and vasculitis (i.e., inflammatory changes within veins as well as arteries) include a spectrum of disease processes caused by immunologic mechanisms. These terms signify a necrotizing transmural inflammation of the vessel wall associated with antigen-antibody immune complex deposition within the endothelium. These conditions show pronounced cellular infiltration in the adventitia, thickened intimal fibrosis, and organized thrombus. These disease processes may clinically mimic atherosclerosis, and most are treated by corticosteroid therapy or chemotherapeutic agents. Even so, it is important to recognize distinguishing characteristics of each disease in order to establish the course of treatment and long-term prognosis. A classification system of systemic vasculitis by vessel size is shown in Table 23-28.

Behçet’s Disease
Behçet’s disease is a rare syndrome characterized by oral and genital ulcerations and ocular inflammation, affecting males in Japan and the Mediterranean. An HLA linkage has been found, indicating a genetic component to the etiology. Vascular involvement is seen in 7% to 38% of patients and is localized to the abdominal aorta, femoral artery, and pulmonary artery. Vascular lesions may also include venous complications such as deep venous thrombosis or superficial thrombophlebitis. Arterial aneurysmal degeneration can occur; however, this is an uncommon, albeit potentially devastating, complication. Multiple true aneurysms and pseudoaneurysms may develop, and rupture of an aortic aneurysm is the major cause of death in patients with Behçet’s disease.

Histologically, degeneration of the vasa vasorum with surrounding perivascular lymphocyte infiltration is seen, along with thickening of the elastic laminae around the tunica media. Aneurysm formation is believed to be associated with a loss of the nutrient flow and elastic component of the vessels, leading to progressive dilatation. Multiple aneurysms are relatively common, with a reported occurrence of 36% in affected Japanese patients. Furthermore, pseudoaneurysm formation after surgical bypass is common at anastomotic suture lines due to the vascular wall fragility and medial destruction. Systemic therapy with corticosteroids and immunosuppressive agents may diminish symptoms related to the inflammatory process; however, they have no effect on the rate of disease progression and arterial degeneration.

Polyarteritis Nodosa
Polyarteritis nodosa (PAN) is another systemic inflammatory disease process, which is characterized by a necrotizing inflammation of medium-sized or small arteries that spares the smallest blood vessels (i.e., arterioles and capillaries). This disease predominantly affects men over women by a 2 to 1 ratio. PAN develops subacutely, with constitutional symptoms that last for weeks to months. Intermittent, low-grade fevers, malaise, weight loss, and myalgias are common presenting symptoms. As medium-sized vessels lie within the deep dermis, cutaneous manifestations occur in the form of livedo reticularis, nodules, ulcers, and digital ischemia. Skin biopsies of these lesions may be sufficient for diagnosis. Inflammation may be seen histologically, with pleomorphic cellular infiltrates and segmental transmural necrosis leading to aneurysm formation.

Neuritis from nerve infarction occurs in 60% of patients, and gastrointestinal complications occur in up to 50%. Additionally, renal involvement is found in 40% and manifests as microscopic hematuria. Cardiac disease is a rare finding except at autopsy, where thickened, diseased coronary arteries may be seen, as well as patchy myocardial necrosis. Patients may succumb to renal failure, intestinal hemorrhage, or perforation. End-organ ischemia from vascular occlusion or aneurysm rupture can be disastrous.
complications with high mortality rates. The mainstay of treatment is steroid and cytotoxic agent therapy. Up to 50% of patients with active PAN will experience remission with high dosing.

### Radiation-Induced Arteritis

Radiation-induced arteritis results from progressive stenosis due to endothelial damage that leads to cellular proliferation and fibrosis. These are well-described complications of combined irradiation and chemotherapy for the treatment of head and neck malignancy. Arterial lesions are known complications of radiation and are similar to those found in atherosclerotic occlusive disease. A history of therapeutic irradiation to the neck can complicate the management of carotid artery occlusive disease. Radiation-induced damage to blood vessels has been well studied. The small capillaries and sinusoids are most susceptible to radiation effects, as endothelial cells are the most radiosensitive cells. The radiation effects on the medium- and large-sized arteries include myointimal proliferation, with or without lipid deposits, and thrombosis. Characteristically, irregular spindle-shaped cells are seen replacing the normal endothelial cells in the healing phase. Occlusive lesions develop in the irradiated carotid arteries and are either the result of vessel wall fibrosis or, more commonly, due to accelerated atherosclerosis. Neurologic complications related to radiation-induced carotid artery disease are similar to those due to nonirradiated atherosclerotic occlusive disease.

Rupture of the carotid artery has been reported following neck irradiation and is likely related to local wound complication and superimposed infection. The diagnosis of radiation arteritis is based on the clinical history and confirmation of the occlusive lesion by duplex ultrasound, MRA, CTA, or subtraction angiography. Irradiated lesions can be confined to the irradiated segment of the internal carotid artery with the remaining part of the vessel spared of disease. Characteristically, the radiation-induced atherosclerotic lesion does not involve the carotid bulb, unlike the nonirradiated atherosclerotic lesions. The indications for intervention in radiation-induced carotid lesions are the same as previously discussed for atherosclerotic carotid occlusive lesions. However, asymptomatic irradiated carotid artery lesions should be considered for intervention because they can be more prone to progression and development of neurologic complications. Endovascular treatment with carotid angioplasty/stenting has become the treatment of choice for radiation-induced lesions, although surgical endarterectomy and bypass have been shown to be safe. The rate of recurrent stenosis is higher in radiation-induced carotid lesions, whether stented or surgically treated.

### Raynaud’s Syndrome

First described in 1862 by Maurice Raynaud, the term Raynaud’s syndrome applies to a heterogeneous symptom array associated with peripheral vasospasm, more commonly occurring in the upper extremities. The characteristically intermittent vasospasm classically follows exposure to various stimuli, including cold temperatures, tobacco, or emotional stress. Formerly, a distinction was made between Raynaud’s “disease” and Raynaud’s “phenomenon” for describing a benign disease occurring in isolation or a more severe disease secondary to another underlying disorder, respectively. However, many patients develop collagen vascular disorders at some point after the onset of vasospastic symptoms; progression to a connective tissue disorder ranges from 11% to 65% in reported series. Therefore, the term Raynaud’s syndrome is now used to encompass both the primary and secondary conditions.

Characteristic color changes occur in response to the arteriolar vasospasm, ranging from intense pallor to cyanosis to redness as the vasospasm occurs. The digital vessels then relax, eventually leading to reactive hyperemia. The majority of patients are young women less than 40 years of age. Up to 70% to 90% of reported patients are women, although many patients with only mild symptoms may never present for treatment. Geographic regions with cooler, damp climates such as the Pacific Northwest and Scandinavian countries have a higher reported prevalence of the syndrome. Certain occupational groups, such as those who use vibrating tools, may be more predisposed to Raynaud’s syndrome or digital ischemia. The exact pathophysiologic mechanism behind the development of such severe vasospasm remains elusive, and much attention has focused on increased levels of α₂-adrenergic receptors and their hypersensitivity in patients with Raynaud’s syndrome, as well as abnormalities in the thermoregulatory response, which is governed by the sympathetic nervous system.

The diagnosis of severe vasospasm may be made using noninvasive measurements in the vascular laboratory. Angiography is usually reserved for those who have digital ulceration and in whom an embolic or obstructive cause is believed to be present and potentially surgically correctable. Different changes in digital blood pressure will occur in patients with Raynaud’s syndrome. Normal individuals will show only a slight decrease in digital blood pressure in response to external cold stimuli, whereas those with Raynaud’s syndrome will show a similar curve until a critical temperature is reached. It is at this point that arterial closure acutely occurs.

There is no cure for Raynaud’s syndrome; thus, all treatments mainly palliate symptoms and decrease the severity and perhaps frequency of attacks. Conservative measures predominate, including the wearing of gloves, use of electric or chemically activated hand warmers, avoiding occupational exposure to vibratory tools, abstinence from tobacco, and relocating to a warmer, dryer climate. The majority (90%) of patients will respond to avoidance of cold and other stimuli. The remaining 10% of patients with more persistent or severe syndromes can be treated with a variety of vasodilatory drugs, albeit with only a 30% to 60% response rate. Calcium channel-blocking agents such as diltiazem and nifedipine are the drugs of choice. The selective serotonin reuptake inhibitor fluoxetine has been shown to reduce the frequency and duration of vasospastic episodes. Intravenous infusions of prostaglandins have been reserved for nonresponders with severe symptoms.

Surgical therapy is limited to debridement of digital ulcerations and amputation of gangrenous digits, which are rare complications. Upper extremity sympathectomy may provide relief in 60% to 70% of patients; however, the results are short-lived with a gradual recurrence of symptoms in 60% of patients within 10 years.

### Fibromuscular Dysplasia

FMD is a vasculopathy of uncertain etiology that is characterized by segmental arterial involvement. Histologically, fibrous tissue proliferation, smooth muscle cell hyperplasia, and elastic fiber destruction alternate with mural thinning. The characteristic beaded appearance of FMD is seen due to areas of medial thinning alternating with areas of stenosis. The most commonly affected
are medium-sized arteries, including the internal carotid, renal, vertebral, subclavian, mesenteric, and iliac arteries. The internal carotid artery is the second most common site of involvement after the renal arteries. FMD occurs most frequently in women (90%) and is recognized at approximately 55 years of age. Only 10% of patients with FMD will have complications attributable to the disease. Pathologically, FMD is a heterogeneous group of four distinct types of lesions that are subgrouped based on the predominant site of involvement within the vessel wall. Of the four types (medial fibroplasia, intimal fibroplasia, medial hyperplasia, and perimedial dysplasia), medial fibroplasia is the most common pathologic type, affecting the internal carotid artery (ICA) and the renal artery, and occurring in 85% of reported cases.

The two main clinical syndromes associated with FMD are TIAs from disease in the internal carotid artery and hypertension from renal artery involvement. Symptoms produced by FMD are generally secondary to associated arterial stenosis and are clinically indistinguishable from those caused by atherosclerotic disease. Often, asymptomatic disease is found incidentally on conventional angiographic studies being performed for other reasons. Within the internal carotid artery, FMD lesions tend to be located higher in the extracranial segment than with atherosclerotic lesions and may not be readily demonstrated by duplex scan.

Clinically, symptoms are due to encroachment on the vessel lumen and a reduction in flow. Additionally, thrombi may form in areas of mural dilatation from a stagnation of flow, leading to distal embolization. Surgical treatment has been favored for symptomatic patients with angiographically proven disease. Due to the distal location of FMD lesions in the extracranial carotid artery, resection and repair are not usually feasible. Instead, graduated luminal dilatation under direct vision has been used successfully in patients, with antiplatelet therapy continued postoperatively. PTA has been used effectively in patients with FMD-induced hypertension. Several series have documented a high technical success rate, with recurrence rates of 8% to 23% at more than 1 year. However, the therapeutic effect of blood pressure control may continue to be observed despite restenosis. Surgical reconstruction of the renal arteries for FMD has good long-term results and is recommended for recurrent lesions after angioplasty. Open balloon angioplasty of the ICA has been described, which allows for precise fluoroscopic guidance, rather than blind dilatation with calibrated metal probes, and back-bleeding after dilatation to eliminate cerebral embolization. Distal neuroprotective devices may allow this procedure to be performed completely percutaneously, thereby lessening the threat of cerebral emboli.

Nonatherosclerotic Disease Affecting the Popliteal Artery Disease

There are three distinct nonatherosclerotic disease entities that may result in lower extremity claudication that predominantly occur in 40- to 50-year-old men. Adventitial cystic disease, popliteal artery entrapment syndrome, and Buerger’s disease should be considered in any young patients presenting with intermittent claudication.

Adventitial Cystic Disease of the Popliteal Artery. The first successful operative repair of popliteal artery occlusion caused by a cyst arising from the adventitia was reported in 1954 by Ejrup and Hierton. Adventitial cystic disease is a rare arterial condition occurring at an incidence of 0.1%, usually in the popliteal artery. This disease affects men in a ratio of approximately 5:1 and appears predominantly in the fourth and fifth decades. The incidence is approximately 1 in 1200 cases of claudication or 1 in 1000 peripheral angiograms. The predominance of reported cases is found in Japan and Europe. However, this disease may affect other vascular sites, such as the femoral, external iliac, radial, ulnar, and brachial arteries. Besides claudication as a symptom, this diagnosis should be considered in young patients who have a mass in a nonaxial vessel in proximity to a related joint. These synovial-like, mucin-filled cysts reside in the subadventitial layer of the vessel wall and have a similar macroscopic appearance to ganglion cysts. Despite this similarity and suggestion of a joint origin for these lesions, histochemical markers have failed to link the cystic lining to synovium.

Patients presenting at a young age with bilateral lower extremity claudication and minimal risk factors for atheroma formation should be evaluated for adventitial cystic disease, as well as the other two nonatherosclerotic vascular lesions described here. Because of luminal encroachment and compression, peripheral pulses may be present in the limb when extended, but then can disappear during knee joint flexion. Noninvasive studies may suggest arterial stenosis with elevated velocities. Color-flow duplex scanning followed by T2-weighted MRI now appears to be the best diagnostic choice. Angiography will demonstrate a smooth, well-defined, crescent-shaped filling defect, the classic “scimitar” sign. There may be associated calcification in the cyst wall and no other evidence of atherosclerotic occlusive disease.

Various therapeutic methods have been described for the treatment of adventitial cystic disease. The recommended treatments are excision of the cyst with the cystic wall, enucleation, or simple aspiration when the artery is stenotic. Retention of the cystic lining leads to continued secretion of the cystic fluid and recurrent lesions. In 30% of patients who have an occluded artery, resection of the affected artery, followed by an interposition graft using autogenous saphenous vein, is recommended.

Popliteal Artery Entrapment Syndrome. Love and colleagues first coined the term popliteal artery entrapment in 1965 to describe a syndrome combining muscular involvement with arterial ischemia occurring behind the knee, with the successful surgical repair having taken place 6 years earlier. This is a rare disorder with an estimated prevalence of 0.16% that occurs with a male-to-female ratio of 15:1. Five types of anatomic entrapment have been defined, according to the position of the medial head of the gastrocnemius muscle, abnormal muscle slips or tendinous bands, or the course of the popliteal artery itself (Table 23-29). Concomitant popliteal vein impingement occurs in up to 30% of cases. Twenty-five percent of cases are bilateral.

The typical patient presents with swelling and claudication of isolated calf muscle groups following vigorous physical activity. Various differential diagnoses must be considered when encountering patients with symptoms and signs suggestive of popliteal artery entrapment syndrome (Table 23-30). In a large series of 240 patients, the median age for surgical treatment was 28.5 years. Noninvasive studies with ABIs should be performed with the knee extended and the foot in a neutral, forced plantar, and dorsiflexed position. A drop in pressure of 50% or greater or dampening of the plethysmographic waveforms in plantar or dorsiflexion is a classic finding. Contraction
Table 23-29
Classification of popliteal entrapment syndrome

<table>
<thead>
<tr>
<th>TYPE</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Popliteal artery is displaced medially around a normal medial head of the gastrocnemius</td>
</tr>
<tr>
<td>II</td>
<td>Medial head of gastrocnemius, which arises lateral to popliteal artery</td>
</tr>
<tr>
<td>III</td>
<td>Popliteal artery is compressed by an accessory slip of muscle from medial head of gastrocnemius</td>
</tr>
<tr>
<td>IV</td>
<td>Entrapment by a deeper popliteal muscle</td>
</tr>
<tr>
<td>V</td>
<td>Any of the above plus popliteal vein entrapment</td>
</tr>
<tr>
<td>VI</td>
<td>Functional entrapment</td>
</tr>
</tbody>
</table>

of the gastrocnemius should compress the entrapped popliteal artery. The sudden onset of signs and symptoms of acute ischemia with absent distal pulses is consistent with popliteal artery occlusion secondary to entrapment. Other conditions resulting from entrapment are thrombus formation with distal emboli or popliteal aneurysmal degeneration. Although CT and MRI have been employed, angiography remains the most widely used test. Angiography performed with the foot in a neutral position may demonstrate classical medial deviation of the popliteal artery or normal anatomic positioning. Coexisting abnormalities may include stenosis, luminal irregularity, delayed flow, aneurysm, or complete occlusion. Diagnostic accuracy is increased with the use of ankle stress view-active plantar flexion and passive dorsiflexion.

The treatment of popliteal artery entrapment consists of surgical decompression of the impinged artery with possible arterial reconstruction. Division of the anomalous musculotendinous insertion site with or without saphenous vein interposition grafting to bypass the damaged arterial segment has been described to be the procedure of choice. The natural history of entrapment is progressive arterial degeneration leading to complete arterial thrombosis. In such instances, thrombolytic therapy is needed with subsequent release of the functional arterial impairment. Lysis will improve distal runoff and may improve limb-salvage and bypass patency rates.

Buerger’s Disease (Thromboangiitis Obliterans)

Buerger’s disease, also known as thromboangiitis obliterans, is a progressive nonatherosclerotic segmental inflammatory disease that most often affects small- and medium-sized arteries, veins, and nerves of the upper and lower extremities. The clinical and pathologic findings of this disease entity were published in 1908 by Leo Buerger in a description of 11 amputated limbs. The typical age range for occurrence is 20 to 50 years, and the disorder is more frequently found in males who smoke. The upper extremities may be involved, and a migratory superficial phlebitis may be present in up to 16% of patients, thus indicating a systemic inflammatory response. In young adults presenting to the Mayo Clinic (1953–1981) with lower limb ischemia, Buerger’s disease was diagnosed in 24%. Conversely, the diagnosis was made in 9% of patients with ischemic finger ulcerations. The cause of thromboangiitis obliterans is unknown; however, use or exposure to tobacco is essential to both the diagnosis and progression of the disease.

Pathologically, thrombosis occurs in small- to medium-sized arteries and veins with associated dense polymorphonuclear leukocyte aggregation, microabscesses, and multinucleated giant cells. The chronic phase of the disease shows a decrease in the hypercellularity and frequent recanalization of the vessel lumen. End-stage lesions demonstrate organized thrombus and blood vessel fibrosis. Although the disease is common in Asia, North American males do not appear to have any particular predisposition, as the diagnosis is made in less than 1% of patients with severe limb ischemia.

Buerger’s disease typically presents in young male smokers, with symptoms beginning prior to age 40. Patients initially present with foot, leg, arm, or hand claudication, which may be mistaken for joint or neuromuscular problems. Progression of the disease leads to calf claudication and eventually ischemic rest pain and ulcerations on the toes, feet, or fingers. A complete history should exclude diabetes, hyperlipidemia, or autoimmune disease as possible etiologies for the occlusive lesions. Because it is likely that multiple limbs are involved, angiography should be performed of all four limbs. Even if symptoms are not yet present in a limb, angiographic findings may be demonstrated. Characteristic angiographic findings show disease confinement to the distal circulation, usually infrapopliteal and distal to the brachial artery. The occlusions are segmental and show “skip” lesions with extensive collateralization, the so-called corkscrew collaterals.

The treatment of thromboangiitis obliterans revolves around strict smoking cessation. In patients who are able to abstain, disease remission is impressive, and amputation avoidance is increased. In the experience reported from the Oregon Health Sciences Center, no disease progression with associated tissue loss occurred after discontinuation of tobacco. The role of surgical intervention is minimal in Buerger’s disease, as there is often no acceptable target vessel for bypass. Furthermore, autogenous vein conduits are limited secondary to coexisting migratory thrombophlebitis. Mills and associates reported their results.

Table 23-30
Differential diagnosis for popliteal entrapment syndrome

<table>
<thead>
<tr>
<th>Vascular Etiologies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atherosclerosis</td>
</tr>
<tr>
<td>Buerger’s disease</td>
</tr>
<tr>
<td>Trauma</td>
</tr>
<tr>
<td>Popliteal aneurysm</td>
</tr>
<tr>
<td>Adventitial cystic disease</td>
</tr>
<tr>
<td>Extrinsic compression</td>
</tr>
<tr>
<td>Cardiac embolism</td>
</tr>
<tr>
<td>Deep vein thrombosis</td>
</tr>
<tr>
<td>Venous entrapment</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Musculoskeletal Etiologies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gastrocnemius or soleus strain</td>
</tr>
<tr>
<td>Periostitis</td>
</tr>
<tr>
<td>Compartment syndrome</td>
</tr>
<tr>
<td>Stress fractures</td>
</tr>
<tr>
<td>Tibialis posterior tendonitis</td>
</tr>
<tr>
<td>Muscular anomalies</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>General Neurologic Etiologies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spinal stenosis</td>
</tr>
</tbody>
</table>
of 31% limb loss in 26 patients over 15 years, thus authenticating the virulence of Buenger’s disease involving the lower extremities. In addition, others have described a significant discrepancy in limb loss in patients who continued to smoke versus those who discontinued tobacco use (67% vs. 35%).

REFERENCES

Entries highlighted in bright blue are key references.


154. Bolia A, Sayers RD, Thompson MM, Bell PR. Subintimal and intraluminal recanalisation of occluded crural


Venous Anatomy

Veins are part of a dynamic and complex system that returns low-nutrient deoxygenated blood to the heart. Venous blood flow is dependent on multiple factors such as gravity, venous valves, the cardiac and respiratory cycles, blood volume, and the calf muscle and feet pumps. Alterations in the intricate balance of these factors can result in venous pathology.

Structure of Veins

Veins are thin-walled, highly distensible, and collapsible. Their structure specifically supports the primary functions of veins to transport blood toward the heart and serve as a reservoir to prevent intravascular volume overload.

The venous intima is composed of a nonthrombogenic endothelium with an underlying basement membrane and an elastic lamina. The endothelium produces endothelium-derived relaxing factors such as nitric oxide and prostacyclin, which help maintain a nonthrombogenic surface through inhibition of platelet aggregation and promotion of platelet disaggregation. The capacitance function of veins is facilitated by circumferential rings of elastic tissue, and smooth muscle located in the media of the vein allows for changes in vein caliber with minimal changes in venous pressure. The adventitia is most prominent in large veins and consists of collagen, elastic fibers, and fibroblasts. When a vein is maximally distended, its diameter may be several times greater than that in the supine position.

In the axial veins, unidirectional blood flow is achieved with multiple venous valves. The inferior vena cava (IVC), common iliac veins, portal venous system, and cranial sinuses are valveless. In the axial veins, valves are more numerous distally in the extremities than proximally. Each valve consists of two thin cusps of a fine connective tissue skeleton covered by endothelium. Venous valves close in response to cephalad-to-caudal blood flow at a velocity of at least 30 cm/s.

Lower Extremity Veins

Lower extremity veins are divided into superficial, deep, and perforating veins. The superficial venous system lies above the uppermost fascial layer of the leg and thigh and consists of the great saphenous vein (GSV) and small saphenous vein (SSV) and their tributaries. The GSV originates from the dorsal pedal venous arch and courses cephalad and medially, anterior to the medial malleolus, entering the common femoral vein approximately 4 cm inferior and lateral to the pubic tubercle. The saphenous nerve accompanies the GSV medially from the ankle to the level of the knee and supplies cutaneous sensation to the medial leg and ankle. The SSV originates laterally from the dorsal pedal venous arch and courses cephalad in the posterior calf. Most often, it penetrates the popliteal fossa, between the medial and lateral heads of the gastrocnemius muscle, to join the popliteal vein. The termination of the SSV may be quite variable, however, with a proximal extension of the SSV (the vein of Giacomini) connecting with the deep femoral vein or GSV. The sural nerve accompanies the SSV laterally along its course and supplies cutaneous sensation to the lateral malleolar region.

The deep veins follow the course of major arteries in the extremities. In the lower leg, paired veins parallel the course of the anterior tibial, posterior tibial, and peroneal arteries, to join behind the knee forming the popliteal vein. Venous bridges connect the paired axial tibial veins in the lower leg. The popliteal vein continues through the adductor hiatus to become the femoral vein. In the proximal thigh, the femoral vein joins with the deep femoral vein to form the common femoral vein, becoming the external iliac vein at the inguinal ligament.

Multiple perforator veins traverse the deep fascia to connect the superficial and deep venous systems. Potentially clinically important perforator veins are the posterior tibial and paratibial perforators (formerly known as the Cockett and Boyd perforators, respectively). The posterior tibial perforator veins drain the medial lower leg and are relatively constant. They
Key Points

1. Thrombolytic therapy, surgical thrombectomy, and placement of inferior vena cava filters are adjunctive treatments that may be indicated in patients with extensive and complicated venous thromboembolism.

2. Deep vein thrombosis (DVT) and pulmonary embolism are well-recognized complications after major abdominal and orthopedic procedures. The risk is further increased in patients with malignancy and a history of venous thromboembolism. Options for DVT prophylaxis include intermittent pneumatic compression, use of graduated compression stockings, and administration of low-dose unfractionated heparin, low molecular weight heparin, fondaparinux, and vitamin K antagonists. Direct thrombin inhibitors and factor Xa inhibitors are approved for prophylactic use only for orthopedic procedures and for recurrent VTE. However, prophylaxis should be stratified based on the patient’s level of risk.

3. In patients with established DVT, unfractionated heparin, low molecular weight heparin, fondaparinux, and some factor Xa inhibitors are options for initial antithrombotic therapy. Vitamin-K antagonists, direct thrombin inhibitors, and factor Xa inhibitors are utilized for long-term anticoagulation.

connect the posterior accessory GSV (formerly known as the posterior arch vein, a tributary to the GSV) and the posterior tibial vein. They may become varicose or incompetent in venous insufficiency states. The posterior accessory GSV has relevance as it represents a connection of the three ankle perforating veins, which are likely of particular importance in the development of a venous stasis ulcers. The paratibial perforator veins connect the GSV to the deep veins approximately 10 cm below the knee and 1 to 2 cm medial to the tibia. Additional perforators in the thigh are known as the perforators of the femoral canal (also known as Hunter’s and Dodd’s perforators).

Venous sinuses are thin-walled, large veins located within the substance of the soleus and gastrocnemius muscles. These sinuses are valveless and are linked by valves, small venous channels that prevent reflux. A large amount of blood can be stored in the venous sinuses before draining into the posterior tibial and peroneal veins. With each contraction of the calf muscle bed, blood is pumped out through the venous channels into the main conduit veins to return to the heart.

Upper Extremity Veins

As in the lower extremity, there are deep and superficial veins in the upper extremity. Deep digital veins form the palmar venous arches of the hand and empty into the paired radial and ulnar veins. These follow the named arteries in the arm and are known as the *venae comitantes*. They become the brachial veins most often near the antecubital fossa and then combine to contribute to forming the axillary vein. Superficial veins of the upper extremity are the cephalic and basilic veins and their tributaries. The cephalic vein originates at the lateral wrist and courses over the lateral ventral surface of the forearm. In the upper arm, the cephalic vein terminates in the infraclavicular fossa, piercing the clavipectoral fascia to empty into the axillary vein. The basilic vein runs medially along the forearm and penetrates the deep fascia as it courses past the elbow in the upper arm. It then joins with the deep brachial veins to become the axillary vein, a landmark for identification of the axillary vein. The median antecubital vein joins the cephalic and the basilic veins on the ventral surface of the elbow.

The axillary vein becomes the subclavian vein at the lateral border of the first rib. At the medial border of the scalenus anterior muscle, the subclavian vein joins with the internal jugular vein to become the brachiocephalic vein, with the subclavian vein coursing anterior to the scalenus anterior muscle. The left and right brachiocephalic veins join to become the superior vena cava, which empties into the right atrium.

The mainstay of treatment for chronic venous insufficiency is compression therapy. Sclerotherapy, perforator vein ligation, and venous reconstruction or ablative techniques may be indicated in patients in whom conservative management fails or as a means to decrease ulcer recurrence.

Lymphedema is categorized as congenital, primary (with early or delayed onset), or secondary. The goals of treatment are to minimize edema and prevent infection. Lymphatic massage, sequential pneumatic compression, use of compression garments, and limb elevation are effective forms of therapy.

The duration and type of long-term anticoagulation should be stratified based on the provoked or unprovoked nature of the DVT, the location of the DVT, previous occurrence of DVT, and presence of concomitant malignancy.

High ligation and stripping, endovenous laser, or radiofrequency ablation and sclerotherapy are effective therapies for patients with saphenous vein valvular insufficiency. Concomitant varicose veins may be managed with compression therapy, sclerotherapy, and phlebectomy. Nonthermal ablative techniques, including the combination of sclerotherapy with endoluminal mechanical injury as well as injection of cyanoacrylate, show early promising results.

EVALUATION OF THE VENOUS SYSTEM

Clinical Evaluation

Evaluation of the venous system begins with a detailed history and physical examination. Risk factors for acute and chronic venous disease are identified. They include increased age, history of venous thromboembolism (VTE), malignancy, trauma and spinal cord injury, hospitalization and immobilization, obesity, nephrotic syndrome, pregnancy, recent postpartum state, oral contraceptive use or hormone replacement therapy, varicose veins, and hypercoagulable states, as well as the postoperative state. Venous pathology is often, but not always, associated with visible or palpable signs that can be identified during the physical examination. There is variation among individuals in the prominence of superficial veins when the person is standing (Fig. 24-1). The superficial veins of a lean athletic person, even when normal, will appear large and easily visualized, but these veins will be far less obvious in the obese individual. Signs of superficial venous abnormalities are listed in Table 24-1.
The deep veins cannot be directly assessed clinically, and abnormalities within them can only be inferred indirectly from changes found on clinical examination.

Chronic venous insufficiency (CVI) may lead to characteristic changes in the skin and subcutaneous tissues in the affected limb. CVI results from incompetence of venous valves, venous obstruction, or both. Most CVI involves venous reflux, and severe CVI often reflects a combination of reflux and venous obstruction. It is important to remember that although CVI originates with abnormalities of the veins, the target organ of CVI is the skin, and the underlying physiologic and biochemical mechanisms leading to the cutaneous abnormalities associated with CVI are poorly understood. A typical leg affected by CVI will be edematous, with edema increasing over the course of the day. The leg may also be indurated and pigmented with eczema and dermatitis. These changes are associated with excessive proteinaceous capillary exudate. Deposition of a pericapillary fibrin cuff may limit nutritional exchange. In addition, an increase in white blood cell trapping within the skin microcirculation in CVI patients may lead to microvascular congestion and thrombosis. Subsequently, white blood cells may migrate into the interstitium and release necrotizing lysosomal enzymes, potentially leading to tissue destruction and eventual ulceration.

Fibrosis can eventually develop from impaired nutrition, chronic inflammation, and fat necrosis (lipodermatosclerosis). Hemosiderin deposition due to the extravasation of red cells and subsequent lysis in the skin contributes to the characteristic pigmentation of chronic venous disease (Fig. 24-2). Ulceration can develop with longstanding venous hypertension and is associated with alterations in microcirculatory and cutaneous lymphatic anatomy and function. The most common location of venous ulceration is approximately 3 cm proximal to the medial malleolus, frequently referred to the “gaiter” region (Fig. 24-3).

Trendelenburg’s test is a clinical test, historically important but now rarely used, that can help determine whether incompetent valves are present and in which of the three venous systems (superficial, deep, or perforator) the valves are abnormal. There are two components to this test. First, with the patient supine, the leg is elevated 45° to empty the veins, and the GSV is occluded with the examiner’s hand or with a rubber tourniquet. Then, with the GSV still occluded, the patient stands, and the superficial veins are observed for blood filling. The compression on the GSV is released, and the superficial veins are observed for filling with blood. A positive result is the sudden filling of veins with standing while the GSV remains occluded, indicating incompetent perforator and deep veins. Additionally, the GSV valves are incompetent if rapid filling is noted following release of compression. A negative result, indicating no clinically relevant venous reflux, is the gradual filling of the veins from arterial inflow. Interpretation of the findings of Trendelenburg’s test is subjective, and therefore, it has largely been supplanted by the more objective noninvasive vascular laboratory tests to localize sites of venous reflux.

**Noninvasive Evaluation.** Before the development of vascular ultrasound, noninvasive techniques to evaluate the venous system were based on plethysmographic techniques. Although a variety of plethysmographic techniques are used in the
Complications of venography include pain, thrombosis, or hematoma at the puncture site. Pain is lower with nonionic low-osmolal contrast media than with conventional contrast agents (with 18% vs. 44% of patients experiencing discomfort, respectively).³ Systemic effects of iodinated contrast media include allergic reaction and risk of renal failure. Postvenography venous thrombosis occurs distal to the venous puncture site in 1% to 9% of patients undergoing venography secondary to intimal damage from the intravenous (IV) contrast agent.⁴ Complications and limitations of IVUS are related to complications at the access site and cost of the catheters.

Figure 24-3. Venous ulceration located proximal to the medial malleolus.

evaluation of both acute and chronic venous disease, they are all based on the detection of volume changes in the limb in response to blood flow.

Duplex ultrasonography (DUS) augmented by color flow imaging is now the most important noninvasive diagnostic method in the evaluation of the venous system. DUS has become standard for the detection of infrainguinal deep vein thrombosis (DVT), with near 100% sensitivity and specificity in symptomatic patients.³ It is also the preferred method of evaluation for upper extremity venous thrombosis and is useful in the evaluation of CVI by documenting the presence of valvular reflux and venous obstruction. Overlying bowel gas and large body habitus many times make DUS less applicable to evaluation of intra-abdominal veins. Magnetic resonance venography (MRV) and computed tomography (CT) venography are alternative noninvasive techniques for evaluation of pelvic and intra-abdominal veins.

**Invasive Evaluation.** Improved accuracy of noninvasive techniques for diagnostic purposes has made the use of invasive procedures more selective. Both venography and intravascular ultrasonic (IVUS) are used as adjuncts to percutaneous or open surgical treatment of venous disorders. When planning endovascular or open surgical treatment, venography may be used to identify areas of obstruction in infrainguinal, intra-abdominal, and upper extremity veins as well as reflux in intra-abdominal and infrainguinal veins. IVUS, with access generally via the common femoral vein, is used primarily to assess for occlusive lesions of the iliac veins and appears more sensitive than venography in detecting iliac vein obstruction.⁴

**VENOUS THROMBOEMBOLISM**

**Epidemiology**

Despite increased awareness and use of prophylactic modalities, DVT or pulmonary embolism (PE), venous thromboembolism (VTE), remain important preventable sources of morbidity and mortality, especially in the surgical patient. The incidence of VTE is approximately 100 per 100,000 people per year in the general population, with 20% of the diagnoses made within 3 months of a surgical procedure. Of the symptomatic patients, one-third will present with PE and two-thirds with DVT.⁶ ⁷ The estimated number of cases of VTE may well be over 600,000 per year in the United States, making it a major U.S. health problem.⁸ Furthermore, death occurs in 6% of DVT and 12% of PE cases within 1 month of diagnosis, although not all deaths are directly secondary to VTE, with many related to the underlying problem leading to the VTE event.⁶ However, not only does VTE pose a veritable threat to life, it also places patients at higher risk for recurrence and post-VTE sequelae such as pulmonary hypertension and postthrombotic syndrome, with 4% and up to 30% incidence, respectively.⁹ ¹¹

**Risk Factors**

Three broadly stated conditions, first described by Rudolf Virchow in 1862, contribute to VTE formation: stasis of blood flow, endothelial damage, and hypercoagulability. Of these risk factors, relative hypercoagulability appears most important in cases of **spontaneous** VTE, or so-called idiopathic VTE, whereas stasis and endothelial damage likely play a greater role in **secondary** VTE, or so-called provoked VTE, occurring in association with transient risk factors such as immobilization, surgical procedures, and trauma. Identifiable risk factors for VTE generally relate to one of the conditions described by Virchow. Often more than one risk factor is present contributing in an exponential, rather than additive, manner. Specific risk factors for VTE are listed in Table 24-2.

The more common acquired VTE risk factors include older age (>40 years), hospitalization and immobilization, hormone replacement and oral contraceptive therapy, pregnancy and the recently postpartum state, prior VTE, malignancy, major surgery, obesity, nephrotic syndrome, trauma and spinal cord injury, long-haul travel (>6 hours), varicose veins, antiphospholipid syndrome, myeloproliferative disorders, and polycythemia. Heritable risk factors include male sex, factor V Leiden mutation; prothrombin 20210A gene variant; antithrombin, protein C, and protein S deficiencies; and dysfibrinogenemias. In some patients, the cause of the thrombophilia may have both a heritable and an acquired component. These mixed causes include homocysteinemia; factors VII, VIII, IX, and XI elevation; hyperfibrinogenemia; and activated protein C resistance in the
absence of factor V Leiden.\textsuperscript{12} There may be a synergistic effect when particular multiple inherited and acquired risk factors are present in the same patient.

Other patient-specific factors associated with venous thrombosis include the traditional cardiovascular risk factors of obesity, hypertension, and diabetes. VTE is more common in whites and African Americans than Asians and Native Americans.\textsuperscript{13,14} Certain gene variants (single nucleotide polymorphisms) are also associated with a mildly increased risk for VTE, and their presence may interact with other risk factors to increase the overall risk for venous thrombosis.\textsuperscript{15}

Anatomic factors may also contribute to development of DVT. At the site where the right iliac artery crosses over the left iliac vein, the left iliac vein may become chronically compressed predisposing to iliofemoral venous thrombosis, so-called May-Thurner syndrome. External compression of major veins by masses of various types can also lead to venous thrombosis.

Many cases of VTE are potentially preventable. Accordingly, in current clinical practice, preoperative VTE risk assessment is becoming increasingly common to identify patients at moderate and high risk. Scoring systems have been developed that take into account the number of VTE risk factors in an individual patient. These risk stratification scores, such as the Rogers score\textsuperscript{16} and Caprini score,\textsuperscript{17} provide individual patient risk stratification and recommendations for prophylactic anticoagulation. The ninth edition of the American College of Chest Physicians (ACCP) Guidelines for Prevention of VTE in Non-Orthopedic Surgical Patients acknowledges both the Rogers and Caprini scores and provides recommendations for VTE prophylaxis (Table 24-3). Orthopedic surgical patients are generally excluded from risk assessment scores because of the disproportionately increased risk of VTE in orthopedic surgery compared with the general and abdominopelvic surgery population.

### Table 24-3

<table>
<thead>
<tr>
<th>LEVEL OF RISK</th>
<th>APPROXIMATE DVT RISK WITHOUT THROMBOPROPHYLAXIS (%)</th>
<th>SUGGESTED THROMBOPROPHYLAXIS OPTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very low risk</td>
<td>(&lt;0.5% ) (Rogers score (&lt;7); Caprini score (0))</td>
<td>No specific thromboprophylaxis, Early ambulation</td>
</tr>
<tr>
<td>Low risk</td>
<td>(~1.5% ) (Rogers score 7–10; Caprini score 1–2)</td>
<td>Mechanical prophylaxis</td>
</tr>
<tr>
<td>Moderate risk</td>
<td>(~3.0% ) (Rogers score (&gt;10); Caprini score 3–4)</td>
<td>LMWH (at recommended doses), LDUH, or mechanical prophylaxis</td>
</tr>
<tr>
<td>High bleeding risk</td>
<td></td>
<td>Mechanical prophylaxis</td>
</tr>
<tr>
<td>High risk</td>
<td>(~6% ) (Caprini score (\geq5))</td>
<td>LMWH (at recommended doses), fondaparinux and mechanical prophylaxis</td>
</tr>
<tr>
<td>High bleeding risk for cancer</td>
<td></td>
<td>Mechanical thromboprophylaxis, Extended-duration LMWH (4 weeks)</td>
</tr>
</tbody>
</table>

DVT = deep vein thrombosis; INR = international normalized ratio; LDUH = low-dose unfractionated heparin; LMWH = low molecular weight heparin; VTE = venous thromboembolism.

Diagnosis

Clinical Evaluation. Early in the course of DVT development, venous thrombosis is thought to begin in an area of relative stasis, such as a soleal sinus vein or immediately downstream of the cusps of a venous valve in an axial calf vein. Isolated proximal DVT without tibial vein thrombosis is unusual. Early in the course of a DVT, there may be no or few clinical findings such as pain or swelling. Even extensive DVT may sometimes be present without signs or symptoms if the patient is nonambulatory or bedbound. History and physical examination are notoriously unreliable in the diagnosis of DVT. In addition, symptoms and signs often associated with DVT, such as extremity pain and/or swelling, are nonspecific. In large studies, DVT has been found by venography or DUS in ≤50% of patients in whom it was clinically suspected.18,19 Objective studies are therefore required to confirm a diagnosis of VTE or to exclude the presence of VTE.

Clinical symptoms may worsen as DVT propagates and involves the major proximal deep veins. Extensive DVT of the major axial deep venous channels of the lower extremity with relative sparing of collateral veins causes a condition called phlegmasia cerulea dolens (Fig. 24-4). This condition is characterized by pain and pitting edema with associated cyanosis. When the thrombosis extends to the collateral veins, massive fluid sequestration and more significant edema ensue, resulting in a condition known as phlegmasia alba dolens.20 The affected extremity in phlegmasia alba dolens is extremely painful and edematous and pale secondary to arterial insufficiency from dramatically elevated below lower knee compartment pressures. Both phlegmasia ceruleolens and phlegmasia alba dolens can be complicated by venous gangrene and the need for amputation.

Vascular Lab and Radiologic Evaluation

Duplex Ultrasound DUS is now the most commonly performed test for the detection of infrainguinal DVT, both above and below the knee, and has a sensitivity and specificity of >95% in symptomatic patients.3 DUS refers to the combination of real-time B-mode ultrasound with compression and flow augmentation maneuvers combined with pulsed Doppler capability. For VTE detection, color flow imaging is an additional extremely useful adjunct in the evaluation of possible axial calf vein DVT and evaluation of intra-abdominal veins. DUS provides the ability to noninvasively visualize venous anatomy, detect occluded and partially occluded venous segments, and demonstrate physiologic flow characteristics.

In the supine patient, normal lower extremity venous flow is phasic (Fig. 24-5), decreasing with inspiration in response to increased intra-abdominal pressure with the descent of the diaphragm and then increasing with expiration as the diaphragm rises and intra-abdominal pressure decreases. When the patient is upright, the decrease in intra-abdominal pressure with expiration cannot overcome the hydrostatic column of pressure existing between the right atrium and the calf. Muscular contractions of the calf, along with the one-way venous valves, are then required to promote venous return to the heart. Flow also can be increased by leg elevation or compression and decreased by sudden elevation of intra-abdominal pressure (Valsalva maneuver).

In a venous DUS examination performed with the patient supine, spontaneous flow, variation of flow with respiration, and response of flow to Valsalva maneuver are all assessed. From the common femoral through the popliteal vein, the primary method of detecting DVT with ultrasound is demonstration of the lack of compressibility of the vein with probe pressure on B-mode imaging. Normally, in transverse section, the vein walls should coapt with pressure. Lack of coaptation indicates thrombus. Axial calf vein thrombi are often best detected by abnormalities in color flow imaging as compressibility is difficult in the calf.

The examination begins at the ankle and continues proximally to the groin. Each vein is visualized, and the flow signal is assessed with distal and proximal compression. Lower extremity DVT can be diagnosed by any of the following DUS findings: lack of spontaneous flow (Fig. 24-6), inability to compress the vein (Fig. 24-7), absence of color filling of the lumen by color flow DUS, loss of respiratory flow variation, and venous distention. Again, lack of venous compression on B-mode imaging is the primary diagnostic variable. Several studies comparing B-mode ultrasound to venography for the detection of femoropopliteal DVT in patients clinically suspected to have DVT report sensitivities of >91% and specificities of >97%.21,22 The ability of DUS to assess isolated calf vein DVT varies greatly, with sensitivities ranging from 50% to 93% and specificities approaching 100%.23,24

Impedance Plethysmography Impedance plethysmography (IPG) was the primary noninvasive method of diagnosing DVT before the widespread use of DUS but is infrequently used today. Changes in electrical resistance resulting from lower extremity blood volume changes are quantified. IPG is less accurate than DUS for the detection of proximal DVT, with 83% sensitivity in symptomatic patients. It is a poor detector of calf vein DVT.25
**Iodine-125 Fibrinogen Uptake**  Iodine-125 fibrinogen uptake (FUT) is a seldom-used technique that involves IV administration of radioactive fibrinogen and monitoring for increased uptake in fibrin clots. An increase of 20% or more in one area of a limb indicates an area of thrombus. FUT can detect DVT in the calf, but high background radiation from the pelvis and the urinary tract limits its ability to detect proximal DVT. It also cannot be used in an extremity that has recently undergone surgery or has active inflammation. In a prospective study, FUT had a sensitivity of 73% and specificity of 71% for identification of DVT in a group of symptomatic and asymptomatic patients. Currently, FUT is primarily a research tool of historic interest.

**Venography**  Venography is the gold standard to which other diagnostic modalities are compared. A small catheter is placed in a dorsal foot vein with injection of a radiopaque contrast agent. Radiographs are obtained in at least two projections. A positive study result is failure to fill the deep system with passage of the contrast medium into the superficial system or demonstration of discrete filling defects (Fig. 24-8). A normal study result virtually excludes the presence of DVT. In a study of 160 patients with a normal venogram followed for 3 months, only two patients (1.3%) subsequently developed DVT, and no patients experienced symptoms of PE. Venography is not routinely used in clinical practice because of its invasiveness and complication risk. It is still, however, sometimes used in research studies evaluating DVT prophylaxis.

**Treatment**  Once the diagnosis of VTE has been made, antithrombotic therapy should be initiated promptly. If clinical suspicion for VTE is high, it may be prudent to start treatment before the diagnosis is objectively confirmed. The goals of VTE treatment are the prevention of mortality and morbidity associated with PE and the prevention of the postthrombotic syndrome (PTS). Treatment regimens may include antithrombotic therapy, temporary or permanent vena cava filter placement, catheter-directed or systemic thrombolytic therapy, and operative thrombectomy.

**Antithrombotic Therapy.**  Most often, antithrombotic therapy for VTE is initiated with IV or subcutaneous (SC) unfractionated heparin or SC low molecular weight heparin. Fondaparinux, a synthetic pentasaccharide, is sometimes also used as an alternative to heparin to initiate therapy. An oral vitamin K antagonist, usually sodium warfarin, is begun shortly after initiation of IV or SC therapy. Either SC or IV therapy is continued until effective oral anticoagulation with warfarin is achieved as indicated by an international normalized ratio (INR) ≥2 for 24 hours. A minimum of 5 days of heparin or fondaparinux therapy is
Recommended. Recently, several oral anticoagulants that function by either directly inhibiting thrombin or inhibiting factor Xa have additionally been approved by the United States Food and Drug Administration (FDA) for both treatment and prophylaxis for VTE. A principle advantage is they do not require monitoring of laboratory parameters for use.

Unfractionated heparin (UFH) binds to antithrombin via a specific 18-saccharide sequence. This increases antithrombin activity over a thousandfold. The antithrombin-heparin complex primarily inhibits factor IIa (thrombin) and factor Xa and, to a lesser degree, factors IXa, XIA, and XIa of the coagulation cascade. In addition, UFH also binds to tissue factor pathway inhibitor, which inhibits the conversion of factors X to Xa, and factors IX to IXa. Finally, UFH catalyzes the inhibition of thrombin by heparin cofactor II via a mechanism independent of antithrombin.

UFH therapy is most commonly administered with an initial IV bolus of 80 units/kg. Weight-based UFH dosages have been shown to be more effective than standard fixed boluses in rapidly achieving therapeutic levels.28 The initial bolus is followed by a continuous IV drip at 18 units/kg per hour. The half-life of IV UFH ranges from 45 to 90 minutes and is dose dependent. The level of antithrombotic therapy should be monitored every 6 hours using the activated partial thromboplastin time (aPTT), with the goal range of 1.5 to 2.5 times control values. This should correspond with plasma heparin anti-Xa activity levels of 0.3 to 0.7 IU/mL.

Initial anticoagulation with UFH may also be administered SC, although this route is less commonly used. Adjusted-dose therapeutic SC UFH is initiated with 17,500 units, followed by 250 units/kg twice daily, and dosing is adjusted to an aPTT goal range similar to that for IV UFH. Fixed-dose unmonitored SC UFH is started with a bolus of 333 units/kg, followed by 250 units/kg twice daily.30

Hemorrhage is the primary complication of UFH therapy. The rate of major hemorrhage (fatal, intracranial, retroperitoneal, or requiring transfusion of >2 units of packed red blood cells) is approximately 5% in hospitalized patients undergoing UFH therapy (1% in medical patients and 8% in surgical patients).30 For patients with UFH-related bleeding complications, cessation of UFH is required, and anticoagulation may be reversed with protamine sulfate. Protamine sulfate binds to UFH and forms an inactive salt compound. Each milligram of protamine neutralizes 90 to 115 units of heparin, and the dosage should not exceed 50 mg IV over any 10-minute period. Side effects of protamine sulfate include hypotension, pulmonary edema, and anaphylaxis. Patients with prior exposure to protamine-containing insulin (NPH) and patients with allergy to fish may have an increased risk of hypersensitivity, although no direct relationship has been established. Protamine administration should be performed judiciously and terminated if any side effects occur.

In addition to hemorrhage, heparin also has other, unique, complications. Heparin-induced thrombocytopenia (HIT) results from heparin-associated antiplatelet antibodies (HAAbs) directed against platelet factor 4 complexed with heparin.31 HIT occurs in 1% to 5% of patients treated with heparin.32,33 In patients with repeat heparin exposure (such as vascular surgery patients), the incidence of HAAbas may be as high as 21%.34 HIT occurs most frequently in the second week of therapy and may lead to disastrous venous or arterial thrombotic complications. Therefore, platelet counts should be monitored periodically in patients receiving continuous heparin therapy.

HIT is diagnosed based on previous exposure to heparin, platelet count less than 100,000, and/or platelet count decline of 50% following exposure. All heparin must be stopped and alternative anticoagulation initiated immediately to avoid thrombotic complications, which may approach 50% over the subsequent 30 days in affected individuals.35

Another complication of prolonged high-dose heparin therapy is osteopenia. Heparin-induced osteopenia results from impairment of bone formation and enhancement of bone resorption by heparin.

Low molecular weight heparins (LMWHs) are derived from the depolymerization of porcine UFH. Like UFH, LMWHs bind to antithrombin via a specific pentasaccharide sequence to expose an active site for the neutralization of factor Xa. However, LMWHs have fewer additional saccharide units. This results in less inactivation of thrombin (factor IIa). In comparison to UFH, LMWHs have increased bioavailability (>90% after SC injection), longer half-lives (approximately 4 to 6 hours), and more predictable elimination rates.

Most patients treated with weight-based once- or twice-daily SC LMWH injections do not require laboratory monitoring for anticoagulant effect, a distinct advantage over continuous IV infusions of UFH. Patients who do require monitoring include those with significant renal insufficiency, pediatric patients, obese patients greater than 120 kg, and pregnant patients. Monitoring may be performed using anti-Xa activity assays. The therapeutic anti-Xa goal range depends on the type of LMWH and the frequency of dosing. There are numerous LMWHs available, and the various preparations differ in their anti-Xa and anti-IIa activities. Treatment dosing for one LMWH, therefore, cannot be extrapolated for use with another. The anticoagulant effect of LMWHs may be partially reversed (approximately 60%) with protamine sulfate.

Numerous well-designed trials comparing SC LMWH with IV and SC UFH for the treatment of DVT have been critically evaluated in several meta-analyses and demonstrate a decrease in thrombotic complications, bleeding, and mortality with LMWHs.36-38 LMWHs also are associated with a decreased rate of HAAb formation and HIT (<2%) compared with UFH (at least in prophylactic doses).30 However, patients with established HIT also should not receive LMWHs because there is cross-reactivity between the drugs.39

A major benefit of LMWHs is that it allows outpatient treatment of VTE.40,41 In a randomized study comparing IV UFH and the LMWH nadroparin calcium,40 there was no significant difference in recurrent thromboembolism (8.6% for UFH vs. 6.9% for LMWH) or major bleeding complications (2.0% for UFH vs. 0.5% for LMWH). There was, however, a 67% reduction in mean days in the hospital for the LMWH group.

Fondaparinux is a synthetic pentasaccharide that has been approved by the FDA for the initial treatment of DVT and PE. Its five-poly saccharide sequence binds and activates antithrombin, causing specific inhibition of factor Xa. In two large noninferiority trials, fondaparinux was compared with the LMWH enoxaparin for the initial treatment of DVT and with IV UFH for the initial treatment of PE.42,43 The rates of recurrent VTE ranged from 3.8% to 5%, with rates of major bleeding of 2% to 2.6%, for all treatment arms. The drug is administered SC once daily with a weight-based dosing protocol: 5 mg, 7.5 mg, or 10 mg for patients weighing <50 kg, 50 to 100 kg, or >100 kg, respectively. The half-life of fondaparinux is approximately...
17 hours in patients with normal renal function. There are rare case reports of fondaparinux-induced thrombocytopenia.  

**Direct thrombin inhibitors (DTIs)** include parenteral forms with recombinant hirudin, argatroban, and bivalirudin, as well as an oral agent, dabigatran. These antithrombotic agents bind to thrombin, inhibiting the conversion of fibrinogen to fibrin as well as thrombin-induced platelet activation. These actions are independent of antithrombin. The parental DTIs should be reserved for (a) patients in whom there is a high clinical suspicion or confirmation of HIT, and (b) patients who have a history of HIT or test positive for heparin-associated antibodies whereas dabigatran can be used as an alternative to Warfarin when INR monitoring is difficult or impractical. In patients with established HIT, DTIs should be administered for at least 7 days, or until the platelet count normalizes. Warfarin may then be introduced slowly, overlapping therapy with a DTI for at least 5 days, or dabigatran may be continued instead of Warfarin.  

Bivalirudin is approved primarily for patients with or without HIT who undergo percutaneous coronary intervention and is rarely used outside of that setting. Argatroban is indicated for the prophylaxis and treatment of thrombosis in HIT. It also is approved for patients with, or at risk for, HIT undergoing percutaneous coronary intervention. Antithrombotic prophylaxis and therapy are initiated with a continuous IV infusion of 2 μg/kg per minute, without the need for a bolus. The half-life ranges from 39 to 51 minutes, and the dosage is adjusted to maintain an aPTT of 1.5 to 3 times normal. Large initial boluses and higher rates of continuous infusion are reserved for patients with coronary artery thrombosis and myocardial infarction. In these patients, therapy is monitored using the activated clotting time. Argatroban is metabolized and excreted by the liver; therefore, dosage adjustments are needed in patients with hepatic impairment. There is no reversal agent for argatroban.  

The oral agent, dabigatran, is US FDA-approved since 2014 for the treatment of VTE and prophylaxis for recurrent VTE. Additionally, limited approval was obtained in 2015 for prophylaxis of VTE after hip replacement surgery. It is administered as a prodrug, dabigatran etexilate, that is converted to the active form, dabigatran, in the liver. The half-life ranges from 12 to 17 hours; it is therefore administered once daily for prophylaxis and twice daily for VTE therapy. Absorption is not dietary dependent, and no drug level monitoring is required. Dabigatran is metabolized in the kidney, and dose adjustment is required for renal insufficiency. Data on use in obese patients is limited; therefore, use is not recommended for patients with a body mass index ≥40 kg/m² or ≥120 kg. Dyspepsia is a common side effect that may limit use in some patients. Dabigatran may be reversed with idarucizumab in emergent situations. It is contraindicated in patients with mechanical heart valves.  

**Direct factor Xa inhibitors**, which are comprised of the oral agents rivaroxaban, apixaban, and edoxaban, are FDA approved for treatment in VTE and prophylaxis for recurrent VTE. Additionally, rivaroxaban and apixaban are approved by the FDA for VTE prophylaxis following knee and hip replacement surgery. These medications function by inactivating circulating and thrombus-bound factor Xa. They are metabolized in the kidney (25–35%) and in the liver; therefore, use is not recommended in patients with renal insufficiency (creatinine clearance <30 mL/min for rivaroxaban, or <15 mL/min for apixaban and edoxaban) or severe hepatic insufficiency. As with the oral DTI, dabigatran, data on use in obese patients is limited; therefore, use is not recommended in these patients. Additionally, these agents are contraindicated in pregnancy. There are no specific reversal agents available for direct factor Xa inhibitors. For severe cases of hemorrhage, indirect partial reversal may be achieved with use of prothrombin complex concentrate administration.  

Rivaroxaban has a half-life of 7 to 17 hours. Therapy does not require monitoring. Prophylactic dosing is 10 mg once daily, and therapeutic dosing is 15 mg twice daily for 21 days, followed by 20 mg once daily thereafter. Apixiban has a half-life of 5 to 9 hours. Therapy does not require monitoring, and there are no dietary restrictions. If monitoring is desired in situations of bleeding or concern for sub- or supratherapeutic dosing, serum anti-Xa levels can be obtained. Prophylactic dosing is 2.5 mg twice daily, and therapeutic dosing is 10 mg daily for 7 days, followed by 5 mg twice daily thereafter.  

Edoxaban has a half-life of 10 to 14 hours. Therapy does not require monitoring. Typical dosing is 60 mg once daily, and 30 mg once daily if creatinine clearance ranges from 15 to 50 mL/min, or body weight ≤60 kg.  

**Vitamin K antagonists**, which include warfarin and other coumarin derivatives, are the traditional mainstay of long-term antithrombotic therapy in patients with VTE. Warfarin inhibits the γ-carboxylation of vitamin K–dependent procoagulants (factors II, VII, IX, and X) and anticoagulants (proteins C and S), resulting in formation of less functional proteins. Warfarin usually requires several days to achieve full effect because normal circulating coagulation proteins must first undergo their normal degradation. Factors X and II have the longest half-lives, in the range of 36 and 72 hours, respectively. A steady-state concentration of warfarin is usually not reached for 4 to 5 days. Warfarin therapy is monitored by measuring the INR, calculated using the following equation:  

\[
\text{INR} = \frac{\text{patient prothrombin time/laboratory normal prothrombin time}}{\text{ISI}}
\]

where **ISI** is the international sensitivity index. The ISI describes the strength of the thromboplastin that is added to activate the extrinsic coagulation pathway. The therapeutic target INR range is usually 2.0 to 3.0, but the response to warfarin is variable and depends on liver function, diet, age, and concomitant medications. In patients receiving anticoagulation therapy without concomitant thrombolysis or venous thrombectomy, the vitamin K antagonist may be started on the same day as the initial parenteral anticoagulant, usually at doses ranging from 5 to 10 mg. Smaller initial doses may be needed in older and malnourished patients, in those with liver disease or congestive heart failure, and in those who have recently undergone major surgery.  

The recommended duration of warfarin antithrombotic therapy is stratified based on whether the DVT was provoked or unprovoked, whether it was the first or a recurrent episode, where the DVT is located, and whether malignancy or thrombophilia is present. Current ACCP recommendations for duration of warfarin therapy are summarized in Table 24-4. In patients with proximal DVT, several randomized clinical trials have demonstrated that shorter-term antithrombotic therapy (4 to 6 weeks) is associated with a higher rate of VTE recurrence than 3 to 6 months of anticoagulation. In these trials, most of the patients with transient risk factors had a low rate of recurrent VTE, and most recurrences were in patients with continuing risk factors. The ACCP recommendation, therefore, is that 3 months of anticoagulation are sufficient to prevent...
Table 24-4

Summary of American College of Chest Physicians recommendations regarding duration of long-term antithrombotic therapy for deep vein thrombosis (DVT)

<table>
<thead>
<tr>
<th>CLINICAL SUBGROUP</th>
<th>ANTITHROMBOTIC TREATMENT DURATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>First episode DVT/transient risk/surgery</td>
<td>VKA or LMWH for 3 months</td>
</tr>
<tr>
<td>First episode DVT/ unprovoked</td>
<td>VKA or LMWH for 3 months</td>
</tr>
<tr>
<td>Distal DVT/unprovoked</td>
<td>VKA for 3 months</td>
</tr>
<tr>
<td>Second episode DVT/ unprovoked DVT and cancer</td>
<td>VKA for extended therapy</td>
</tr>
<tr>
<td>Distal DVT/unprovoked</td>
<td>LMWH for extended therapy over VKA</td>
</tr>
</tbody>
</table>


recurrent VTE in patients with DVT occurring around the time of a transient risk factor (e.g., hospitalization or orthopedic or major general surgery).

In contrast to patients with thrombosis related to transient risk factors, patients with unprovoked VTE are much more likely to develop recurrence (rates as high as 40% at 10 years). In this latter group of patients, numerous clinical trials have compared 3 to 6 months of anticoagulation therapy with extended-duration warfarin therapy, both at low intensity (INR of 1.5 to 2.0) and at conventional intensity (INR of 2.0 to 3.0). In patients with idiopathic DVT, extended-duration antithrombotic therapy is associated with a relative reduction in the rate of recurrent VTE by 75% to >90%. In addition, conventional-intensity warfarin reduces the risk even further compared with low-intensity warfarin (0.7 events per 100 person-years vs. 1.9 events per 100 person-years) without an increase in bleeding complications.

In patients with VTE in association with a hypercoagulable condition, the optimal duration of anticoagulation therapy is influenced more by the clinical circumstances at the time of the VTE (idiopathic vs. secondary) than by the actual presence or absence of the more common thrombophilic conditions. In patients with VTE related to malignancy, increasing evidence suggests that longer-term therapy with LMWH (up to 6 months) is associated with a lower VTE recurrence than treatment using conventional vitamin K antagonists. The primary complication of warfarin therapy is hemorrhage, and the risk is related to the magnitude of INR prolongation. Depending on the INR and the presence of bleeding, warfarin anticoagulation may be reversed by (a) omitting or decreasing subsequent dosages, (b) administering oral or parenteral vitamin K, or (c) administering fresh frozen plasma, prothrombin complex concentrate, or recombinant factor VIIa.

Warfarin therapy rarely may be associated with the development of skin necrosis and limb gangrene. These conditions occur more commonly in women (4:1), and the most commonly affected areas are the breast, buttoks, and thighs. This complication, which usually occurs in the first days of therapy, is occasionally, but not exclusively, associated with protein C or S deficiency and malignancy. Patients who require continued anticoagulation may restart low-dose warfarin (2 mg) while receiving concomitant therapeutic heparin. The warfarin dosage is then gradually increased over a 1- to 2-week period.

Systemic and Catheter-Directed Thrombolysis. Patients with extensive proximal, iliofemoral DVT may benefit from systemic thrombolysis or catheter-directed thrombolysis (CDT). CDT appears to be more effective (see later in chapter) and potentially reduces acute congestive lower extremity symptoms more rapidly than anticoagulation alone and decreases the development of PTS.

Several thrombolytic agents are available, including streptokinase, urokinase, alteplase (recombinant tissue plasminogen activator), reteplase, and tenecteplase. All share the ability to convert plasminogen to plasmin, which leads to the degradation of fibrin. They differ with regard to their half-lives, their potential for inducing fibrinogenolysis (generalized lytic state), their potential for antigenicity, and their FDA-approved indications for use.

Streptokinase is purified from B-hemolytic Streptococcus and is approved for the treatment of acute myocardial infarction, PE, DVT, arterial thromboembolism, and occluded central lines and arterovenous shunts. It is not specific for fibrin-bound plasminogen, however, and its use is limited by its significant rates of antigenicity. Fevers and shivering occur in 1% to 4% of patients.

Urokinase is derived from human neonatal kidney cells grown in tissue culture. Currently, it is only approved for lysis of massive PE or PE associated with unstable hemodynamics.

Alteplase, reteplase, and tenecteplase all are recombinant variants of tissue plasminogen activator. Alteplase is indicated for the treatment of acute myocardial infarction, acute ischemic stroke, and acute massive PE. However, it is often used for CDT of DVT. Reteplase and tenecteplase are indicated only for the treatment of acute myocardial infarction.

Systemic thrombolysis was evaluated in numerous older prospective and randomized clinical trials, and its efficacy was summarized in a recent Cochrane Review. In 12 studies involving over 700 patients, systemic thrombolysis was associated with significantly more clot lysis (relative risk [RR] 0.24 to 0.57) and significantly less PTS (RR 0.66). However, venous function was not significantly improved. In addition, more bleeding complications occurred (RR 1.73).

In an effort to minimize bleeding complications and increase efficacy, CDT techniques were developed for the treatment of symptomatic primarily iliofemoral DVT. With catheter-directed therapy, venous access may be achieved through percutaneous catheterization of the ipsilateral popliteal vein, retrograde catheterization through the contralateral femoral vein, or retrograde cannulation from the internal jugular vein. Multi–side-hole infusion catheters, with or without infusion wires, are used to deliver the lytic agent directly into the thrombus. Lytic agents may be administered alone or, now more commonly, in combination with catheter-based methods to
physically break up the clot—so-called pharmacomechanical thrombolysis. One commonly used device to perform pharmacomechanical thrombolysis is the AngioJet, which utilizes pulsed injection of thrombolytic via a percutaneously inserted catheter followed by active aspiration to remove the thrombus.

The efficacy of CDT for the treatment of symptomatic iliofemoral DVT has been reported previously in a large, multicenter, randomized control trial. Two-hundred and nine patients with proximal DVT identified within 21 days of onset of symptoms were assigned to conventional anticoagulant therapy vs. conventional anticoagulant therapy plus CDT. In the CDT group, placement of a venous stent was permitted for any identified iliac vein stenotic lesion. At 6 months, iliocaval patency was significantly improved in the thrombolysis group (65.9% vs. 47.4%). At 2 years, in the CDT group, there was an absolute risk reduction of nearly 15% for development of PTS, translating to a number needed to treat of seven patients to prevent one case of PTS.60 However, these results were, in part, contradictory to the results reported the more recent Acute Venous Thrombosis: Thrombus Removal with Adjunctive Catheter-Directed Thrombolysis (ATTRACT) trial, a prospective, randomized, multicenter trial evaluating nearly 700 patients comparing anticoagulant use with CDT in patients with acute femoropopliteal, and/or iliac vein DVT.61 However, the purpose of this trial was to see if indications for CDT should be extended for isolated femoropopliteal DVT (43% of the patients in this study) and therefore, was underpowered to compare treatment efficacy for iliofemoral DVT, which known to have a higher risk of PTS. The study found that PTS occurred with equal frequency in the two groups (47% vs. 48%, \( P = \text{NS} \)), but patients who were treated with pharmacomechanical CDT plus anticoagulation were less likely to develop moderate-to-severe PTS than those treated with anticoagulation alone (18% vs. 24%, \( P = .003 \)). There was no difference between the two groups in quality of life. There was an increase in both overall hemorrhage (4.5% vs. 1.7%) and major hemorrhage (1.7% vs 0.3%) with CDT but no fatal or intracranial hemorrhage in either cohort. Taken in combination, the findings from these trials support selective use of CDT with anticoagulation in young patients with acute iliofemoral DVT and anticoagulation alone in the remaining patient with DVT.

There are contraindications to thrombolytic therapy. Absolute contraindications include prior history of ischemic or hemorrhagic stroke within 3 months, head trauma within 3 months, neurologic surgery within 6 months, known intracranial neoplasm, internal bleeding within 6 weeks, active or known bleeding disorder, traumatic cardiopulmonary resuscitation within 3 weeks or suspected aortic dissection. Fortunately, serious remote bleeding is uncommon, and intracranial hemorrhage rarely occurs. The majority of bleeding complications are limited to the venous access site. Symptomatic pulmonary embolism occurs uncommonly and is very rarely fatal.

**Inferior Vena Caval Filters.** Since the introduction of the Kimray-Greenfield filter in the United States in 1973, numerous vena caval filters have been developed. Although the designs are variable, they all are designed to prevent pulmonary emboli, while allowing continuation of venous blood flow through the IVC. Early filters were placed surgically through the femoral vein. Currently, less invasive techniques allow percutaneous filter placement through a femoral vein, internal jugular vein, or a peripheral vein under fluoroscopic or ultrasound guidance.

Placement of an IVC filter is indicated for patients who have manifestations of lower extremity VTE and absolute contraindications to anticoagulation, those that have a bleeding complication from anticoagulation therapy of acute VTE, or those who develop recurrent DVT or PE despite adequate anticoagulation therapy and for patients with severe pulmonary hypertension.

When possible, anticoagulation therapy should be continued in patients with vena cava filters. The duration of anticoagulation is determined by the underlying VTE and not by the presence of the IVC filter itself. Practically speaking, however, many patients who require an IVC filter for recurrent VTE are the same ones who would benefit most from indefinite anticoagulation. In patients who are not able to receive anticoagulants due to recent surgery or trauma, the clinician should continually reassess if anticoagulation may be started safely at a later date.

Placement of permanent IVC filters has been evaluated as an adjunct to routine anticoagulation in patients with proximal DVT.62 Routine IVC filter placement has not been shown to prolong early or late survival in patients with proximal DVT but did decrease the rate of PE (HR, 0.22; 95% CI, 0.05–0.90); however, there is an increased rate of recurrent DVT in patients with IVC filters (HR, 1.87; 95% CI, 1.10–3.20).

IVC filters are associated with acute and late complications. Acute complications include thrombosis or bleeding at the insertion site and misplacement of the filter. Late complications include thrombosis of the IVC, DVT, breaking, migration, or erosion of the filter through the IVC (Fig. 24-9). The rate of fatal complications is <0.12%.63 As a result of the increasing number

---

**Figure 24-9.** Preoperative computed tomography imaging and intraoperative photo demonstrating erosion of IVC filter through the IVC wall.
of reported complications with IVC filters, the FDA issued a warning in 2010 recommending removal of IVC filters as soon as they are no longer needed. This was followed by an update in 2014 where the recommendation was made to remove IVC filters within 29 and 54 days after implantation based upon a mathematical model that suggested an increased risk-to-benefit ratio at this time point.

In some patients, the need for an IVC filter may be self-limited. Such patients can be treated with so-called removable IVC filters. Depending on the device, removable IVC filters are potentially removable by percutaneous endovascular techniques for up to several months after their initial implantation, assuming the filter is no longer required and does not have large amounts of trapped thrombi. IVC filters that have been in place for an extended period of time may require adjunctive techniques, including laser-assisted removal or open surgical removal when they are embedded within the vena cava. All temporary IVC filters are approved for permanent implantation, and many so-called temporary filters end up as permanent devices with all the potential complications of permanent IVC filters.

**Operative Venous Thrombectomy.** In patients with acute iliofemoral DVT, surgical therapy is generally reserved for patients who worsen with anticoagulation therapy and those with phlegmasia cerulea dolens and impending venous gangrene. If the patient has phlegmasia cerulea dolens, a fasciotomy of the calf compartments is first performed. In iliofemoral DVT, a longitudinal venotomy is made in the common femoral vein, and a venous balloon embolectomy catheter is passed through the thrombus into the IVC and pulled back several times until no further thrombus can be extracted. The distal thrombus in the leg is removed by manual pressure beginning in the foot. This is accomplished by application of a tight rubber elastic wrap beginning at the foot and extending to the thigh. If the thrombus in the femoral vein is old and cannot be extracted, the vein may be ligated. For a thrombus that extends into the IVC, the IVC is exposed transperitoneally and controlled below the renal veins. The IVC is opened, and the thrombus is removed by gentle massage. An intraoperative completion venogram determines if any residual thrombus or stenosis is present. If a residual iliac vein stenosis is present, intraoperative angioplasty and stenting can be performed. In most cases, an arteriovenous fistula is then created by anastomosing the great saphenous vein (GSV) end to side with the superficial femoral artery in an effort to maintain patency of the thrombectomized iliofemoral venous segment. Heparin is administered postoperatively for several days. Warfarin anticoagulation is maintained for at least 6 months after thrombectomy. Complications of iliofemoral thrombectomy include PE in up to 20% of patients and death in <1% of patients.

One study followed 77 limbs for a mean of 8.5 years after thrombectomy for acute iliofemoral DVT. In limbs with successful thrombectomy, valvular competence in the thrombectomized venous segment was 80% at 5 years and 56% at 10 years. More than 90% of patients had minimal or no symptoms of PTS. There were 12 (16%) early thrombectomy failures. Patients were required to wear compression stockings for at least 1 year after thrombectomy.

Survival rates for surgical pulmonary embolectomy have improved over the past 20 years with the addition of cardiopulmonary bypass. Emergency pulmonary embolectomy for acute PE is rarely indicated. Patients with preterminal massive PE (Fig. 24-10) for whom thrombolysis has failed or who have contraindications to thrombolytics may be candidates for this procedure. Open pulmonary artery embolectomy is performed through a posterolateral thoracotomy with direct visualization of the pulmonary arteries. Mortality rates range between 20% and 40%.

Percutaneous catheter-based techniques for removal of a PE involve mechanical thrombus fragmentation or embolectomy using suction devices. Mechanical clot fragmentation is followed by CDT. Results of catheter-based fragmentation are based on small case series. In a study in which a fragmentation device was used in 10 patients with acute massive PE, fragmentation was successful in 7 patients with a mortality rate of 20%. Transvenous catheter pulmonary suction embolectomy has also been performed for acute massive PE with a reported 76% successful extraction rate and a 30-day survival of 70%.

**Prophylaxis**

Patients who undergo major general surgical, gynecologic, urologic, and neurosurgical procedures without thromboprophylaxis have a significant incidence of perioperative DVT. An estimated one-third of the 150,000 to 200,000 VTE-related deaths per year in the United States occur following surgery. The goal of prophylaxis is to reduce the mortality and morbidity associated with VTE. The first manifestation of VTE may be a life-threatening PE (Fig. 24-11), and as indicated earlier, clinical evaluation to detect DVT before PE is unreliable.

Effective methods of VTE prophylaxis involve the use of one or more pharmacologic or mechanical modalities. Currently available pharmacologic agents include low-dose UFH, LMWH, synthetic pentasaccharides, and vitamin K antagonists. Mechanical methods include intermittent pneumatic compression (IPC) and graduated compression stockings. There is insufficient evidence to consider aspirin alone as adequate DVT prophylaxis. Methods of prophylaxis vary with regard to efficacy, and the 2012 ACCP Clinical Practice Guidelines stratify their uses according to the patient’s level of VTE risk, bleeding risk, and the values and preferences of individual patients (see Table 24-3).

**Venous Thromboembolism Prophylaxis in Nonorthopedic Surgery.** The risk for VTE associated with a surgical procedure depends on the type of operation, type of anesthesia, duration of surgery, and other risk factors, such as patient age, presence of cancer, prior VTE, obesity, presence of infection, and known thrombophilic disorders. VTE risk can be stratified according to
the previously mentioned risk assessment models, the Caprini score and Rogers score. These risk assessment models are included in the prophylaxis guidelines for nonorthopedic surgery (Tables 24-5 and 24-6). A composite score is created using assigned values for each risk factor. The cumulative score for each patient is then used to predict thrombosis risk and provide recommendations regarding VTE prophylaxis.

Patients at very low risk (<0.5%; Rogers score <7; Caprini score 0) who undergo general or abdominopelvic procedures do not require pharmacologic or mechanical prophylaxis; however, early ambulation is required. Patients at low risk (<1.5%; Rogers score 7–10; Caprini score 1–2) should receive mechanical prophylaxis. Patients at moderate risk (3%; Rogers score >10; Caprini score 3–4) should receive LMWH at recommended doses, low-dose UFH, or mechanical prophylaxis. Patients at high risk (6%; Caprini score ≥5) should receive LMWH at recommended doses or low-dose UFH and mechanical prophylaxis. Thromboprophylaxis should continue until discharge, except in select high-risk patients with malignancy in whom extended-duration prophylaxis (up to 4–6 weeks) may be beneficial. Patients with significant risk for bleeding should receive mechanical prophylaxis until this risk subsides.75

Overall, low-dose UFH and LMWH reduce the risk for symptomatic and asymptomatic VTE by 60% to 70%. The risks for bleeding differ, depending on the dosage. Lower dosages of LMWH appear to be associated with less bleeding risk than low-dose UFH, but the latter produces less bleeding risk than higher prophylactic dosages of LMWH.76 Other advantages of LMWH include once-daily dosing protocols and a lower rate of heparin-associated antibody formation.

Fondaparinux has been compared with the LMWH dalteparin in patients who undergo high-risk major abdominal surgery. It also has been compared with IPC alone in patients undergoing non–high-risk abdominal surgery.77,78 Fondaparinux demonstrated rates of VTE prevention, bleeding complications, and mortality similar to those of LMWH. It was more beneficial than IPC alone in reducing VTE but with a higher rate of bleeding (1.6% vs. 0.2%).

Prophylactic insertion of IVC filters has been suggested for VTE prophylaxis in high-risk trauma patients, bariatric surgical patients, and some patients with malignancy who have contraindications for LMWH therapy.79 A 5-year study of prophylactic IVC filter placement in 132 trauma patients at high risk of PE (head injury, spinal cord injury, pelvic or
### Table 24-6
**Caprini risk assessment model**

<table>
<thead>
<tr>
<th>1 POINT</th>
<th>2 POINTS</th>
<th>3 POINTS</th>
<th>5 POINTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age 41–60</td>
<td>Age 61–74</td>
<td>Age ≥75</td>
<td>Stroke (&lt;1 month)</td>
</tr>
<tr>
<td>Minor surgery</td>
<td>Arthroscopic surgery</td>
<td>History of VTE</td>
<td>Elective arthroplasty</td>
</tr>
<tr>
<td>BMI &gt;25 kg/m²</td>
<td>Major open surgery (&gt; 45 minutes)</td>
<td>Family history of VTE</td>
<td>Hip, pelvis, or leg fracture</td>
</tr>
<tr>
<td>Swollen legs</td>
<td>Laparoscopic surgery (&gt; 45 minutes)</td>
<td>Factor V Leiden</td>
<td>Acute spinal cord injury (&lt;1 month)</td>
</tr>
<tr>
<td>Varicose veins</td>
<td>Malignancy</td>
<td>Prothrombin 20210A</td>
<td></td>
</tr>
<tr>
<td>Pregnancy or postpartum</td>
<td>Confined to bed (&gt;72 hours)</td>
<td>Lupus anticoagulant</td>
<td></td>
</tr>
<tr>
<td>History of unexplained or recurrent spontaneous abortion</td>
<td>Immobilizing plaster cast</td>
<td>Anticardiolipin antibody</td>
<td></td>
</tr>
<tr>
<td>Oral contraceptives of hormone replacement</td>
<td>Central venous access</td>
<td>Elevated serum homocysteine</td>
<td></td>
</tr>
<tr>
<td>Sepsis (&lt;1 month)</td>
<td></td>
<td>Heparin-induced thrombocytopenia</td>
<td></td>
</tr>
<tr>
<td>Serious lung disease, including pneumonia (&lt;1 month)</td>
<td></td>
<td>Other congenital or acquired thrombophilia</td>
<td></td>
</tr>
<tr>
<td>Abnormal pulmonary function test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute myocardial infarction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>History of inflammatory bowel disease</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical patient at bed rest</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

BMI = body mass index; VTE = venous thromboembolism.


---

long bone fractures) reported a 0% incidence of symptomatic PE in patients with a correctly positioned IVC filter. In 47 patients with a malpositioned IVC filter (strut malposition or filter tilt), there was a 6.3% incidence of symptomatic PE with three deaths. DVT occurred at the insertion site in 3.1% of the patients. IVC patency was 97.1% at 3 years.

Fatal and nonfatal PE can still occur in patients with vena cava interruption. As noted earlier, long-term complications associated with permanent IVC filters include IVC thrombosis and DVT. Currently, the ACCP recommends IVC filters be placed only if a proximal DVT is present and anticoagulation therapy is contraindicated. Placement of an IVC filter in the setting of severe pulmonary embolism development while anticoagulated remains controversial. IVC filter insertion is not recommended for primary prophylaxis.

Removable IVC filters may be placed in patients with a temporarily increased risk of PE. The best patient groups for retrievable filter placement may include young trauma patients with transient immobility, patients undergoing surgical procedures associated with a high risk of PE, and patients with hypercoagulable states who cannot receive anticoagulation therapy for a short period of time. Careful follow-up is required to assure all potentially removable filters are in fact removed.

---

**OTHER VENOUS THROMBOTIC DISORDERS**

**Superficial Vein Thrombophlebitis**

Superficial vein thrombophlebitis (SVT) most commonly occurs in varicose veins but can occur in normal veins. When SVT recurs at variable sites in normal superficial veins, *thrombophlebitis migrans*, it may signify a hidden visceral malignancy or a systemic disorder such as a blood dyscrasia and/or a collagen vascular disease. SVT also frequently occurs as a complication of indwelling catheters, with or without associated extravasation of injected material. Upper extremity vein thrombosis has been reported to occur in 38% of patients with peripherally inserted central catheters; 57% of these developed in the cephalic vein (Fig. 24-12). Suppurative SVT may occur in veins with indwelling catheters and may be associated with generalized sepsis.

Clinical signs of SVT include redness, warmth, and tenderness along the distribution of the affected veins, often associated with a palpable cord. Patients with suppurative SVT may have fever and leukocytosis. DUS should be performed in patients with signs and symptoms of acute SVT to confirm the diagnosis and to determine if any associated DVT is present. Concomitant lower extremity DVT may be present in 5% to 40% of patients with SVT; most occur in patients with greater
VENOUS AND LYMPHATIC DISEASE

CHAPTER 24

Figure 24-12. Duplex ultrasound of a brachial vein containing thrombus and percutaneously inserted central catheter (PICC).

Saphenous vein SVT within 1 cm of the saphenofemoral junction. A follow-up DUS should be performed in 5 to 7 days in patients with SVT in the proximal GSV but without deep vein involvement. Approximately 10% to 20% of patients with SVT involving the proximal GSV experience progression to deep venous involvement within 1 week.83,84

Treatment of SVT is quite variable. A Cochrane Review reported that LMWHs and nonsteroidal anti-inflammatory drugs both reduce the rate of SVT extension or recurrence.85 Additionally, a multicenter, randomized, blinded trial comparing use of fondaparinux to placebo in lower extremity SVT found use of fondaparinux reduced the rate of VTE formation by 85%, though the incidence of VTE formation was low in both groups (0.2% vs. 1.3%, \( P < .001 \)). The number needed to treat to prevent one episode of VTE was 88 patients. There was no difference in mortality between the two arms of the trial.85

Topical medications appear to improve local symptoms. Surgical treatment, combined with the use of graduated compression stockings, is associated with a lower rate of VTE and SVT progression.86 The treatment is individualized and depends on the location of the thrombus and the severity of symptoms. In patients with SVT not within 1 cm of the saphenofemoral junction, treatment consists of compression and administration of an anti-inflammatory medication such as indomethacin. In patients with suppurative SVT, antibiotics and removal of any existing indwelling catheters are mandatory. Excision of the vein may be necessary but is usually reserved for patients with systemic symptoms or when excision of the involved vein is straightforward. If the SVT extends proximally to within 1 cm of the saphenofemoral junction, extension into the common femoral vein is more likely to occur. In these patients, anticoagulation therapy for 6 weeks and GSV ligation appear equally effective in preventing thrombus extension into the deep venous system.87,88

Upper Extremity Vein Thrombosis

Axillary-subclavian venous thrombosis (ASVT) is classified into two forms. Primary ASVT occurs in only a small minority of all patients with ASVT. In the primary form, no clear cause for the thrombosis is readily identifiable at initial evaluation. Patients with primary ASVT often give a history of performing prolonged, repetitive motion activities, which results in damage to the subclavian vein, usually where it passes between the head of the clavicle and the first rib in association with the subclavious muscle. This condition is also known as venous thoracic outlet syndrome, effort thrombosis, and Paget-Schroetter syndrome. Secondary ASVT is more common and is associated with an easily identified cause such as an indwelling catheter or a hypercoagulable state. Over 30% of patients with tunneled subclavian vein access devices develop ASVT.89

A patient with ASVT may be asymptomatic or may present with varying degrees of upper extremity edema, tenderness, and conspicuous superficial venous enlargement. DUS can be performed initially to confirm the diagnosis, but limitations to the exam by the clavicle and collateralization can lead to a false-negative study. Venography is recommended when there is nonconcordance between the duplex study and clinical suspicion. Anticoagulation therapy should be initiated once ASVT is diagnosed to prevent PE and decrease symptoms.

Treatment of patients with primary upper extremity venous thrombosis is controversial because the natural history of the disease may vary from minimal to no symptoms to significant symptoms with vigorous upper extremity activities. In recent years, patients presenting with acute symptomatic primary ASVT are often considered candidates for CDT therapy with the goal of minimizing long-term symptoms of venous congestion. Venography is performed through a catheter placed using an ultrasound-guided percutaneous basilic vein approach to document the extent of the thrombus (Fig. 24-13). A guidewire is traversed through the thrombus, and a catheter is placed within the thrombus. Typically, tissue plasminogen activator is administered through a multi–side-hole infusion catheter. Various catheter-based mechanical techniques may also be employed to speed thrombus removal. Heparin is administered concurrently with the thrombolytic infusion. After completion of thrombolytic therapy, a follow-up venogram is obtained. Correctable anatomic abnormalities may then be considered for treatment. Adjuvant procedures after thrombolytic therapy may include

Figure 24-13. Upper extremity venogram showing stenosis of the right subclavian vein at the first rib (arrow).
cervical or first rib resection for thoracic outlet abnormalities, scalenectomy, surgical venous reconstruction, and balloon angioplasty of residual venous stenosis. The ACCP guidelines recommend the same intensity and duration of anticoagulant therapy in patients who undergo thrombolysis as in patients who are treated with anticoagulation alone.

**Mesenteric Vein Thrombosis**

Five percent to 15% of cases of acute mesenteric ischemia occur as a result of mesenteric vein thrombosis (MVT). Mortality rates in patients with MVT may approach 50%. The usual presenting symptom is nonspecific abdominal pain and distention, often accompanied by nausea, vomiting, and diarrhea. Peritoneal signs, suggesting intestinal infarction, are present in fewer than half of MVT patients. MVT is more common in patients with a hypercoagulable states, malignancy, and cirrhosis. MVT also occurs as a rare complication of laparoscopic surgery.

Most cases of MVT are diagnosed with contrast-enhanced CT scanning or magnetic resonance imaging (MRI) in the course of an evaluation for abdominal pain. The sensitivity and specificity for CT and MRI approach 100% and 98%, respectively. Ultrasound can also be used and has reported sensitivity and specificity of 93% and 99%, respectively.

Patients with MVT are treated with fluid resuscitation, heparin anticoagulation, and bowel rest. Once the patient’s clinical status improves, oral intake can be carefully started. The patient is transitioned to oral anticoagulation over 3 to 4 days and, depending on the etiology of the MVT, continued for 3 to 6 months or indefinitely. Most patients with MVT can be treated nonoperatively, but urgent laparotomy is indicated in patients with peritoneal findings. Broad-spectrum antibiotics are administered perioperatively. Operative findings consist of edema and cyanotic discoloration of the mesentery and bowel wall. In more advanced cases, thrombus involves the distal mesenteric veins. The arterial supply to the involved bowel is usually intact. Nonviable bowel is resected, and primary anastomosis can be performed. If the viability of the remaining bowel is in question, a second-look operation is performed within 24 to 48 hours.

**VARICOSE VEINS**

Varicose veins are common and are present in at least 10% of the general population. The findings of varicose veins may include dilated and tortuous veins, telangiectasias, and fine reticular varicosities. Risk factors for varicose veins include obesity, female sex, inactivity, and family history. Varicose veins can be classified as primary or secondary. Primary varicose veins result from intrinsic abnormalities of the venous wall, whereas secondary varicose veins are associated with deep and superficial venous insufficiency.

Patients with varicose veins may complain of unsightly appearance, aching, heaviness, pruritus, and early fatigue of the affected leg. These symptoms worsen with prolonged standing and sitting and are relieved by elevation of the leg above the level of the heart. A mild amount of edema is often present. More severe signs include thrombophlebitis, hyperpigmentation, lipodermatosclerosis, ulceration, and bleeding from attenuated vein clusters.

An important component of treatment for patients with varicose veins is the use of elastic compression stockings. Patients may be prescribed elastic stockings with compression ranging from 20 to 30, 30 to 40, or even 40 to 50 mmHg. Stockings range in length from knee high to waist high, and they should cover the symptomatic varices. Elastic compression provides sufficient relief of symptoms in many symptomatic patients.

Cosmetic concerns may lead to intervention. Additionally, interventions are warranted in patients whose symptoms worsen or are unrelieved despite compression therapy or who have lipodermatosclerosis or venous ulcer. Randomized trials of symptomatic patients with varicose veins have demonstrated improved quality of life with interventional treatment. Interventional management includes injection sclerotherapy, thermal ablation, surgical therapy, or a combination of these techniques. Injection sclerotherapy alone can be successful in varicose veins and in telangiectatic vessels. A recent multicenter, randomized trial compared foam sclerotherapy versus placebo for symptomatic varicose veins found significant symptom relief and improved cosmetic appearance with sclerotherapy. Sclerotherapy acts by destroying the venous endothelium. Sclerosing agents include hypertonic saline, sodium tetradecyl sulfate, and polidocanol. Concentrations of 11.7% to 23.4% hypertonic saline, 0.125% to 0.250% sodium tetradecyl sulfate, and 0.5% polidocanol are used for telangiectasias. Larger varicose veins require higher concentrations: 23.4% hypertonic saline, 0.50% to 1% sodium tetradecyl sulfate, and 0.75% to 1.0% polidocanol.

Elastic bandages are wrapped around the leg after injection and worn continuously for 3 to 5 days to produce apposition of the inflamed vein walls and prevent thrombus formation. After the bandages are removed, elastic compression stockings should be worn for a minimum of 2 weeks. Complications from sclerotherapy include allergic reaction, local hyperpigmentation, thrombophlebitis, DVT, and possible skin necrosis.

Newer devices combine sclerotherapy with catheter-based mechanical endoluminal injury to achieve nonthermal ablation. Additional nonthermal, nonsclerosant ablative techniques using proprietary adhesive formulations with cyanoacrylate are current being evaluated and have demonstrated promising early results.

Patients with symptomatic GSV or SSV reflux may be treated with endovenous ablation techniques or surgical removal of the affected vein. Endovenous laser and radiofrequency ablation (RFA) techniques have gained in popularity in the past several years. Such techniques are generally associated with equally effective but more rapid postprocedure recovery than traditional open surgical stripping of the GSV.

With either endoluminal technique, the distal thigh or proximal calf GSV is punctured with a 21-gauge needle under ultrasound guidance. A sheath is placed over a guidewire, and the laser fiber or RFA catheter is advanced until it is near to, but not at, the saphenofemoral junction. Tumescent anesthetic is administered around the GSV, and the vein is treated as the catheter is withdrawn. Endovenous laser treatment and RFA result in durable ablation of the GSV, with rates of varicose vein recurrence and clinical severity scores comparable to those seen with open surgery.

Risks of endovenous ablation include DVT, ecchymosis, and saphenous nerve injury.

Saphenous vein ligation and stripping may still be the preferred therapy for patients with GSVs of very large diameter (>2 cm). Surgical removal of the GSV usually is performed via small incisions placed medially in the groin and just below the knee. The GSV is removed using a blunt tip catheter or an invagination pin stripper. Complications associated with GSV stripping include ecchymosis, hematoma, lymphocele, DVT, infection, and saphenous nerve injury. GSV stripping is associated with a lower rate of recurrence of varicose veins and a better quality of life than saphenofemoral junction ligation alone.
Larger varicose veins are best treated by surgical excision using the “stab avulsion” technique. Stab avulsions are performed by making 2-mm incisions directly over branch varicosities, and the varicosity is dissected from the surrounding subcutaneous tissue as far proximally and distally as possible through the small incisions (Fig. 24-14). In most cases the vein is simply avulsed with no attempt at ligation. Bleeding is easily controlled with leg elevation, manual compression, and preprocedure tumescent anesthesia.

**CHRONIC VENOUS INSUFFICIENCY**

Chronic venous insufficiency (CVI) affects an estimated 600,000 people in the United States. Patients complain of leg fatigue, discomfort, and heaviness. Signs of CVI may include varicose veins, pigmentation, lipodermatosclerosis, and venous ulceration. Importantly, severe CVI is not necessarily associated with varicose veins. Chronic venous ulcers carry significant negative physical, financial, and psychological implications. A quality-of-life study reported that 65% of patients with chronic leg ulcers had severe pain, 81% had decreased mobility, and 100% experienced a negative impact of their disease on their work capacity. The socioeconomic impact of chronic venous leg ulcers is staggering, with an estimated 2 million workdays lost per year. The annual healthcare cost in the United States to treat CVI is estimated at $1 billion.

CVI can be primary or secondary. Primary CVI results from intrinsic abnormalities of the vein wall, whereas secondary CVI, so-called postthrombotic syndrome (PTS), occurs as a result of DVT. The signs and symptoms of CVI can therefore be attributed to venous reflux, venous obstruction, calf muscle pump dysfunction, or a combination of these factors, as well as loss of venous wall elasticity. In the majority of patients with CVI, the most important factor appears to be venous reflux, although most severe cases tend to have an obstructive etiology as well. Venous reflux results from abnormalities of the venous valve. Primary valvular reflux or incompetence is diagnosed when there is no known underlying cause of valvular dysfunction. Secondary valvular reflux is diagnosed when an identifiable cause is present. The most frequent secondary cause is DVT.

**Evaluation of Venous Insufficiency**

Early diagnostic studies to evaluate CVI required invasive measurements of ambulatory venous pressure (AVP) and venous recovery time (VRT). To measure AVP and VRT, a needle is inserted into a dorsal foot vein and connected to a pressure transducer. The patient is asked to perform 10 tiptoe exercises. Initially there is often a slight upward deflection of pressure with the onset of exercise followed by a decline in pressure with each subsequent tiptoe maneuver. After approximately 10 tiptoes, the measured pressure stabilizes and reflects a balance of venous inflow and outflow. The pressure at this point is the AVP, which is measured in millimeters of mercury. The patient is then asked to stop exercising to allow the vein to fill with return of the venous pressure to baseline. The time required for the venous pressure to return from the AVP level to 90% of baseline pressure is referred to as the VRT. A normal VRT typically ranges from 20 to 60 seconds. Values less than 20 seconds indicate significant reflux with increasing severity. To distinguish between superficial and deep venous reflux, a thigh tourniquet can be placed inflated to 50 mmHg to eliminate influence of the superficial venous. A VRT that remains below 20 seconds after tourniquet inflation suggests both superficial and deep venous reflux. Elevations of AVP indicate venous hypertension. The magnitude of AVP reflects the severity of CVI. There is an 80% incidence of venous ulceration in patients with an AVP of >80 mmHg.

**Plethysmography.** Noninvasive plethysmography has been used to evaluate CVI. Venous photoplethysmography indirectly evaluates venous function through the use of infrared light. A light-emitting diode is placed just above the medial malleolus, and the patient then performs a series of tiptoe maneuvers. Photoplethysmography provides a measurement of VRT. In limbs with CVI, VRT is shortened compared with that in a normal limb. AVP and VRT are only measures of the overall venous function of a lower extremity venous system. They cannot localize the site of reflux or evaluate the function of the calf pump.

Air plethysmography is a theoretically attractive but not widely used method to assess calf pump function, venous reflux, and overall lower extremity venous function. An air-filled plastic pressure bladder is placed on the calf to detect volume changes in the leg during a standard set of maneuvers. The patient is first supine, and then the leg is elevated and the minimum volume of venous blood recorded. The patient is then asked to assume an upright position with the examined leg non-weight bearing. The venous volume of the examined leg is determined when the volume curve flattens. The venous filling index (VFI), a measure of reflux, is calculated by dividing the maximum venous volume by the time required to achieve maximum venous volume. Next, the patient performs a single tiptoe maneuver, and the ejection fraction (EF) is determined. The EF is the volume change between the recorded volume before and after the tiptoe maneuver and is a measure of calf pump function. At this point, the veins of the leg are allowed to refill. The patient then performs 10 tiptoe maneuvers, and the residual volume fraction is calculated by dividing the venous volume in the leg after 10 tiptoe exercises by the venous volume present before the exercises. The residual volume fraction is a reflection of overall venous function. Theoretically, patients with increased VFIs and normal EFs (indicating the presence of reflux with normal calf pump function) would benefit from anti-reflux surgery, whereas patients with normal VFIs and diminished EFs would not.

**Venous Duplex Ultrasound.** Venous DUS has become the gold standard for evaluation of venous function largely supplanting venographic and plethysmographic techniques.
The principle advantage of DUS is that it can be used to evaluate reflux in individual venous segments targeting abnormal areas for treatment. The examination has been validated when performed with the patient in the standing position and the examined leg non-weight bearing. Pneumatic pressure cuffs of appropriate size are placed around the thigh, calf, and forefoot. The ultrasound transducer is positioned over the venous segment to be examined, just proximal to the pneumatic cuff (Fig. 24-15). The cuff is then inflated to a standard pressure for 3 seconds and then rapidly deflated. Ninety-five percent of normal venous valves close within 0.5 second.\textsuperscript{109} The presence of reflux for >0.5 second is considered abnormal. Typically, the common femoral, femoral, popliteal, and posterior tibial veins, as well as the GSV and SSV, are evaluated in a complete examination. The presence of reflux is considered abnormal. Typically, the common femoral, femoral, popliteal, and posterior tibial veins, as well as the GSV and SSV, are evaluated in a complete examination.

**Nonoperative Treatment of Chronic Venous Insufficiency**

**Compression Therapy.** Compression therapy is the mainstay of CVI management. Compression can be achieved using a variety of techniques, including elastic compression stockings, paste gauze boots (Unna’s boots), multilayer elastic wraps or dressings (Fig. 24-16), and pneumatic compression devices. Nonelastic compression bandages generally achieve higher and more prolonged degrees of compression than elastic compression bandages. The exact mechanism by which compression therapy can improve CVI remains uncertain. An improvement in skin and subcutaneous tissue microcirculatory hemodynamics as well as a direct effect on subcutaneous pressure have been hypothesized as the mechanisms of efficacy of compression therapy.\textsuperscript{110} Additionally, compression therapy has demonstrated quantifiable differences in ulcer healing with decreases in matrix metalloproteins and inflammatory cytokines.\textsuperscript{111,112} Clinically, routine use of elastic and nonelastic bandages reduces lower extremity edema in patients with CVI. In addition, supine perimalleolar subcutaneous pressure has been demonstrated to increase with elastic compression.\textsuperscript{111} With edema reduction, cutaneous metabolism may improve due to enhanced diffusion of oxygen and other nutrients to the cellular elements of skin and subcutaneous tissues. Increases in subcutaneous tissue pressure with elastic compression bandages may counteract transcapillary Starling forces, which favor leakage of fluid out of the capillary.

Before the initiation of therapy for CVI, patients must be educated about their chronic disease and the need to comply with their treatment plan to heal ulcers and prevent recurrence. A definitive diagnosis of venous ulceration must be made before treatment is initiated. A detailed history should be obtained from a patient presenting with lower extremity ulcerations, including medications used and associated medical conditions that may promote lower extremity ulceration. Arterial insufficiency is assessed by physical examination or noninvasive studies. In addition, systemic conditions that affect wound healing and leg edema, such as diabetes mellitus, immunosuppression, malnutrition, and congestive heart failure, should be improved as much as possible.

Compression therapy is most commonly achieved with graduated elastic compression stockings. Elastic compression stockings are available in various compositions, strengths, and lengths, and can be customized for a particular patient. The benefits of elastic compression stocking therapy for the treatment of CVI and healing of ulcerations have been well documented.\textsuperscript{114-117} In a retrospective study involving 113 venous ulcer patients,\textsuperscript{115} the use of below-knee, 30- to 40-mmHg elastic compression stockings, after edema and cellulitis were first resolved if present, resulted in 93% healing. Complete ulcer healing occurred in 99 of 102 patients (97%) who were compliant with stocking use vs. 6 of 11 patients (55%) who were noncompliant ($P < .0001$). The mean time to ulcer healing was 5 months. The rate of ulcer recurrence was lower in patients who were compliant with their compression therapy. By life table analysis, ulcer recurrence was 29% at 5 years for compliant patients and 100% at 3 years for noncompliant patients.

In addition to promoting ulcer healing, elastic compression therapy can also improve quality of life in patients with CVI. In one prospective study,\textsuperscript{116} 112 patients with CVI documented by DUS were administered a questionnaire to quantify the symptoms of swelling, pain, skin discoloration, cosmesis, activity tolerance, depression, and sleep alterations. Patients were treated with 30- to 40-mmHg elastic compression stockings. There were overall improvements in symptom severity scores at 1 month after initiation of treatment. Further improvements were noted at 16 months after treatment.

Patient compliance with compression therapy is crucial in treating CVI and especially venous leg ulcers. Many patients are initially intolerant of compression in areas of hypersensitivity adjacent to an active ulcer or at sites of previously healed ulcers. They may also have difficulty applying elastic stockings. To improve compliance, patients should be instructed to wear their stockings initially only as long as it is easily tolerated and...
then gradually to increase the amount of time the stockings are worn. Alternatively, patients can be fitted with lower-strength stockings initially followed by introduction of higher-strength stockings over a period of several weeks. Many commercially available devices, such as silk inner toe liners, adjustable elastic compression with Velcro, stockings with zippered sides (Fig. 24-17), and metal fitting aids (Fig. 24-18), are available to assist patients in applying elastic stockings. However, despite all these available adjuncts, many patients remain noncompliant with elastic compression therapy.

Another method of compression was developed by the German dermatologist Paul Gerson Unna. Unna’s boot has been used for many years to treat venous ulcers and is available in many versions. A typical Unna’s boot consists of a three-layer dressing and requires application by trained personnel. A rolled gauze bandage impregnated with calamine, zinc oxide, glycerin, sorbitol, gelatin, and magnesium aluminum silicate is first applied with graded compression from the forefoot to just below the knee. The next layer consists of a 10-cm-wide continuous gauze dressing followed by an outer layer of elastic wrap, also applied with graded compression. The bandage becomes stiff after drying, and the rigidity may aid in preventing edema formation. Unna’s boot is changed weekly or sooner if the patient experiences significant drainage and soiling of the dressing.

Once applied, Unna’s boot requires minimal patient involvement and provides continuous compression and topical therapy. However, Unna’s boot has several disadvantages. It is bulky and can be uncomfortable, which may affect patient compliance. In addition, the ulcer cannot be monitored after the boot is applied, the technique is labor intensive, and the degree of compression provided is operator dependent. Occasionally, patients may also develop contact dermatitis to the components of Unna’s boot.

The efficacy of Unna’s dressing has been studied. A retrospective 15-year survey encompassing 998 patients with one or more venous ulcers treated weekly with Unna’s dressing reported that 73% of ulcers healed in patients who returned for more than one treatment. The median time to healing for individual ulcers was 9 weeks. Unna’s dressing has been compared to other forms of treatment. A randomized, prospective study comparing Unna’s boot to polyurethane foam dressing in 36 patients with venous ulcers demonstrated superior healing over 12 months in patients treated with Unna’s boot (94.7% vs. 41.2%). A recent Cochrane Review of 39 randomized controlled trials demonstrated that compression increases ulcer healing rates compared with no compression, multicomponent systems are more effective than single-component systems, and multicomponent systems that include an elastic bandage are more effective than those composed mainly of inelastic constituents.

Other forms of compression dressing available to treat CVI include multilayered dressings and legging orthoses. The purported advantages of multilayered dressings include maintenance of compression for a longer period of time, more even distribution of compression, and better absorption of wound...
exudates. However, the efficacy of multilayered dressings depends on the wrapping technique of healthcare personnel. A commercially available legging orthosis consisting of multiple adjustable loop-and-hook closure compression bands provides compression similar to that of Unna’s boot and can be applied daily by the patient.\textsuperscript{122}

**Skin Substitutes.** Several types of skin substitutes are commercially available or under clinical study in the United States.\textsuperscript{105} Bioengineered skin ranges in composition from acellular skin substitutes to partial living skin substitutes. Their mechanism of action in healing venous ulcers is uncertain; however, they may serve as delivery vehicles for various growth factors and cytokines important in wound healing.

Apligraf\textsuperscript{123} is a commercially available bilayered living skin construct that closely approximates human skin for use in the treatment of venous ulcers. It contains a protective stratum corneum and a keratinocyte-containing epidermis overlaying a dermis consisting of dermal fibroblasts in a collagen matrix.\textsuperscript{122} Apligraf is between 0.5 mm and 1.0 mm thick and is supplied as a disk of living tissue on an agarose gel nutrient medium. It must be used within 5 days of release from the manufacturer\textsuperscript{123} (Fig. 24-19). The disk is easily handled and applied and conforms to irregularly contoured ulcer beds though it is very costly.

A prospective randomized study comparing multilayer compression therapy alone to treatment with Apligraf in addition to multilayered compression therapy has been performed to assess the efficacy of Apligraf in the treatment of venous ulcers.\textsuperscript{118} More patients treated with Apligraf had ulcer healing at 6 months (63% vs. 49%, $P = .02$). The median time to complete ulcer closure was significantly shorter in patients treated with Apligraf (61 days vs. 181 days, $P = .003$). The ulcers that showed the greatest benefit with the living skin construct were those that were large and deep (>1000 mm$^2$) or were longstanding (>6 months). No evidence of rejection or sensitization has been reported in response to Apligraf application.

**Surgical/Interventional Treatment of Chronic Venous Insufficiency Perforator Vein Ligation.** Incompetence of the perforating veins connecting the superficial and deep venous systems of the lower extremities has been implicated in the development of venous ulcers. The classic open technique described by Linton in 1938 for perforator vein ligation has a high incidence of wound complications and has largely been abandoned.\textsuperscript{124} A minimally invasive technique termed subfascial endoscopic perforator vein surgery (SEPS) evolved with improvement of endoscopic equipment.

DUS is performed preoperatively in patients undergoing SEPS to document deep venous competence and to identify perforating veins in the posterior compartment. The patient is positioned on the operating table with the affected leg elevated at 45° to 60°. An Esmarch bandage and a thigh tourniquet are used to exsanguinate the limb. The knee is then flexed, and two small incisions are made in the proximal medial leg away from areas of maximal induration at the ankle. Laparoscopic trocars are then positioned, and the subfascial dissection is performed with a combination of blunt and sharp dissection. Carbon dioxide is then used to insufflate the subfascial space. The thigh tourniquet is inflated to prevent air embolism. The perforators are then identified and doubly clipped and divided. After completion of the procedure, the leg is wrapped in a compression bandage for 5 days postoperatively.

The efficacy of SEPS as a stand-alone procedure in treatment of venous insufficiency is controversial and unproven. In a report from a large North American registry of 146 patients undergoing SEPS\textsuperscript{125} (Fig. 24-20), healing was achieved in 88% of ulcers (75 of 85) at 1 year. Adjunctive procedures, primarily superficial vein stripping, were performed in 72% of patients. Ulcer recurrence was predicted to be 16% at 1 year and 28% at 2 years by life table analysis. These results are similar to those achieved in some studies with compression therapy alone. A review of several studies from 2003 to 2011 demonstrated, when taken in aggregate, that 2059 limbs with 896 ulcers underwent SEPS and concomitant saphenous vein ablation (70%) with a 0% to 16% complication rate and achieved ulcer healing in 90% of patients.\textsuperscript{126} There has been a multicenter, prospective, European trial performed in patients with venous ulcers to evaluate the efficacy of SEPS. Post hoc analysis suggested possible benefit for SEPS in certain categories of patients with venous ulcer. Overall, however, primary analysis of the study’s end points indicated no advantage to SEPS in addition to superficial venous surgery and compression in the healing of venous ulcers.\textsuperscript{127} The technique appears to have fallen out of favor in most institutions with injection sclerotherapy preferred due to ease of use.

**Superficial Venous Surgery.** Currently it is accepted that superficial venous surgery in addition to compression therapy has a role in the treatment of patients with venous ulcer. The ESCHAR trial was a randomized prospective trial performed in the United Kingdom to evaluate the combination of superficial

![Figure 24-19](image1.png) Apligraf skin graft material supplied as a disk on an agarose gel nutrient medium.

![Figure 24-20](image2.png) Trocar placement for subfascial endoscopic perforator vein surgery. (Used with permission from Dr. Pankaj Patel.)
venous surgery and compression vs. compression alone in the treatment of venous ulcer. Superficial venous surgery had no additive effect to compression alone in the healing of a venous ulcer but significantly reduced venous ulcer recurrence at 4 years. Based on the results of this trial, it is reasonable to offer ablation or removal of the GSV in addition to compression therapy in patients with abnormal saphenous veins and signs and/or symptoms of severe CVI.128

Deep Venous Valvular Reconstruction. In the absence of significant deep vein valvular incompetence, saphenous vein stripping and possibly perforator vein ligation can be effective in the treatment of CVI. However, in patients with a combination of superficial and deep vein valvular incompetence, the addition of deep vein valvular reconstruction theoretically may improve ulcer healing.130 Numerous techniques of deep vein valve correction have been reported. These techniques consist of repair of existing valves, transplant of venous segments from the arm, transposition of an incompetent vein onto an adjacent competent vein, and implantation of cryopreserved vein segments including competent valves.

Successful long-term outcomes of 60% to 80% have been reported for venous valve reconstructions by internal suture repair.129,130 However, among patients who initially had ulceration, 40% to 50% still had persistence or recurrence of ulcers in the long term.127,129

Valve transplantation involves replacement of a segment of incompetent femoral vein or popliteal vein with a segment of axillary or brachial vein with competent valves. Early results are similar to those for venous valve reconstruction.129,130 However, in the long term, the transplanted venous segments tend to develop incompetence, intimal hyperplasia, and cusp sinus thrombosis with long-term outcomes that are poorer than those for venous valve reconstructions. The outcomes for venous transposition are similar to those for valve transplantation.

Currently, reconstruction techniques for deep venous insufficiency and associated CVI are rarely performed.

Venous Stenting. Currently there is great interest in the role of venous stents in the treatment of CVI. Stenotic lesions of the iliac veins, primarily documented with IVUS, are being reported in a very high percentage of patients with edema, lipo-dermatosclerosis, or ulceration secondary to venous disease. It appears possible to percutaneously place stents in the iliac veins with near 100% technical success and excellent patency of the stent out to 4 years. Retrospective case series suggest favorable effects on ulcer healing, symptoms of CVI, and quality of life in patients with CVI. The role of venous stenting as an independent procedure in the treatment of patients with CVI remains an area of active investigation.131

LYMPHEDEMA

Pathophysiology

Lymphedema is extremity swelling that results from a reduction in lymphatic transport and accumulation of lymph within the interstitial space. It is caused by anatomic and/or physiologic abnormalities such as lymphatic hypoplasia, functional insufficiency, or absence of lymphatic valves.

The original classification system, described by Allen, is based on the cause of the lymphedema. Primary lymphedema is further subdivided into congenital lymphedema, lymphedema praecox, and lymphedema tarda. Congenital lymphedema may involve a single lower extremity, multiple limbs, the genitalia, or the face. The edema typically develops before 2 years of age and may be associated with specific hereditary syndromes (Turner syndrome, Milroy syndrome, Klippel-Trénaunay-Weber syndrome). Lymphedema praecox is the most common form of primary lymphedema, accounting for 94% of cases. Lymphedema praecox is far more common in women, with the gender ratio favoring women 10:1. The onset is during childhood or the teenage years, and the swelling involves the foot and calf. Lymphedema tarda is uncommon, accounting for <10% of cases of primary lymphedema. The onset of edema is after 35 years of age.

Secondary lymphedema is far more common than primary lymphedema. Secondary lymphedema develops as a result of lymphatic obstruction or disruption. Axillary node dissection leading to lymphedema of the arm is the most common cause of secondary lymphedema in the United States. Other causes of secondary lymphedema include radiation therapy, trauma, infection, and malignancy. Globally, filariasis (an infection caused by Wuchereria bancrofti, Brugia malayi, and Brugia timori) and environmental exposure to minerals in volcanic soil resulting in podoconiosis in barefoot populations are the most common causes of secondary lymphedema.

Clinical Diagnosis

In most patients, the diagnosis of lymphedema can be made based on the history and physical examination alone. Patients commonly complain of heaviness and fatigue in the affected extremity. The limb size increases throughout the day and decreases to some extent, usually minimally, over the course of the night when the patient is recumbent. The limb, however, never completely normalizes. In the lower extremity, the swelling classically involves the dorsum of the foot, and the toes have a squared-off appearance. In advanced cases, hyperkeratosis of the skin develops, and fluid weeps from lymph-filled vesicles (Fig. 24-21).
Recurrent cellulitis is a common complication of lymphedema. Repeated infection results in further lymphatic damage, worsening existing disease. The clinical presentation of cellulitis ranges from subtle erythema and worsening of edema to a rapidly progressive soft tissue infection with systemic toxicity.

Many medical conditions can cause edema. If the symptoms are mild, distinguishing lymphedema from other causes of leg swelling can be difficult. Venous insufficiency is often confused with lymphedema. However, patients with advanced venous insufficiency typically have lipodermatosclerosis in the gaiter region, skin ulceration, and/or varicose veins. Bilateral pitting edema is typically associated with congestive heart failure, renal failure, or a hypoproteinemic state.

**Radiologic Diagnosis**

**Duplex Ultrasound.** When a patient is evaluated for edema, it is often difficult to distinguish the early stages of lymphedema from venous insufficiency. DUS of the venous system can determine if there is concomitant venous thrombosis or venous reflux, perhaps contributing to extremity edema. The diagnostic modalities discussed in the following sections have limited use in clinical practice. They are invasive and tedious and rarely change the management of a patient with lymphedema. Most physicians rely on the patient’s history and physical examination alone to make the diagnosis of lymphedema.

**Lymphoscintigraphy.** Lymphoscintigraphy has become the most commonly, but still overall uncommonly, used diagnostic test to identify lymphatic abnormalities. It has largely replaced lymphangiography. A radiolabeled sulfur colloid (technetium 99m sulfur colloid) is injected into the subdermal, interdigital region of the affected limb. The lymphatic transport is monitored with a whole-body gamma camera, and major lymphatics and nodes can be visualized (Fig. 24-22). In normal individuals, tracer activity may be detected in the inguinal region within 15 to 60 minutes. Within 3 hours, uptake should be present in the pelvic and abdominal lymph nodes. In patients with lymphedema, various patterns may be seen on lymphoscintigraphy. There may be delayed or absent transport to the inguinal nodes. Increased cutaneous collaterals may be seen with obstruction of the primary axial channels. There may also be localized regions of reduced uptake in patients with prior node dissection or radiation therapy.

**Lymphangiography.** Radiologic lymphangiography is performed by first visualizing the lymphatics by injecting colored dye into the hand or foot. The visualized lymphatic segment is exposed through a small incision and cannulated with a 27- to 30-gauge needle. An oil-based dye is then injected slowly into the lymphatics over several hours. The lymphatic channels and nodes are then visualized with traditional radiographs (Figs. 24-23 and 24-24). Lymphangiography is reserved for patients with lymphangiectasia or lymphatic fistulas, and patients who are being considered for microvascular reconstruction.

**Management**

An important aspect of the management of lymphedema is patient understanding that there is no cure for lymphedema. The primary goals of treatment are to minimize swelling and to prevent recurrent infections. Controlling the chronic limb swelling can improve discomfort, heaviness, and tightness, and potentially reduce the progression of disease.132
Compression Garments. Graded compression stockings are widely used in the treatment of lymphedema. The stockings reduce the amount of swelling in the involved extremity by decreasing edema accumulation while the extremity is dependent. When worn daily, compression stockings have been associated with long-term maintenance of reduced limb circumference.\textsuperscript{134} They may also protect the tissues against chronically elevated intrinsic pressures, which lead to thickening of the skin and subcutaneous tissue.\textsuperscript{135} Compression stockings also offer a degree of protection against external trauma that may lead to cellulitis.

The amount of compression required for controlling lymphedema ranges from 20 to 60 mmHg and varies among patients. The stockings can be custom made or prefabricated and are available in above- and below-knee lengths. The stockings should be worn during waking hours. The garments should be replaced approximately every 6 months when they lose elasticity.

Bedrest and Leg Elevation. Elevation is an important aspect of controlling lower extremity swelling and is often the first recommended intervention. However, continuous elevation throughout the day can interfere with quality of life more than lymphedema itself. Elevation is an adjunct to lymphedema therapy but is not the mainstay of treatment.

Intermittent Pneumatic Compression Therapy. The use of IPC with a single-chamber or multichamber pump temporarily reduces edema and provides another adjunct to the use of compression stockings. These devices have been shown to be effective in reducing limb volume; however, use of compression stockings is necessary to maintain the volume reduction when the patient is no longer supine because fluid transport is not associated with the transport of macromolecules (proteins) from the tissue. Typically, IPC is used for 4 to 6 hours per day at home when the patient is supine, with pressure ranges between 30 and 60 mmHg demonstrated to be most effective.\textsuperscript{135}

Lymphatic Massage. Manual lymphatic drainage is a form of massage developed by Vodder\textsuperscript{136} that is directed at reducing edema. In combination with the use of compression stockings, manual lymphatic drainage is associated with a long-term reduction in edema and fewer infections per patient per year.\textsuperscript{137}

Antibiotic Therapy. Patients with lymphedema are at increased risk of developing cellulitis in the affected extremity due to microscopic breakdown in the skin barrier either secondary to swelling or unrecognized and untreated tinea pedis. Recurrent infection can damage the lymphatics, aggravating the edema and increasing the risk for subsequent infection. \textit{Staphylococcus} and \textit{β}-hemolytic \textit{Streptococcus} are the most common organisms causing soft tissue infection. Aggressive antibiotic therapy and elevation with compression are recommended at the earliest signs or symptoms of cellulitis. The drug of choice is penicillin or a cephalosporin active against \textit{Streptococcus} for 5 days. In patients with recurrent cellulitis despite methods to reduced edema, treatment with monthly intramuscular injections of benzathine penicillin 1.2 MU, twice-daily erythromycin 250 mg, or penicillin V 1 g daily has proven effective at suppression.\textsuperscript{138}

Surgery. A variety of surgical procedures have been devised for the treatment of lymphedema. Surgical treatment involves either excision of extra tissue\textsuperscript{139} or anastomosis of a lymphatic vessel to another lymphatic or vein.\textsuperscript{140} In excisional procedures, part or all of the edematous tissue is removed. This does not improve lymphatic drainage but debulks redundant tissue. The microsurgical procedures involve the creation of a lymphatic or lymphaticovenous anastomosis, which theoretically improves lymphatic drainage. No long-term follow-up data are available for these interventions, and therefore operative therapy for lymphedema is not well accepted worldwide. Furthermore, operative intervention has the potential to further obliterate lymphatic channels, worsening the edema.\textsuperscript{141}

SUMMARY

Lymphedema is a chronic condition caused by ineffective lymphatic transport, which results in edema and skin damage. Lymphedema is not curable, but the symptoms and long-term effects can be controlled with a combination of elastic compression stockings, limb elevation, pneumatic compression, and massage. Controlling the edema protects the skin and potentially prevents cellulitis.

REFERENCES

Entries highlighted in bright blue are key references.


SURGICAL ANATOMY

The esophagus is a muscular tube that starts as the continuation of the pharynx and ends as the cardia of the stomach. When the head is in a normal anatomic position, the transition from pharynx to esophagus occurs at the lower border of the sixth cervical vertebra. Topographically this corresponds to the cricoid cartilage anteriorly and the palpable transverse process of the sixth cervical vertebra laterally (Fig. 25-1). The esophagus is firmly attached at its upper end to the cricoid cartilage and at its lower end to the diaphragm; during swallowing, the proximal points of fixation move cranially the distance of one cervical vertebral body.

The esophagus lies in the midline, with a deviation to the left in the lower portion of the neck and upper portion of the thorax, and returns to the midline in the midportion of the thorax near the bifurcation of the trachea (Fig. 25-2). In the lower portion of the thorax, the esophagus again deviates to the left and anteriorly to pass through the diaphragmatic hiatus.
Key Points

1. Benign esophageal disease is common and is best evaluated with thorough physiologic testing (high resolution esophageal motility, 24-hour ambulatory pH measurement, and/or esophageal impedance testing) and anatomic testing (esophagoscopy, video esophagography, and/or computed tomography [CT] scanning).

2. Gastroesophageal reflux disease (GERD) is the most common disease of the gastrointestinal tract for which patients seek medical therapy. When GERD symptoms (heartburn, regurgitation, chest pain, and/or supraesophageal symptoms) are troublesome despite adequately dosed PPI, surgical correction may be indicated.

3. Barrett’s esophagus is the transformation of the distal esophageal epithelium from squamous to a specialized columnar epithelium capable of further neoplastic progression. The detection of Barrett’s esophagus on esophagoscopy and biopsy increases the future risk of cancer by >40x compared to individuals without Barrett’s esophagus.

4. Giant hiatal hernia, otherwise known as paraesophageal hernia, should be repaired when symptomatic or associated with iron deficiency anemia. Laparoscopic hiatal hernia repair with fundoplication is the most common approach to repair.

5. Achalasia is the most common primary esophageal motor disorder. It is characterized by an absence of peristalsis and a hypertensive nonrelaxing lower esophageal sphincter. It is best treated with laparoscopic Heller myotomy and partial fundoplication.

6. Most esophageal cancer presents with dysphagia, at which time it has invaded the muscularis of the esophagus and is often associated with lymph node metastases. The preferred treatment at this stage is multimodality therapy with chemoradiation therapy followed by open or minimally invasive esophagectomy.

Three normal areas of esophageal narrowing are evident on the barium esophagogram or during esophagoscopy. The uppermost narrowing is located at the entrance into the esophagus and is caused by the cricopharyngeal muscle. Its luminal diameter is 1.5 cm, and it is the narrowest point of the esophagus. The middle narrowing is due to an indentation of the anterior and left lateral esophageal wall caused by the crossing of the left main stem bronchus and aortic arch. The luminal diameter at this point is 1.6 cm. The lowermost narrowing is at the hiatus of the diaphragm and is caused by the gastroesophageal sphincter mechanism. The luminal diameter at this point varies somewhat, depending on the distention of the esophagus by the passage of food, but has been measured at 1.6 to 1.9 cm. These normal constrictions tend to hold up swallowed foreign objects, and the overlying mucosa is subject to injury by swallowed corrosive liquids due to their slow passage through these areas.

Figure 25-3 shows the average distance in centimeters measured during endoscopic examination between the incisor teeth and the cricopharyngeus, aortic arch, and cardia of the stomach. Manometrically, the length of the esophagus between the lower border of the cricopharyngeus and upper border of the lower sphincter varies according to the height of the individual.

Figure 25-1. A. Topographic relationships of the cervical esophagus: (a) hyoid bone, (b) thyroid cartilage, (c) cricoid cartilage, (d) thyroid gland, (e) sternoclavicular. B. Lateral radiographic appearance with landmarks identified as labeled in A. The location of C6 is also included (f). (Reproduced with permission from Shields TW: General Thoracic Surgery, 3rd ed. Philadelphia, PA: Lea & Febiger; 1989.)
The pharyngeal musculature consists of three broad, flat, overlapping fan-shaped constrictors (Fig. 25-4). The opening of the esophagus is collared by the cricopharyngeal muscle, which arises from both sides of the cricoid cartilage of the larynx and forms a continuous transverse muscle band without an interruption by a median raphe. The fibers of this muscle...
blend inseparably with those of the inferior pharyngeal constrictor above and the inner circular muscle fibers of the esophagus below. Some investigators believe that the cricopharyngeus is part of the inferior constrictor; that is, that the inferior constrictor has two parts, an upper or retrothyroid portion having diagonal fibers, and a lower or retrocricoid portion having transverse fibers. Keith in 1910 showed that these two parts of the same muscle serve totally different functions. The retrocricoid portion serves as the upper sphincter of the esophagus and relaxes when the retrothyroid portion contracts, to force the swallowed bolus from the pharynx into the esophagus.

The cervical portion of the esophagus is approximately 5 cm long and descends between the trachea and the vertebral column, from the level of the sixth cervical vertebra to the level of the interspace between the first and second thoracic vertebrae posteriorly, or the level of the suprasternal notch anteriorly. The recurrent laryngeal nerves lie in the right and left grooves between the trachea and the esophagus. The left recurrent nerve lies somewhat closer to the esophagus than the right, owing to the slight deviation of the esophagus to the left, and the more lateral course of the right recurrent nerve around the right subclavian artery. Laterally, on the left and right sides of the cervical esophagus are the carotid sheaths and the lobes of the thyroid gland.

The thoracic portion of the esophagus is approximately 20 cm long. It starts at the thoracic inlet. In the upper portion of the thorax, it is in intimate relationship with the posterior wall of the trachea and the prevertebral fascia. Just above the tracheal bifurcation, the esophagus passes to the right of the aorta. This anatomic positioning can cause a notch indentation in its left lateral wall on a barium swallow radiogram. Immediately below this notch, the esophagus crosses both the bifurcation of the trachea and the left main stem bronchus, owing to the slight deviation of the terminal portion of the trachea to the right by the aorta (Fig. 25-5). From there down, the esophagus passes over the posterior surface of the subcarinal lymph nodes (LN), and then descends over the pericardium of the left atrium to reach the diaphragmatic hiatus (Fig. 25-6). From the bifurcation of the trachea downward, both the vagal nerves and the esophageal nerve plexus lie on the muscular wall of the esophagus.

Dorsally, the thoracic esophagus follows the curvature of the spine and remains in close contact with the vertebral bodies. From the eighth thoracic vertebra downward, the esophagus moves vertically away from the spine to pass through the hiatus of the diaphragm. The thoracic duct passes through the hiatus of the diaphragm on the anterior surface of the vertebral column behind the aorta and under the right crus. In the thorax, the thoracic duct lies dorsal to the esophagus between the azygos vein on the right and the descending thoracic aorta on the left.

The abdominal portion of the esophagus is approximately 2 cm long and includes a portion of the lower esophageal sphincter (LES). It starts as the esophagus passes through the diaphragmatic hiatus and is surrounded by the phrenoesophageal membrane, a fibroelastic ligament arising from the subdiaphragmatic fascia as a continuation of the transversalis fascia lining the abdomen (Fig. 25-7). The upper leaf of the membrane attaches itself in a circumferential fashion around the esophagus, about 1 to 2 cm above the level of the hiatus. These fibers blend in with the elastic-containing adventitia of the abdominal esophagus and the cardia of the stomach. This portion of the esophagus is subjected to the positive-pressure environment of the abdomen.

The musculature of the esophagus can be divided into an outer longitudinal and an inner circular layer. The upper 2 to 6 cm of the esophagus contains only striated muscle fibers. From then on, smooth muscle fibers gradually become more abundant. Most clinically significant esophageal motility disorders involve only the smooth muscle in the lower two-thirds of the esophagus. When a long surgical esophageal myotomy is indicated, the incision needs to extend only this distance.

The longitudinal muscle fibers originate from a crico-esophageal tendon arising from the dorsal upper edge of the anteriorly located cricoid cartilage. The two bundles of muscle diverge and meet in the midline on the posterior wall of the esophagus about 3 cm below the cricoid (see Fig. 25-4). From this point on, the entire circumference of the esophagus is

---

**Figure 25-5.** A, Cross-section of the thorax at the level of the tracheal bifurcation. B, Computed tomographic scan at same level viewed from above: (a) ascending aorta, (b) descending aorta, (c) tracheal carina, (d) esophagus, (e) pulmonary artery. (Reproduced with permission from Shields TW: General Thoracic Surgery, 3rd ed. Philadelphia, PA: Lea & Febiger; 1989.)
covered by a layer of longitudinal muscle fibers. This configuration of the longitudinal muscle fibers around the most proximal part of the esophagus leaves a V-shaped area in the posterior wall covered only with circular muscle fibers. Contraction of the longitudinal muscle fibers shortens the esophagus. The circular muscle layer of the esophagus is thicker than the outer longitudinal layer. In situ, the geometry of the circular muscle is helical and makes the peristalsis of the esophagus assume a wormlike drive, as opposed to segmental and sequential squeezing. As a consequence, severe motor abnormalities of the esophagus assume a corkscrew-like pattern on the barium swallow radiogram.

The cervical portion of the esophagus receives its main blood supply from the inferior thyroid artery. The thoracic portion receives its blood supply from the bronchial arteries, with 75% of individuals having one right-sided and two left-sided branches. Two esophageal branches arise directly from the aorta. The abdominal portion of the esophagus receives its blood supply from the ascending branch of the left gastric artery and from inferior phrenic arteries (Fig. 25-8). On entering the wall of the esophagus, the arteries assume a T-shaped division to form a longitudinal plexus, giving rise to an intramural vascular network in the muscular and submucosal layers. As a consequence, the esophagus can be mobilized from the stomach to the level of the aortic arch without fear of devascularization and ischemic necrosis. Caution, however, should be exercised as to the extent of esophageal mobilization in patients who have had a previous thyroidectomy with ligation of the inferior thyroid arteries proximal to the origin of the esophageal branches.

Blood from the capillaries of the esophagus flows into a submucosal venous plexus, and then into a periesophageal

---

**Figure 25-6.** A. Cross-section of the thorax at the midleft atrial level. B. Computed tomographic scan at same level viewed from above: (a) aorta, (b) esophagus, (c) left atrium, (d) right atrium, (e) left ventricle, (f) right ventricle, (g) pulmonary vein. *(Reproduced with permission from Shields TW: General Thoracic Surgery, 3rd ed. Philadelphia, PA: Lea & Febiger; 1989.)*

**Figure 25-7.** Attachments and structure of the phrenoesophageal membrane. Transversalis fascia lies just above the parietal peritoneum. *(Reproduced with permission from Shields TW: General Thoracic Surgery, 3rd ed. Philadelphia, PA: Lea & Febiger; 1989.)*

**Figure 25-8.** Arterial blood supply of the esophagus. *(Reproduced with permission from Shields TW: General Thoracic Surgery, 3rd ed. Philadelphia, PA: Lea & Febiger; 1989.)*
venous plexus from which the esophageal veins originate. In the cervical region, the esophageal veins empty into the inferior thyroid vein; in the thoracic region, they empty into the bronchial, azygos, or hemiazygos veins; and in the abdominal region, they empty into the coronary vein (Fig. 25-9). The submucosal venous networks of the esophagus and stomach are in continuity with each other, and, in patients with portal venous obstruction, this communication functions as a collateral pathway for portal blood to enter the superior vena cava via the azygos vein.

The parasympathetic innervation of the pharynx and esophagus is provided mainly by the vagus nerves. The constrictor muscles of the pharynx receive branches from the pharyngeal plexus, which is on the posterior lateral surface of the middle constrictor muscle, and is formed by pharyngeal branches of the vagus nerves with a small contribution from cranial nerves IX and XI (Fig. 25-10). The cricopharyngeal sphincter and the cervical portion of the esophagus receive branches from both recurrent laryngeal nerves, which originate from the vagus nerves—the right recurrent nerve at the lower margin of the subclavian artery and the left at the lower margin of the aortic arch. They are slung dorsally around these vessels and ascend in the groove between the esophagus and trachea, giving branches to each. Damage to these nerves interferes not only with the function of the vocal cords but also with the function of the cricopharyngeal sphincter and the motility of the cervical esophagus, predisposing the individual to pulmonary aspiration on swallowing.

Afferent visceral sensory pain fibers from the esophagus end without synapse in the first four segments of the thoracic spinal cord, using a combination of sympathetic and vagal pathways. These pathways are also occupied by afferent visceral sensory fibers from the heart; hence, both organs have similar symptomatology.

The lymphatics located in the submucosa of the esophagus are so dense and interconnected that they constitute a single plexus (Fig. 25-11). There are more lymph vessels than blood capillaries in the submucosa. Lymph flow in the submucosal plexus runs in a longitudinal direction, and, on injection of a contrast medium, the longitudinal spread is seen to be about six times that of the transverse spread. In the upper two-thirds of the esophagus, the lymphatic flow is mostly cephalad, and, in the lower third, caudad. In the thoracic portion of the esophagus,
the submucosal lymph plexus extends over a long distance in
a longitudinal direction before penetrating the muscle layer to
enter lymph vessels in the adventitia. As a consequence of this
nonsegmental lymph drainage, a primary tumor can extend for
a considerable length superiorly or inferiorly in the submucosal
plexus. Consequently, free tumor cells can follow the submu-
cosal lymphatic plexus in either direction for a long distance
before they pass through the muscularis and on into the regional
LNs. The cervical esophagus has more direct segmental lymph
drainage into the regional nodes, and, as a result, lesions in this
portion of the esophagus have less submucosal extension and a
more regionalized lymphatic spread.

The efferent lymphatics from the cervical esophagus drain
into the paratracheal and deep cervical LNs, and those from the
upper thoracic esophagus empty mainly into the paratracheal
LNs. Efferent lymphatics from the lower thoracic esophagus
drain into the subcarinal nodes and nodes in the inferior pulmo-
nary ligaments. The superior gastric nodes receive lymph not
only from the abdominal portion of the esophagus, but also from
the adjacent lower thoracic segment.

PHYSIOLOGY

Swallowing Mechanism
The act of alimentation requires the passage of food and drink
from the mouth into the stomach. One-third of this distance con-
sists of the mouth and hypopharynx, and two-thirds is made up
by the esophagus. To comprehend the mechanics of alimenta-
tion, it is useful to visualize the gullet as a mechanical model
in which the tongue and pharynx function as a piston pump
with three valves, and the body of the esophagus and cardia
function as a worm-drive pump with a single valve. The three
valves in the pharyngeal cylinder are the soft palate, epiglottis,
and cricopharyngeus. The valve of the esophageal pump is the
LES. Failure of the valves or the pumps leads to abnormali-
ties in swallowing—that is, difficulty in food propulsion from
mouth to stomach—or regurgitation of gastric contents into the
esophagus or pharynx.

Food is taken into the mouth in a variety of bite sizes,
where it is broken up, mixed with saliva, and lubricated. Once
initiated, swallowing is entirely a reflex act. When food is
ready for swallowing, the tongue, acting like a piston, moves
the bolus into the posterior oropharynx and forces it into the
hypopharynx (Fig. 25-12). Concomitantly with the posterior
movement of the tongue, the soft palate is elevated, thereby
closing the passage between the oropharynx and nasopharynx.
This partitioning prevents pressure generated in the oropharynx
from being dissipated through the nose. When the soft palate is
paralyzed, for example, after a cerebrovascular accident, food
is commonly regurgitated into the nasopharynx. During swal-
loving, the hyoid bone moves upward and anteriorly, elevating
the larynx and opening the retroesophageal space, bringing the
epiglottis under the tongue (see Fig. 25-12). The backward
tilt of the epiglottis covers the opening of the larynx to prevent
aspiration. The entire pharyngeal part of swallowing occurs within
1.5 seconds.

During swallowing, the pressure in the hypopharynx rises
abruptly, to at least 60 mmHg, due to the backward movement
of the tongue and contraction of the posterior pharyngeal con-
strictors. A sizable pressure difference develops between the
hypopharyngeal pressure and the less-than-atmospheric mides-
ophageal or intrathoracic pressure (Fig. 25-13). This pressure
gradient speeds the movement of food from the hypopharynx
into the esophagus when the cricopharyngeus or upper esopha-
geal sphincter relaxes. The bolus is both propelled by peristaltic
contraction of the posterior pharyngeal constrictors and sucked
into the thoracic esophagus. Critical to receiving the bolus is
the compliance of the cervical esophagus; when compliance is
lost due to muscle pathology, dysphagia can result. The upper
esophageal sphincter closes within 0.5 seconds of the initiation
of the swallow, with the immediate closing pressure reaching

Figure 25-12. Sequence of events during the oropharyngeal phase
of swallowing. (Reproduced with permission from Zuidema GD,
Orringer MB: Shackelford’s Surgery of the Alimentary Tract, 3rd ed.

Figure 25-13. Resting pressure profile of the foregut showing the
pressure differential between the atmospheric pharyngeal pressure
(P) and the less-than-atmospheric midesophageal pressure (E) and
greater-than-atmospheric intragastric pressure (G), with the inter-
posed high-pressure zones of the cricopharyngeus (C) and distal
esophageal sphincter (DES). The necessity for relaxation of the cri-
copharyngeus and DES pressure to move a bolus into the stomach
is apparent. Esophageal work occurs when a bolus is pushed from
the midesophageal area (E), with a pressure less than atmospheric,
into the stomach, which has a pressure greater than atmospheric
(G). (Reproduced with permission from Waters PF, DeMeester TR:
Foregut motor disorders and their surgical management, Med Clin
approximately twice the resting level of 30 mmHg. The postrelaxation contraction continues down the esophagus as a peristaltic wave (Fig. 25-14). The high closing pressure and the initiation of the peristaltic wave prevents reflux of the bolus from the esophagus back into the pharynx. After the peristaltic wave has passed farther down the esophagus, the pressure in the upper esophageal sphincter returns to its resting level.

Swallowing can be started at will, or it can be reflexively elicited by the stimulation of areas in the mouth and pharynx, among them the anterior and posterior tonsillar pillars or the posterior lateral walls of the hypopharynx. The afferent sensory nerves of the pharynx are the glossopharyngeal nerves and the superior laryngeal branches of the vagus nerves. Once aroused by stimuli entering via these nerves, the swallowing center in the medulla coordinates the complete act of swallowing. Progress of the wave in the esophagus is caused by sequential activation of its muscles, initiated by efferent vagal nerve fibers arising in the swallowing center.

Continuity of the esophageal muscle is not necessary for sequential activation if the nerves are intact. If the muscles, but not the nerves, are cut across, the pressure wave begins distally below the cut as it dies out at the proximal end above the cut. This allows a sleeve resection of the esophagus to be done without destroying its normal function. Afferent impulses from receptors within the esophageal wall are not essential for progression of the coordinated wave. Afferent nerves, however, do go to the swallowing center from the esophagus because if the esophagus is distended at any point, a contraction wave begins with a forceful closure of the upper esophageal sphincter and sweeps down the esophagus. This secondary contraction occurs without any movements of the mouth or pharynx. Secondary peristalsis can occur as an independent local reflex to clear the esophagus of ingested material left behind after the passage of the primary wave. Current studies suggest that secondary peristalsis is not as common as once thought.

Despite the powerful occlusive pressure, the propulsive force of the esophagus is relatively feeble. If a subject attempts to swallow a bolus attached by a string to a counterweight, the maximum weight that can be overcome is 5 to 10 g. Orderly contractions of the muscular wall and anchoring of the esophagus at its inferior end are necessary for efficient aboral propulsion to occur. Loss of the inferior anchor, as occurs with a large hiatal hernia, can lead to inefficient propulsion.

The LES provides a pressure barrier between the esophagus and stomach and acts as the valve on the worm-drive pump of the esophageal body. Although an anatomically distinct LES has been difficult to identify, microdissection studies show that, in humans, the sphincter-like function is related to the

---

**Figure 25-14.** Intraluminal esophageal pressures in response to swallowing. (Reproduced with permission from Waters PF, DeMeester TR: Foregut motor disorders and their surgical management, Med Clin North Am. 1981 Nov;65(6):1235-1268.)
architecture of the muscle fibers at the junction of the esophageal tube with the gastric pouch (Fig. 25-15). The sphincter actively remains closed to prevent reflux of gastric contents into the esophagus and opens by a relaxation that coincides with a pharyngeal swallow (see Fig. 25-14). The LES pressure returns to its resting level after the peristaltic wave has passed through the esophagus. Consequently, reflux of gastric juice that may occur through the open valve during a swallow is cleared back into the stomach.

If the pharyngeal swallow does not initiate a peristaltic contraction, then the coincident relaxation of the LES is unguarded and reflux of gastric juice can occur. This may be an explanation for the observation of spontaneous lower esophageal relaxation, thought by some to be a causative factor in gastroesophageal reflux disease (GERD). The power of the worm-drive pump of the esophageal body is insufficient to force open a valve that does not relax. In dogs, a bilateral cervical parasympathetic blockade abolishes the relaxation of the LES that occurs with pharyngeal swallowing or distention of the esophagus. Consequently, vagal function appears to be important in coordinating the relaxation of the LES with esophageal contraction.

The antireflux mechanism in human beings is composed of three components: a mechanically effective LES, efficient esophageal clearance, and an adequately functioning gastric reservoir. A defect of any one of these three components can lead to increased esophageal exposure to gastric juice and the development of mucosal injury.

**Physiologic Reflux**

On 24-hour esophageal pH monitoring, healthy individuals have occasional episodes of gastroesophageal reflux. This physiologic reflux is more common when awake and in the upright position than during sleep in the supine position. When reflux of gastric juice occurs, normal subjects rapidly clear the acid gastric juice from the esophagus regardless of their position.

There are several explanations for the observation that physiologic reflux in normal subjects is more common when they are awake and in the upright position than during sleep in the supine position. First, reflux episodes occur in healthy volunteers primarily during transient losses of the gastroesophageal barrier, which may be due to a relaxation of the LES or intragastric pressure overcoming sphincter pressure. Gastric juice can also reflux when a swallow-induced relaxation of the LES is not protected by an oncoming peristaltic wave. The average frequency of these “unguarded moments” or of transient losses of the gastroesophageal barrier is far less while asleep and in the supine position than while awake and in the upright position. Consequently, there are fewer opportunities for reflux to occur in the supine position. Second, in the upright position, there is a 12-mmHg pressure gradient between the resting, positive intra-abdominal pressure measured in the stomach and the most negative intrathoracic pressure measured in the esophagus at midthoracic level. This gradient favors the flow of gastric juice up into the thoracic esophagus when upright. The gradient diminishes in the supine position. Third, the LES pressure in normal subjects is significantly higher in the supine position than in the upright position. This is due to the apposition of the hydrostatic pressure of the abdomen to the abdominal portion of the sphincter when supine. In the upright position, the abdominal pressure surrounding the sphincter is negative compared with atmospheric pressure, and, as expected, the abdominal pressure gradually increases the more caudally it is measured. This pressure gradient tends to move the gastric contents toward the cardia and encourages the occurrence of reflux into the esophagus when the individual is upright. In contrast, in the supine position, the gastroesophageal pressure gradient diminishes, and the abdominal hydrostatic pressure under the diaphragm increases, causing an increase in sphincter pressure and a more competent cardia.

The LES has intrinsic myogenic tone, which is modulated by neural and hormonal mechanisms. \(\alpha\)-Adrenergic neurotransmitters or \(\beta\)-blockers stimulate the LES, and \(\alpha\)-blockers and \(\beta\)-stimulants decrease its pressure. It is not clear to what extent cholinergic nerve activity controls LES pressure. The vagus nerve carries both excitatory and inhibitory fibers to the esophagus and sphincter. The hormones gastrin and motilin have been shown to increase LES pressure; and cholecystokinin, estrogen, glucagon, progesterone, somatostatin, and secretin decrease LES pressure. The peptides bombesin, l-enkephalin, and substance P increase LES pressure; and calcitonin gene-related peptide, gastric inhibitory peptide, neuropeptide Y, and vasoactive intestinal polypeptide decrease LES pressure. Some pharmacologic agents such as antacids, cholinergics, agonists, domperidone, metoclopramide, and prostaglandin \(E_2\) are known to increase LES pressure; and anticholinergics, barbiturates, calcium channel blockers, caffeine, diazepam, dopamine, meperidine, prostaglandin \(E_1\) and \(E_2\), and theophylline decrease LES pressure. Peppermint, chocolate, coffee, ethanol, and fat are all associated with decreased LES pressure and may be responsible for esophageal symptoms after a sumptuous meal.
ASSESSMENT OF ESOPHAGEAL FUNCTION

A thorough understanding of the patient’s underlying anatomic and functional deficits before making therapeutic decisions is fundamental to the successful treatment of esophageal disease. The diagnostic tests, as presently used, may be divided into four broad groups: (a) tests to detect structural abnormalities of the esophagus; (b) tests to detect functional abnormalities of the esophagus; (c) tests to detect increased esophageal exposure to gastric juice; and (d) tests of duodenogastric function as they relate to esophageal disease.

Tests to Detect Structural Abnormalities

Endoscopic Evaluation. The first diagnostic test in patients with suspected esophageal disease is usually upper gastrointestinal endoscopy. This allows assessment and biopsy of the mucosa of the stomach and the esophagus, as well as the diagnosis and assessment of obstructing lesions in the upper gastrointestinal tract. In any patient complaining of dysphagia, esophagoscopy is indicated, even in the face of a normal radiographic study.

For the initial endoscopic assessment, the flexible fiberoptic esophagoscope is the instrument of choice because of its technical ease, patient acceptance, and the ability to simultaneously assess the stomach and duodenum. Rigid endoscopy is now only rarely required, mainly for the disimpaction of difficult foreign bodies impacted in the esophagus, and few individuals now have the skill set and experience to use this equipment.

When GERD is the suspected diagnosis, particular attention should be paid to detecting the presence of esophagitis and Barrett’s columnar-lined esophagus (CLE). When endoscopic esophagitis is seen, severity and the length of esophagitis involved are recorded. Whilst many different grading systems have been proposed, the commonest system now in use is the Los Angeles (LA) grading system. In this system, mild esophagitis is classified LA grade A or B—one or more erosions limited to the mucosal fold(s) and either less than or greater than 5 mm in longitudinal extent respectively (Fig. 25-16). More severe esophagitis is classified LA grade C or D. In grade C, erosions extend over the mucosal folds but over less than three-quarters of the esophageal circumference; in grade D, confluent erosions extend across more than three-quarters of the esophageal circumference. In addition to these grades, more severe damage can lead to the formation of a stricture. A stricture’s severity can be assessed by the ease of passing a standard endoscope. When a stricture is observed, the severity of the esophagitis above it should be recorded. The absence of esophagitis above a stricture suggests the possibility of a chemical-induced injury or a neoplasm as a cause. The latter should always be considered and is ruled out only by evaluation of a tissue biopsy of adequate size. It should be remembered that gastroesophageal reflux is not always associated with visible mucosal abnormalities, and patients can experience significant reflux symptoms, despite an apparently normal endoscopy examination.

Barrett’s esophagus (BE) is a condition in which the tubular esophagus is lined with columnar epithelium, as opposed to the normal squamous epithelium (see Fig. 25-16). Histologically, it appears as intestinal metaplasia (IM). It is suspected at endoscopy when there is difficulty in visualizing the squamocolumnar junction at its normal location, and by the appearance of a redder, salmon-colored mucosa in the lower esophagus, with a clearly visible line of demarcation at the top of the Barrett’s esophagus segment. Its presence is confirmed by biopsy. Multiple biopsy specimens should be taken in a cephalad direction to confirm the presence of IM, and to evaluate the Barrett’s epithelium for dysplastic changes. BE is susceptible to ulceration, bleeding, stricture formation, and, most important, malignant degeneration. The earliest sign of the latter is high grade dysplasia or intramucosal adenocarcinoma (see Fig. 25-16). These dysplastic changes have a patchy distribution, so a minimum of four biopsy samples spaced 2 cm apart should be taken from the Barrett’s-lined portion of the esophagus. Changes seen in one biopsy are significant. Nishimaki has determined that the tumors occur in an area of specialized columnar epithelium near the squamocolumnar junction in 85% of patients, and within 2 cm of the squamocolumnar junction in virtually all patients. Particular attention should be focused on this area in patients suspected of harboring a carcinoma.

Abnormalities of the gastroesophageal flap valve can be visualized by retroflexion of the endoscope. Hill has graded the appearance of the gastroesophageal valve from I to IV according to the degree of unfolding or deterioration of the normal valve architecture (Fig. 25-17). The appearance of the valve correlates with the presence of increased esophageal acid exposure, occurring predominantly in patients with grade III and IV valves.

A hiatal hernia is endoscopically confirmed by finding a pouch lined with gastric rugal folds lying 2 cm or more above the margins of the diaphragmatic crura, identified by having the patient sniff. A hernia is best demonstrated with the stomach fully insufflated and the gastroesophageal junction observed with a retroflexed endoscope. A prominent sliding hiatal hernia is frequently associated with increased esophageal exposure to gastric juice. When a paraesophageal hernia (PEH) is observed, particular attention is taken to exclude gastric (Cameron’s) ulcers or gastritis within the pouch. The intragastric retroflex or J maneuver is important in evaluating the full circumference of the mucosal lining of the herniated stomach.

When an esophageal diverticulum is seen, it should be carefully explored with the flexible endoscope to exclude ulceration or neoplasia. When a submucosal mass is identified, biopsy specimens are usually not performed. At the time of surgical resection, a submucosal leiomyoma or reduplication cyst can generally be dissected away from the intact mucosa, but if a biopsy sample is taken, the mucosa may become fixed to the underlying abnormality. This complicates the surgical dissection by increasing the risk of mucosal perforation. Endoscopic ultrasound provides a better method for evaluating these lesions.

Radiographic Evaluation. Barium swallow evaluation is undertaken selectively to assess anatomy and motility. The anatomy of large hiatal hernias is more clearly demonstrated by contrast radiology than endoscopy, and the presence of coordinated esophageal peristalsis can be determined by observing several individual swallows of barium traversing the entire length of the organ, with the patient in the horizontal position. Hiatal hernias are best demonstrated with the patient prone because the increased intra-abdominal pressure produced in this position promotes displacement of the esophagogastric junction above the diaphragm. To detect lower esophageal narrowing, such as rings and strictures, fully distended views of the esophagogastric region are crucial. The density of the barium used to study the esophagus can potentially affect the accuracy of the examination. Esophageal disorders shown clearly by a full-column technique include circumferential carcinomas, peptic strictures, large esophageal ulcers, and hiatal hernias. A small hiatal hernia is usually not associated with significant symptoms or illness, and its presence is an irrelevant finding unless the hiatal hernia is large (Fig. 25-18) or the hernia
Lesions extrinsic but adjacent to the esophagus can be reliably detected by the full-column technique if they contact the distended esophageal wall. Conversely, a number of important disorders may go undetected if this is the sole technique used to examine the esophagus. These include small esophageal neoplasms, mild esophagitis, and esophageal varices. Thus, the full-column technique should be supplemented with mucosal relief or double-contrast films to enhance detection of these smaller or more subtle lesions.

Motion-recording techniques greatly aid in evaluating functional disorders of the pharyngoesophageal and esophageal phases of swallowing. The technique and indications for cine- and videoradiography will be discussed in the section entitled “Video- and Cineradiography,” as they are more useful to evaluate function and seldom used to detect structural abnormalities.

The radiographic assessment of the esophagus is not complete unless the entire stomach and duodenum have been examined. A gastric or duodenal ulcer, partially obstructing gastric neoplasm, or scarred duodenum and pylorus may contribute significantly to symptoms otherwise attributable to an esophageal abnormality.

When a patient’s complaints include dysphagia and no obstructing lesion is seen on the barium swallow, it is useful to have the patient swallow a barium-impregnated marshmallow, a barium-soaked piece of bread, or a hamburger mixed with barium. This test may bring out a functional disturbance in esophageal transport that can be missed when liquid barium is used.

Tests to Detect Functional Abnormalities

In many patients with symptoms of an esophageal disorder, standard radiographic and endoscopic evaluation fails to demonstrate a structural abnormality. In these situations, esophageal function tests are necessary to identify a functional disorder.

Esophageal Motility. Esophageal motility is a widely used technique to examine the motor function of the esophagus and
Figure 25-17. A. Grade I flap valve appearance. Note the ridge of tissue that is closely approximated to the shaft of the retroflexed endoscope. It extends 3 to 4 cm along the lesser curve. B. Grade II flap valve appearance. The ridge is slightly less well defined than in grade I and it opens rarely with respiration and closes promptly. C. Grade III flap valve appearance. The ridge is barely present, and there is often failure to close around the endoscope. It is nearly always accompanied by a hiatal hernia. D. Grade IV flap valve appearance. There is no muscular ridge at all. The gastroesophageal valve stays open all the time, and squamous epithelium can often be seen from the retroflexed position. A hiatal hernia is always present. (Reproduced with permission from Hill LD, Kozarek RA, Kraemer SJ, et al: The gastroesophageal flap valve: in vitro and in vivo observations, Gastrointest Endosc. 1996 Nov;44(5):541-547.)
its sphincters. The esophageal motility study (EMS) is indicated whenever a motor abnormality of the esophagus is suspected on the basis of complaints of dysphagia, odynophagia, or noncardiac chest pain, and the barium swallow or endoscopy does not show a clear structural abnormality. EMS is particularly necessary to confirm the diagnosis of specific primary esophageal motility disorders (i.e., achalasia, diffuse esophageal spasm [DES], nutcracker esophagus, and hypertensive LES). It also identifies nonspecific esophageal motility abnormalities and motility disorders secondary to systemic disease such as scleroderma, dermatomyositis, polymyositis, or mixed connective tissue disease. In patients with symptomatic GERD, manometry of the esophageal body can identify a mechanically defective LES and evaluate the adequacy of esophageal peristalsis and contraction amplitude. EMS has become an essential tool in the preoperative evaluation of patients before antireflux surgery, guiding selection of the appropriate procedure based upon the patient’s underlying esophageal function and excluding patients with achalasia who can be misdiagnosed with gastroesophageal reflux when clinical and endoscopic parameters alone are used for diagnosis.

EMS is performed using electronic, pressure-sensitive transducers located within the catheter, or water-perfused catheters with lateral side holes attached to transducers outside the body. The traditional water perfused catheter has largely been replaced by high resolution motility (HRM), but knowledge of traditional methods of assessing esophageal motility is helpful for understanding esophageal physiology.

As the pressure-sensitive station is brought across the gastroesophageal junction (GEJ), a rise in pressure above the gastric baseline signals the beginning of the LES. The respiratory inversion point is identified when the positive excursions that occur in the abdominal cavity with breathing change to negative deflections in the thorax. The respiratory inversion point serves as a reference point at which the amplitude of LES pressure and the length of the sphincter exposed to abdominal pressure are measured. As the pressure-sensitive station is withdrawn into the body of the esophagus, the upper border of the LES is identified by the drop in pressure to the esophageal baseline. From these measurements, the pressure, abdominal length, and overall length of the sphincter are determined (Fig. 25-19). To

---

Figure 25-17. (Continued)

Figure 25-18. Radiogram of an intrathoracic stomach. This is the end stage of a large hiatal hernia, regardless of its initial classification.

account for the asymmetry of the sphincter (Fig. 25-20), the pressure profile is repeated with each of the five radially oriented transducers, and the average values for sphincter pressure above gastric baseline, overall sphincter length, and abdominal length of the sphincter are calculated.

Table 25-1 shows the values for these parameters in 50 normal volunteers without subjective or objective evidence of a foregut disorder. A mechanically defective sphincter is identified by having one or more of the following characteristics: an average LES pressure of <6 mmHg, an average length exposed to the positive-pressure environment in the abdomen of 1 cm or less, and/or an average overall sphincter length of 2 cm or less.

**High-Resolution Manometry.** Esophageal manometry was introduced into clinical practice in the 1970s and, until recently, has changed little. In 1991, Ray Clouse introduced the concept of improving conventional manometry by increasing the number of recording sites and adding a three-dimensional assessment. This “high-resolution manometry” is a variant of the conventional manometry in which multiple, circumferential recording sites are used, in essence creating a “map” of the esophagus and its sphincters. High-resolution catheters contain 36 miniaturized pressure sensors positioned every centimeter along the length of the catheter. The vast amount of data generated by these sensors is then processed and presented in traditional linear plots or as a visually enhanced spatiotemporal video tracing that is readily interpreted. The function of the esophageal body is assessed with 10 to 15 wet swallows. Amplitude, duration, and morphology of contractions following each swallow are visually displayed (Fig. 25-21).

The relationship of the esophageal contractions following a swallow is classified as peristaltic or simultaneous. The data are used to identify motor disorders of the esophagus.

The position, length, and function of the lower esophageal sphincter (LES) are demonstrated by a high-pressure zone that should relax at the inception of swallowing and contract after the water or solid bolus passes through the LES. Simultaneous acquisition of data for the upper esophageal sphincter, esophageal body, LES, and gastric pressure minimizes the movement artifacts and study time associated with conventional esophageal manometry. This technology significantly enhances esophageal diagnostics, bringing it into the realm of “image”-based studies. High-resolution manometry may allow the identification of focal motor abnormalities previously overlooked. It has enhanced the ability to predict bolus propagation and increased sensitivity in the measurement of pressure gradients.

**Esophageal Impedance.** Newer technology introduced into the clinical realm a decade ago allows measurement of esophageal function and gastroesophageal reflux in a way that was previously not possible. An intraluminal electrical impedance catheter is used to measure GI function. Impedance is the ratio of voltage to current, and is a measure of the electrical conductivity of a hollow organ and its contents. Intraluminal electrical impedance is inversely proportional to the electrical conductivity of the luminal contents and the cross-sectional area of the lumen. Air has a very low electrical conductivity and, therefore, high impedance. Saliva and food cause an impedance decrease because of their increased conductivity. Luminal dilatation results in a decrease in impedance, whereas luminal contraction yields an impedance increase. Investigators have established the impedance waveform characteristics that define esophageal bolus transport. This allows for the characterization of both esophageal function, via quantification of bolus transport, and gastroesophageal reflux (Fig. 25-22). The probe measures impedance between adjacent electrodes, with measuring segments located at 2, 4, 6, 8, 14, and 16 cm from the distal tip. An extremely low electric current of 0.00025 μW is transmitted across the electrodes at a frequency of 1 to 2 kHz and is limited.
Figure 25-21A. Normal high-resolution manometry motility study. Pressure measurements are recorded with color coding (red = high; blue = low). LES = lower esophageal sphincter; PIP = pressure inversion point; UES = upper esophageal sphincter.
Figure 25-21B. High-resolution manometry motility study in patient with mechanically defective lower esophageal sphincter. Note the absence of lower esophageal sphincter tone. Pressure measurements are recorded with color coding (red = high; blue = low). LES = lower esophageal sphincter; PIP = pressure inversion point; UES = upper esophageal sphincter.
Figure 25-21C. High-resolution manometry motility study in patient with deficient esophageal body peristalsis. Note the very weak peristalsis in the lower two-thirds of the esophagus. Pressure measurements are recorded with color coding (red = high; blue = low). LES = lower esophageal sphincter; PIP = pressure inversion point; UES = upper esophageal sphincter.
Figure 25-21D. High-resolution manometry motility study in patient with achalasia. Note the complete absence of esophageal body peristalsis, and the lack of relaxation of the lower esophageal sphincter. Pressure measurements are recorded with color coding (red = high; blue = low). LES = lower esophageal sphincter; PIP = pressure inversion point; UES = upper esophageal sphincter.
Figure 25-21E. High-resolution manometry motility study in patient with diffuse esophageal spasm. Note the very high amplitude contractions in the esophageal body. Pressure measurements are recorded with color coding (red = high; blue = low). LES = lower esophageal sphincter; PIP = pressure inversion point; UES = upper esophageal sphincter.
to 8 μA. This is below the stimulation threshold for nerves and muscles and is three orders of magnitude below the threshold of cardiac stimulation. A standard pH electrode is located 5 cm from the distal tip so that the acidic or nonacidic nature of refluxate can be correlated with the number of reflux events.

Esophageal impedance has been validated as an appropriate method for the evaluation of gastrointestinal function and is used selectively for the diagnosis of gastroesophageal reflux. It has been compared to cineradiography showing that impedance waves correspond well with actual bolus transport illustrated by radiography. Bolus entry, transit, and exit can be clearly identified by impedance changes in the corresponding measuring segments. Studies comparing standard esophageal manometry with impedance measurements in healthy volunteers have shown that esophageal impedance correlates with peristaltic wave progression and bolus length.

Figure 25-22. Esophageal impedance probe measures electrical resistance between evenly spaced electrodes. LES = lower esophageal sphincter.

Twenty-four-hour pH monitoring, the historical gold standard for diagnosing and quantifying gastroesophageal reflux, has some significant limitations. With 24-hour ambulatory pH testing, reflux is defined as a drop in the pH below 4, which effectively “blinds” the test to reflux occurring at higher pH values. Furthermore, in patients with persistent symptoms on proton pump inhibitor (PPI) therapy, pH monitoring has limited use as it can only detect abnormal acid reflux (pH <4), the occurrence of which has been altered by the antisecretory medication. Given that PPI antisecretory therapy is highly effective in neutralizing gastric acid, the question of whether persistent symptoms are a result of persistent acid reflux, nonacid reflux, or are not reflux related becomes a key issue in surgical decision making. Until recently, this differentiation could not be made. Detection of both acid and nonacid reflux has potential to define these populations of patients and thus improve patient selection for antireflux surgery. Multichannel intraluminal impedance technology allows the measurement of both acid and nonacid reflux, with potential to enhance diagnostic accuracy.

Using this technology, Balaji and colleagues showed that most gastroesophageal reflux remains despite acid suppression. Impedance pH may be particularly useful in evaluating patients with persistent symptoms despite PPI treatment, patients with respiratory symptoms, and postoperative patients who are having symptoms that are elusive to diagnosis.

Esophageal Transit Scintigraphy. The esophageal transit of a 10-mL water bolus containing technetium-99m (99mTc) sulfur colloid can be recorded with a gamma camera. Using this technique, delayed bolus transit has been shown in patients with a variety of esophageal motor disorders, including achalasia, scleroderma, DES, and nutcracker esophagus.

Video- and Cineradiography

High-speed cinematic or video recording of radiographic studies allows re-evaluation by reviewing the studies at various speeds. This technique is more useful than manometry in the evaluation of the pharyngeal phase of swallowing. Observations suggesting oropharyngeal or cricopharyngeal dysfunction include misdirection of barium into the trachea or nasopharynx, prominence of the cricopharyngeal muscle, a Zenker’s diverticulum, a narrow pharyngoesophageal segment, and stasis of the contrast medium in the valleculae or hypopharyngeal recesses (Fig. 25-23). These findings are usually not specific, but rather common manifestations of neuromuscular disorders affecting the pharyngoesophageal area. Studies using liquid barium, barium-impregnated solids, or radiopaque pills aid the evaluation of normal and abnormal motility in the esophageal body. Loss of the normal stripping wave or segmentation of the barium column with the patient in the recumbent position correlates with abnormal motility of the esophageal body. In addition, structural abnormalities such as small diverticula, webs, and minimal extrinsic pressure on the esophagus may be recognized only with motion-recording techniques. The simultaneous capture of videofluoroscopic images and manometric tracings is now available and is referred to as manofluorography. Manofluorographic studies allow precise correlation of the anatomic events, such as opening of the upper esophageal sphincter, with manometric observations, such as sphincter relaxation. Manofluorography, although not widely available, is presently the best means available to evaluate complex functional abnormalities.

Tests to Detect Increased Exposure to Gastric Juice

Twenty-Four-Hour Ambulatory pH Monitoring. The most direct method of measuring increased esophageal exposure to gastric juice is by an indwelling pH electrode, or, more recently, via a radiotelemetric pH monitoring capsule that can be clipped to the esophageal mucosa. The latter consists of an antimony pH electrode fitted inside a small, capsule-shaped device accompanied by a battery and electronics that allow 48-hour monitoring and transmission of the pH data via transcutaneous radio telemetry to a waist-mounted data logger. The device can be introduced either transorally or transnasally, and it can be clipped to the esophageal mucosa using endoscopic fastening techniques. It passes spontaneously within 1 to 2 weeks. Prolonged monitoring of esophageal pH is performed by placing the pH probe or telemetry capsule 5 cm above the manometrically measured upper border of the distal sphincter for 24 hours. It measures the actual time the esophageal mucosa is exposed to gastric juice, measures the ability of the esophagus to clear refluxed acid, and correlates esophageal acid exposure with the patient’s symptoms. A 24- to 48-hour period is necessary so that measurements can be made over one or two complete circadian cycles. This allows measuring the effect of physiologic activity, such as eating or sleeping, on the reflux of gastric juice into the esophagus (Fig. 25-24).
The 24-hour esophageal pH monitoring should not be considered a test for reflux, but rather a measurement of the esophageal exposure to gastric juice. The measurement is expressed by the time the esophageal pH was below a given threshold during the 24-hour period (Table 25-3). This single assessment, although concise, does not reflect how the exposure has occurred; that is, did it occur in a few long episodes or several short episodes? Consequently, two other assessments are necessary: the frequency of the reflux episodes and their duration.

The units used to express esophageal exposure to gastric juice are: (a) cumulative time the esophageal pH is below a chosen threshold, expressed as the percentage of the total, upright, and supine monitored time; (b) frequency of reflux episodes below a chosen threshold, expressed as number of episodes per 24 hours; and (c) duration of the episodes, expressed as the number of episodes >5 minutes per 24 hours, and the time in minutes of the longest episode recorded. Table 25-2 shows the normal values for these components of the 24-hour record at the whole-number pH threshold derived from 50 normal asymptomatic subjects. The upper limits of normal were established at the 95th percentile. Most centers use pH 4 as the threshold.

Based on these studies and extensive clinical experience, 48-hour esophageal pH monitoring is considered to be the gold standard for the diagnosis of GERD.

The Bravo pH Capsule (Medtronics, Minneapolis, MN) measures pH levels in the esophagus and transmits continuous

### Figure 25-23. Esophagograms from a patient with cricopharyngeal achalasia. A. Anteroposterior film showing retention of the contrast medium at the level of the vallecula and piriform recesses, with no barium passing into the esophagus. B. Lateral film, taken opposite the C5–C6 vertebrae, showing posterior indentation of the cricopharyngeus, retention in the hypopharynx, and tracheal aspiration. (Reproduced with permission from DeMeester TR, Matthews H: International Trends in General Thoracic Surgery. Vol 3. Benign Esophageal Disease. St. Louis, Mo: Mosby; 1987.)

### Table 25-2

<table>
<thead>
<tr>
<th>COMPONENT</th>
<th>MEAN</th>
<th>SD</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total time</td>
<td>1.51</td>
<td>1.36</td>
<td>4.45</td>
</tr>
<tr>
<td>Upright time</td>
<td>2.34</td>
<td>2.34</td>
<td>8.42</td>
</tr>
<tr>
<td>Supine time</td>
<td>0.63</td>
<td>1.0</td>
<td>3.45</td>
</tr>
<tr>
<td>No. of episodes</td>
<td>19.00</td>
<td>12.76</td>
<td>46.90</td>
</tr>
<tr>
<td>No. &gt;5 min</td>
<td>0.84</td>
<td>1.18</td>
<td>3.45</td>
</tr>
<tr>
<td>Longest episode</td>
<td>6.74</td>
<td>7.85</td>
<td>19.80</td>
</tr>
</tbody>
</table>

SD = standard deviation.

esophageal pH readings to a receiver worn on the patient’s belt or waistband (Fig. 25-25). Symptoms that the patient experiences are recorded in a diary and/or by pressing buttons on the receiver unit. Generally, 48 hours of pH data are measured with this probe. A recent study has shown that the addition of a second day of pH monitoring increased the sensitivity of pH measurement by 22%. The capsule eventually detaches and passes through the digestive tract in 5 to 7 days.

**Radiographic Detection of Gastroesophageal Reflux.** The definition of radiographic gastroesophageal reflux varies depending on whether reflux is spontaneous or induced by various maneuvers. In only about 40% of patients with classic symptoms of GERD is spontaneous reflux (i.e., reflux of barium from the stomach into the esophagus with the patient in the upright position) observed by the radiologist. In most patients who show spontaneous reflux on radiography, the diagnosis of increased esophageal acid exposure is confirmed by 24-hour esophageal pH monitoring. Therefore, the radiographic demonstration of spontaneous regurgitation of barium into the esophagus in the upright position is a reliable indicator that reflux is present. However, failure to see this does not indicate the absence of disease, and for this reason this test is rarely used for clinical diagnosis.

**Tests of Duodenogastric Function**

Esophageal disorders are frequently associated with abnormalities of duodenogastric function. Abnormalities of the gastric reservoir or increased gastric acid secretion can be responsible for increased esophageal exposure to gastric juice. Reflux of alkaline duodenal juice, including bile salts, pancreatic enzymes, and bicarbonate, is thought to have a role in the pathogenesis of esophagitis and complicated Barrett’s esophagus. Furthermore, functional disorders of the esophagus are often not confined to

---

**Table 25-3**

<table>
<thead>
<tr>
<th>pH THRESHOLD</th>
<th>95TH PERCENTILE</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1</td>
<td>14.2</td>
</tr>
<tr>
<td>&lt;2</td>
<td>17.37</td>
</tr>
<tr>
<td>&lt;3</td>
<td>14.10</td>
</tr>
<tr>
<td>&lt;4</td>
<td>14.72</td>
</tr>
<tr>
<td>&lt;5</td>
<td>15.76</td>
</tr>
<tr>
<td>&lt;6</td>
<td>12.76</td>
</tr>
<tr>
<td>&gt;7</td>
<td>14.90</td>
</tr>
<tr>
<td>&gt;8</td>
<td>8.50</td>
</tr>
</tbody>
</table>

the esophagus alone, but are associated with functional disorders of the rest of the foregut (i.e., stomach and duodenum). Tests of duodenogastric function that are helpful to investigate esophageal symptoms include gastric emptying studies, gastric acid analysis, and cholescintigraphy (for the diagnosis of pathologic duodenogastric and/or duodenogastroesophageal reflux).

**Gastric Emptying Study.** Gastric emptying studies are performed with radionuclide-labeled meals. Emptying of solids and liquids can be assessed simultaneously when both phases are marked with different tracers. After ingestion of a labeled standard meal, gamma camera images of the stomach are obtained at 5- to 15-minute intervals for 2 to 4 hours. After correction for decay, the counts in the gastric area are plotted as the percentage of total counts at the start of the imaging. The resulting emptying curve can be compared with data obtained in normal volunteers. In general, normal subjects will empty 59% of a meal within 90 minutes. Although delayed gastric emptying is often associated with gastroesophageal reflux, in general delayed emptying does not correlate with a poorer clinical outcome after antireflux surgery, and it should not be considered a contraindication to surgical treatment.

**GASTROESOPHAGEAL REFLUX DISEASE**

GERD was not recognized as a significant clinical problem until the mid-1930s and was not identified as a precipitating cause for esophagitis until after World War II. In the early 21st century, it has grown to be a very common problem and now accounts for a majority of esophageal pathology. It is recognized as a chronic disease, and when medical therapy is required, it is often lifelong treatment. Recent efforts at the development of various endoscopic antireflux interventions, although innovative, have not been successful in consistently controlling gastroesophageal reflux. Antireflux surgery is an effective and long-term therapy and is the only treatment that is able to restore the gastroesophageal barrier. Despite the common prevalence of GERD, it can be one of the most challenging diagnostic and therapeutic problems in clinical medicine. A contributing factor to this is the lack of a universally accepted definition of the disease.

The most simplistic approach is to define the disease by its symptoms. However, symptoms thought to be indicative of GERD, such as heartburn or acid regurgitation, are very common in the general population and many individuals consider them to be normal and do not seek medical attention. Even when excessive, these symptoms are not specific for gastroesophageal reflux. They can be caused by other diseases such as achalasia, DES, esophageal carcinoma, pyloric stenosis, cholelithiasis, gastritis, gastric or duodenal ulcer, and coronary artery disease.

A thorough, structured evaluation of the patient’s symptoms is essential before any therapy, particularly any form of esophageal surgery. The presence and severity of both typical symptoms of heartburn, regurgitation, and dysphagia, and atypical symptoms of cough, hoarseness, chest pain, asthma, and aspiration should be discussed with the patient in detail. Many of these atypical symptoms may not be esophageal related and hence will not improve and may even worsen with antireflux surgery.

Heartburn is generally defined as a *subternal burning-type discomfort*, beginning in the epigastrium and radiating upward. It is often aggravated by meals, spicy or fatty foods, chocolate, alcohol, and coffee and can be worse in the supine position. It is commonly, although not universally, relieved by antacid or antisecretory medications. Epidemiologic studies have shown that heartburn occurs monthly in as many as 40% to 50% of the Western population. The occurrence of heartburn at night and its effect on quality of life have recently been highlighted by a Gallup poll conducted by the American Gastroenterologic Society (Table 25-4).

**Table 25-4**

**American Gastroenterologic Association Gallup poll on nighttime gastroesophageal reflux disease symptoms**

- 50 million Americans have nighttime heartburn at least 1/wk
- 80% of heartburn sufferers had nocturnal symptoms—65% both day & night
- 63% report that it affects their ability to sleep and impacts their work the next day
- 72% are on prescription medications
- Nearly half (45%) report that current remedies do not relieve all symptoms

Regurgitation, the effortless return of acid or bitter gastric contents into the chest, pharynx, or mouth, is highly suggestive of foregut pathology. It is often particularly severe at night when supine or when bending over and can be secondary to either an incompetent or obstructed GEJ. With the latter, as in achalasia, the regurgitant is often bland, as if food was put into a blender. When questioned, most patients can distinguish the two. It is the regurgitation of gastric contents that may result in associated pulmonary symptoms, including cough, hoarseness, asthma, and recurrent pneumonia. Bronchospasm can be precipitated by esophageal acidification and cough by either acid stimulation or distention of the esophagus.

Dysphagia, or difficulty swallowing, is a relatively non-specific term but arguably the most specific symptom of foregut disease. It can be a sign of underlying malignancy and should be aggressively investigated until a diagnosis is established. Dysphagia refers to the sensation of difficulty in the passage of food from the mouth to the stomach and can be divided into oropharyngeal and esophageal etiologies. Oropharyngeal dysphagia is characterized by difficulty transferring food out of the mouth into the esophagus, nasal regurgitation, and/or aspiration. Esophageal dysphagia refers to the sensation of food sticking in the lower chest or epigastrium. This may or may not be accompanied by pain (odynophagia) that will be relieved by the passage of the bolus.

Chest pain, although commonly and appropriately attributed to cardiac disease, is frequently secondary to esophageal pathology as well. Nearly 50% of patients with severe chest pain, normal cardiac function, and normal coronary arteriograms have positive 24-hour pH studies, implicating gastroesophageal reflux as the underlying etiology. Exercise-induced gastroesophageal reflux is well known to occur, and may result in exertional chest pain similar to angina. It can be quite difficult, if not impossible, to distinguish between the two etiologies, particularly on clinical grounds alone. Nevens and colleagues evaluated the ability of experienced cardiologists to differentiate pain of cardiac vs. esophageal origin. Of 248 patients initially seen by cardiologists, 185 were thought to have typical angina, and 63 were thought to have atypical chest pain. Forty-eight (26%) of those thought to have classic angina had normal coronary angiograms, and 16 of the 63 with atypical pain had abnormal angiogram. Thus, the cardiologists’ clinical impression was wrong 25% of the time. Finally, Pope and associates investigated the ultimate diagnosis in 10,689 patients presenting to an
emergency department with acute chest pain. Approximately 17% were found to have acute ischemia, 6% had stable angina, 21% had other cardiac causes, and 55% had noncardiac causes. The investigators concluded that the majority of people presenting to the emergency department with chest pain do not have an underlying cardiac etiology for their symptoms. Chest pain precipitated by meals, occurring at night while supine, nonradiating, responsive to antacid medication, or accompanied by other symptoms suggesting esophageal disease such as dysphagia or regurgitation should trigger the thought of possible esophageal origin. Furthermore, the distinction between heartburn and chest pain is also difficult and largely dependent upon the individual patient. One person’s heartburn is another’s chest pain.

The precise mechanisms accounting for the generation of symptoms secondary to esophageal pathology remain unclear. Considerable insight has been acquired, however. Investigations into the effect of luminal content, esophageal distention and muscular function, neural pathways, and brain localization have provided a basic understanding of the stimuli responsible for symptom generation. It is also clear that the visceroneural pathways of the foregut are complexly intertwined with that of the tracheobronchial tree and heart. This fact accounts for the common overlap of clinical presentations with diverse disease processes in upper GI, cardiac, and pulmonary systems.

The Human Antireflux Mechanism and the Pathophysiology of Gastroesophageal Reflux Disease

There is a high-pressure zone located at the esophagogastric junction in humans. Although this is typically referred to as the lower esophageal “sphincter,” there are no distinct anatomical landmarks that define its beginning and end. Architecturally speaking, there is a specialized thickening in this region that is made up of the collar sling musculature and the clasp fibers. The collar sling is located on the greater curvature side of the junction, and the clasp fibers are located on the lesser curvature side. These muscles remain in tonic opposition until the act of swallowing, whereupon receptive relaxation occurs allowing passage of a food bolus into the stomach. In addition, the LES will also open when the gastric fundus is distended with gas and liquid, thus resulting in an unfolding of the valve and enabling venting of gas (a belch). Whether physiologic or pathologic, the common denominator for symptom generation. It is also clear that the visceroneural pathways of the foregut are complexly intertwined with that of the tracheobronchial tree and heart. This fact accounts for the common overlap of clinical presentations with diverse disease processes in upper GI, cardiac, and pulmonary systems.

The Lower Esophageal Sphincter. As defined by esophageal manometry, there are three characteristics of the LES that work in unison to maintain its barrier function. These characteristics include the resting LES pressure, its overall length, and the intra-abdominal length that is exposed to the positive pressure environment of the abdomen (Table 25-5). The resistance to gastroesophageal reflux is a function of both the resting LES pressure and length over which this pressure is exerted. Thus, as the sphincter becomes shorter, a higher pressure will be required in order to prevent a given amount of reflux (Fig. 25-26). Much like the neck of a balloon as it is inflated, as the stomach fills and distends, sphincter length decreases. Therefore, if the overall length of the sphincter is permanently short from repeated distention of the fundus secondary to large volume meals, then with minimal episodes of gastric distention and pressure, there will be insufficient sphincter length for the barrier to maintain competence, and reflux will occur.

Table 25-5

<table>
<thead>
<tr>
<th>PARAMETER</th>
<th>MEDIAN VALUE</th>
<th>2.5TH PERCENTILE</th>
<th>97.5TH PERCENTILE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pressure (mmHg)</td>
<td>13</td>
<td>5.8</td>
<td>27.7</td>
</tr>
<tr>
<td>Overall length (cm)</td>
<td>3.6</td>
<td>2.1</td>
<td>5.6</td>
</tr>
<tr>
<td>Abdominal length (cm)</td>
<td>2</td>
<td>0.9</td>
<td>4.7</td>
</tr>
</tbody>
</table>

A third characteristic of the LES that impacts its ability to prevent reflux is its position about the diaphragm. It is important that a portion of the total length of the LES be exposed to the effects of an intra-abdominal pressure. That is, during periods of elevated intra-abdominal pressure, the resistance of the barrier would be overcome if pressure were not applied equally to both the LES and stomach simultaneously. Thus, in the presence of a hiatal hernia, the sphincter resides entirely within the chest cavity and cannot respond to an increase in intra-abdominal pressure because the pinch valve mechanism is lost and gastroesophageal reflux is more liable to occur.

Therefore, a permanently defective sphincter is defined by one or more of the following characteristics: an LES with a mean resting pressure of less than 6 mmHg, an overall sphincter length of <2 cm, and intra-abdominal sphincter length of <1 cm. Compared to normal subjects without GERD these values are below the 2.5 percentile for each parameter. The most common cause of a defective sphincter is an inadequate abdominal length.

Once the sphincter is permanently defective, this condition is irreversible, and although esophageal mucosal injury may be healed with antisecretory medication, reflux will continue to occur. Additionally, the presence of a defective LES may be associated with reduced esophageal body function and thus decrease clearance times of refluxed material. In addition, the progressive loss of effective esophageal clearance may predispose the patient to severe mucosal injury, volume regurgitation, aspiration, and pulmonary injury. Reflux may occur in the face of a normal LES resting pressure. This condition is usually due to a functional problem of gastric emptying or excessive air swallowing. These conditions may lead to gastric distention, increased intra-gastric pressure, a resultant shortening or

![Figure 25-26. As the esophageal sphincter becomes shorter, increased pressure is necessary to maintain competence. LES = lower esophageal sphincter.](image-url)
unfolding of the LES, and subsequent reflux. The mechanism by which gastric distention contributes to LES unfolding provides a mechanical explanation for “transient LES relaxation.” It is thought that with repeated gastric distention secondary to large meal volume or chronic air swallowing, there is repeated unfolding of the LES and subsequent attenuation of the collar sling musculature. It is at this point that the physiologic and normal mechanism of gastric venting is replaced with pathologic and severe postprandial reflux disease. In addition, patients with GERD will increase the frequency of swallowing in an effort to neutralize the refluxed acid with their saliva (pH 7.0). This phenomenon leads to increased air swallowing and further gastric distention, thus compounding the problem. Therefore, GERD may have its origins in the stomach secondary to gastric distention due to overeating/drinking, air swallowing, or consumption of carbonated liquids, and this may be further compounded by the ingestion of fatty meals, which result in delayed gastric emptying.

**Relationship Between Hiatal Hernia and Gastroesophageal Reflux Disease.** As the collar sling musculature and clasp fibers become attenuated with repeated gastric distention, the esophagogastric junction begins to assume an “upside down funnel” appearance, with progressive opening of the acute angle of His. This in turn may result in attenuation and stretching of the phrenoesophageal ligament, with subsequent enlargement of the hiatal opening and axial herniation. There is a high degree of correlation between reflux threshold and the degree of hiatal herniation (Fig. 25-27).

**Summary.** It is believed that GERD has its origins within the stomach. Distention of the fundus occurs because of overeating and delayed gastric emptying secondary to a high-fat diet. The resultant distention causes “unrolling” of the sphincter by the expanding fundus, and this subsequently exposes the squamous epithelium in the region of the distal LES to gastric juice. Repeated exposure results in inflammation and the development of columnar epithelium at the cardia. This is the initial step of the development of carditis and explains why in early disease esophagitis is mild and commonly limited to the very distal aspect of the esophagus. The patient attempts to compensate for this by increased swallowing, allowing the saliva to neutralize the refluxed gastric juice and thus, alleviate the discomfort induced by the reflux event. The increased swallowing results in aerophagia, bloating, and belching. This in turn creates a vicious cycle of increased gastric distention and thus further exposure and repetitive injury to the distal esophagus. The development of carditis explains the complaint of epigastric pain often experienced by patients with early reflux disease. Additionally, this process can lead to a fibrotic mucosal ring located at the squamocolumnar junction, which is termed a “Schatzki ring” and which may result in dysphagia. This inflammatory process may extend into muscularis propria and thus result in a progressive loss in the length and pressure of the LES. This explanation for the pathophysiology of GERD is supported by the observation that severe esophagitis is almost always associated with a defective LES.

**Complications Associated With Gastroesophageal Reflux Disease**

The complications of gastroesophageal reflux disease may result from the direct injurious effects of gastric fluid on the mucosa, larynx, or respiratory epithelium. Complications due to repetitive reflux are esophagitis, stricture, and BE; repetitive aspiration may lead to progressive pulmonary fibrosis. The severity of the complications is directly related to the prevalence of a structurally defective sphincter (Table 25-6). The observation that a structurally defective sphincter occurs in 42% of patients without complications (most of whom have one or two components failed) suggests that disease may be confined to the sphincter due to compensation by a vigorously contracting esophageal body. Eventually, all three components of the sphincter fail, allowing unrestricted reflux of gastric juice into the esophagus and overwhelming its normal clearance mechanisms. This leads to esophageal mucosal injury with progressive deterioration of esophageal contractility, as is commonly seen in patients with strictures and BE. The loss of esophageal clearance increases the potential for regurgitation into the pharynx with aspiration.

![Figure 25-27. Yield pressure of the lower esophageal sphincter decreases as hiatal hernia size increases.](image)

**Table 25-6**

<table>
<thead>
<tr>
<th>Complication</th>
<th>No.</th>
<th>Structurally Normal Sphincter (%)</th>
<th>Structurally Defective Sphincter (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>59</td>
<td>58</td>
<td>42</td>
</tr>
<tr>
<td>Erosive esophagitis</td>
<td>47</td>
<td>23</td>
<td>77a</td>
</tr>
<tr>
<td>Stricture</td>
<td>19</td>
<td>11</td>
<td>89</td>
</tr>
<tr>
<td>Barrett’s esophagus</td>
<td>25</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>150</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Grade more severe with defective cardia.*

The potential injurious components that reflux into the esophagus include gastric secretions such as acid and pepsin, as well as biliary and pancreatic secretions that regurgitate from the duodenum into the stomach. There is a considerable body of experimental evidence to indicate that maximal epithelial injury occurs during exposure to bile salts combined with acid and pepsin. These studies have shown that while acid alone does minimal damage to the esophageal mucosa, the combination of acid and pepsin is highly deleterious. Similarly, the reflux of duodenal juice alone does little damage to the mucosa, although the combination of duodenal juice and gastric acid is particularly noxious.

Complications of gastroesophageal reflux such as esophagitis, stricture, and Barrett’s metaplasia occur in the presence of two predisposing factors: a mechanically defective LES and an increased esophageal exposure to fluid containing duodenal content that includes bile and pancreatic juice. The duodenal origin of esophageal contents in patients with an increased exposure to a pH >7 has previously been confirmed by esophageal aspiration studies (Fig. 25-28). Studies have clarified and expanded these observations by measuring esophageal bilirubin exposure over a 24-hour period as a marker for the presence of duodenal juice. Direct measurement of esophageal bilirubin exposure as a marker for duodenal juice has shown that 58% of patients with GERD have increased esophageal exposure to duodenal juice and that this exposure occurs most commonly when the esophageal pH is between 4 and 7 (Fig. 25-29). These earlier studies have been confirmed by other studies that measure volume reflux using impedance technology (Fig. 25-30).

If reflux of gastric juice is allowed to persist and sustained or repetitive esophageal injury occurs, two sequelae can result. First, a luminal stricture can develop from submucosal and eventually intramural fibrosis. Second, the tubular esophagus may become replaced with columnar epithelium. The columnar epithelium is resistant to acid and is associated with the alleviation of the complaint of heartburn. This columnar epithelium often becomes intestinalized, identified histologically by the presence

**Figure 25-28.** Sample bile acid concentration and esophageal pH plotted against time to obtain detailed profiles; in this case showing both significant bile acid (vertical bars) and acid (linear plot) reflux. (Reproduced with permission from Nehra D, Watt P, Pye JK, et al. Automated oesophageal reflux sampler: a new device used to monitor bile acid reflux in patients with gastroesophageal reflux disease, J Med Eng Technol. 1997 Jan-Feb;21(1):1-9.)


**Figure 25-30.** Prevalence of abnormal esophageal bilirubin exposure in healthy subjects and in patients with gastroesophageal reflux disease with varied degrees of mucosal injury. (*P <.03 vs. all other groups; **P <.03 vs. healthy subjects.) (Reproduced with permission from Kauer WK, Peters JH, DeMeester TR, et al: Mixed reflux of gastric and duodenal juices is more harmful to the esophagus than gastric juice alone. The need for surgical therapy re-emphasized, Ann Surg. 1995 Oct;222(4):525-531.)
of goblet cells. This specialized IM is currently required for the diagnosis of BE. Endoscopically, BE can be quiescent or associated with complications of esophagitis, stricture, Barrett’s ulceration, and dysplasia. The complications associated with BE may be due to the continuous irritation from refluxed duodenogastric juice. This continued injury is pH dependent and may be modified by medical therapy. The incidence of metaplastic Barrett’s epithelium becoming dysplastic and progressing to adenocarcinoma is approximately 0.2% to 0.5% per year.

An esophageal stricture can be associated with severe esophagitis or BE. In the latter situation, it occurs at the site of maximal inflammatory injury (i.e., the columnar-squamous epithelial interface). Patients who have a stricture in the absence of Barrett’s esophagus should have the presence of gastroesophageal reflux documented before the presence of the stricture is ascribed to reflux esophagitis. In patients with normal acid exposure and no endoscopic or CT evidence of cancer, the stricture may be a result of a drug-induced chemical injury, the latter resulting from the lodgment of a capsule or tablet in the distal esophagus. In such patients, dilation usually corrects the problem of dysphagia. It is also possible for drug-induced injuries to occur in patients who have underlying esophagitis and a distal esophageal stricture secondary to gastroesophageal reflux. In this situation, a long, string-like stricture progressively develops as a result of repetitive caustic injury from capsule or tablet lodgment on top of an initial reflux stricture. These strictures are often resistant to dilation. The incidence of this problem has lessened since the introduction of proton pump inhibitor medication.

**Metaplastic (Barrett’s Esophagus) and Neoplastic (Adenocarcinoma) Complications**

The condition whereby the tubular esophagus is lined with columnar epithelium rather than squamous epithelium was first described by Norman Barrett in 1950. He incorrectly believed it to be congenital in origin. It is now realized that it is an acquired abnormality, occurs in 10% to 15% of patients with GERD, and represents the end stage of the natural history of this disease. It is also distinctly different from the congenital condition in which islands of gastric fundic epithelium are found in the upper half of the esophagus.

The definition of BE has evolved considerably over the past decade. Traditionally, BE was identified by the presence of columnar mucosa extending at least 3 cm into the esophagus. It is now recognized that the specialized, intestinal-type epithelium, or intestinal metaplasia (IM) found in the Barrett’s mucosa, is the only tissue predisposed to malignant degeneration. Consequently, the diagnosis of BE is presently made given any length of endoscopically identifiable columnar mucosa that proves, on biopsy, to show IM. Although long segments of columnar mucosa without IM do occur, they are uncommon and might be congenital in origin.

The hallmark of IM is the presence of intestinal goblet cells. There is a high prevalence of biopsy-demonstrated IM at the cardia, on the gastric side of the squamocolumnar junction, in the absence of endoscopic evidence of a CLE. Evidence is accumulating that these patches of what appears to be Barrett’s in the cardia have a similar malignant potential as in the longer segments, and are precursors for carcinoma of the cardia.

The long-term relief of symptoms remains the primary reason for performing antireflux surgery in patients with BE. Healing of esophageal mucosal injury and the prevention of disease progression are important secondary goals. In this regard, patients with BE are no different than the broader population of patients with gastroesophageal reflux. They should be considered for antireflux surgery when patient data suggest severe disease or predict the need for long-term medical management. Most patients with BE are symptomatic. Although it has been argued that some patients with BE may not have symptoms, careful history taking will reveal the presence of symptoms in most, if not all, patients.

Patients with BE have a spectrum of disease ranging from visually identifiable but short segments, to long segments of classic BE. In general, however, they represent a relatively severe stage of gastroesophageal reflux, usually with markedly increased esophageal acid exposure, deficient LES characteristics, poor esophageal body function, and a high prevalence of duodenogastric reflux. Gastric hypersecretion occurs in 44% of patients. Most will require long-term PPI therapy for relief of symptoms and control of coexistent esophageal mucosal injury. Given such profound deficits in esophageal physiology, antireflux surgery is an excellent means of long-term control of reflux symptoms for most patients with BE.

The typical complications in BE include ulceration in the columnar-lined segment, stricture formation, and a dysplasia-cancer sequence. Barrett’s ulceration is unlike the erosive ulceration of reflux esophagitis in that it more closely resembles peptic ulceration in the stomach or duodenum, and has the same propensity to bleed, penetrate, or perforate. Fortunately, this complication occurs very rarely. The strictures found in BE occur at the squamocolumnar junction, and they are typically higher than peptic strictures in the absence of BE. Ulceration and stricture in association with BE were commonly reported before 1975, but with the advent of potent acid suppression medication, they have become less common. In contrast, the complication of adenocarcinoma developing in Barrett’s mucosa has become more common. Adenocarcinoma developing in Barrett’s mucosa was considered a rare tumor before 1975. Today, it occurs at approximately 0.2% to 0.5% per year of follow-up, which represents a risk 40 times that of the general population. Most, if not all, cases of adenocarcinoma of the esophagus arise in Barrett’s epithelium (Fig. 25-31). About one-third of all patients with BE present with malignancy.

The long-term risk of progression to dysplasia and adenocarcinoma, although not the driving force behind the decision to perform antireflux surgery, is a significant concern for both patient and physician. Although to date, there have been no prospective randomized studies documenting that antireflux surgery has an effect on the risk of progression to dysplasia and carcinoma, complete control of reflux of gastric juice into the esophagus is clearly a desirable goal.

**Respiratory Complications**

A significant proportion of patients with GERD will have associated respiratory symptoms. These patients may have laryngopharyngeal reflux-type symptoms, adult-onset asthma, or even idiopathic pulmonary fibrosis. These symptoms and organ injury may occur in isolation or in conjunction with typical reflux symptoms such as heartburn and regurgitation. Several studies have demonstrated that up to 50% of patients with asthma have either endoscopically evident esophagitis or abnormal distal esophageal acid exposure. These findings support a causal relationship between GERD and aerodigestive symptoms and complications in a proportion of patients.
Etiology of Reflux-Induced Respiratory Symptoms. There are two mechanisms that have been proposed as the cause of reflux-induced respiratory symptoms. The reflux theory suggests that these symptoms are the direct result of laryngopharyngeal exposure and aspiration of gastric contents. The reflex theory suggests that the vagal-mediated afferent fibers result in bronchoconstriction during episodes of distal esophageal acidification. The evidence supporting a mechanism of direct exposure to the aerodigestive system is based in clinical studies that have documented a strong correlation between idiopathic pulmonary fibrosis and hiatal hernia. In addition, the presence of GERD was demonstrated to be highly associated with several pulmonary diseases in a recent Department of Veteran Affairs multivariate analysis. Next, with ambulatory pH testing, acid exposure within the proximal esophagus is more frequently identified in patients with gastroesophageal reflux and respiratory symptoms than in patients who have gastroesophageal reflux symptoms alone. These findings are supported by scintigraphic studies, which have demonstrated aspiration of ingested radioisotope in patients with both gastroesophageal reflux and pulmonary symptoms. In animal studies, tracheal instillation of acid has been demonstrated to profoundly increase airway resistance. Finally, in patients who have undergone multichannel intraluminal impedance testing with a catheter configured to detect laryngopharyngeal reflux, a correlation between proximal fluid movement and laryngopharyngeal symptoms, such as cough, can be demonstrated.

The reflex mechanism is supported by the bronchoconstriction that occurs with the infusion of acid into the distal esophagus. There is a shared embryologic origin of the tracheoesophageal tract and vagus nerve, and this reflex is thought to be an afferent fiber–mediated reflex that protects the aerodigestive system from the aspiration of refluxate. In patients with respiratory symptoms and documented gastroesophageal reflux without proximal esophageal acid exposure, pulmonary symptoms will often times significantly improve or completely resolve after undergoing laparoscopic fundoplication. It is likely that both of the proposed mechanisms work simultaneously to cause these symptoms in the face of GERD.

The most difficult clinical challenge in formulating a treatment plan for reflux-associated respiratory symptoms resides in establishing the diagnosis. Although the diagnosis may be straightforward in patients with predominately typical reflux symptoms and secondary respiratory complaints, a substantial number of patients will have respiratory symptoms that dominate the clinical scenario. Typical gastroesophageal reflux
symptoms, such as heartburn and regurgitation, may often be completely absent only to be uncovered with objective esophageal physiology testing. Traditionally, the diagnosis of reflux-induced respiratory injury is established using ambulatory dual probe pH monitoring, with one probe positioned within the distal esophagus and the other at a proximal location. Proximal probe positioning has included multiple locations such as the trachea, pharynx, and proximal esophagus. Although ambulatory esophageal pH monitoring allows a direct correlation between esophageal acidification and respiratory symptoms, sensitivity of this testing modality is poor, and the temporal relationship between laryngeal or pulmonary symptoms and reflux events is complex. In addition, as the refluxed gastric fluid travels proximally, it may be neutralized by saliva and therefore go undetected with pH monitoring. Impedance testing may also be used to detect the movement of fluid throughout the entire esophageal column regardless of pH content.

**Treatment.** Once the diagnosis is established, treatment may be initiated with either PPI therapy or antireflux surgery. A trial of high-dose PPI therapy may help establish that reflux is partly or completely responsible for the respiratory symptoms. It is important to note that the persistence of symptoms in the face of aggressive PPI treatment does not necessarily rule out reflux as a possible cofactor or sole etiology.

Although there is probably some element of a placebo effect, relief of respiratory symptoms can be anticipated in up to 50% of patients with reflux-induced asthma treated with antisecretory medications. However, when examined objectively, <15% of patients can be expected to have improvement in their pulmonary function with medical therapy. In properly selected patients, antireflux surgery improves respiratory symptoms in nearly 90% of children and 70% of adults with asthma and reflux disease. Improvements in pulmonary function can be demonstrated in around 30% of patients. Uncontrolled studies of the two forms of therapy (PPI and surgery) and the evidence from the two randomized controlled trials of medical vs. surgical therapy indicate that surgical valve reconstruction is the most effective therapy for reflux-induced asthma. The superiority of the surgery over PPI is most noticeable in the supine position, which corresponds with the nadir of PPI blood levels and resultant acid breakthrough and is the time in the circadian cycle when asthma symptoms are at their worst.

In asthmatic patients with an esophageal motility disorder, performing an antireflux operation will not prevent the regurgitation and possible aspiration of swallowed liquid or food “upstream” to the valve reconstruction. It is critical that esophageal body function be considered prior to surgical intervention in this patient population.

**Medical Therapy for Gastroesophageal Reflux Disease.**

With the widespread availability of over-the-counter antisecretory medications, most patients with mild or moderate symptoms will carry self-medication. When initially identified with mild symptoms of uncomplicated GERD, patients can be placed on 12 weeks of simple antacids before diagnostic testing is initiated. This approach may successfully and completely resolve the symptoms. Patients should be counseled to elevate the head of the bed; avoid tight-fitting clothing; eat small, frequent meals; avoid eating the nighttime meal immediately prior to bedtime; and avoid alcohol, coffee, chocolate, and peppermint, which are known to reduce resting LES pressure and may aggravate symptoms.

Used in combination with simple antacids, alginic acid may augment the relief of symptoms by creating a physical barrier to reflux, as well as by acid reduction. Alginic acid reacts with sodium bicarbonate in the presence of saliva to form a highly viscous solution that floats like a raft on the surface of the gastric contents. When reflux occurs, this protective layer is refluxed into the esophagus, and acts as a protective barrier against the noxious gastric contents. Medications to promote gastric emptying, such as metoclopramide or domperidone, are beneficial in early disease but of little value in more severe disease.

In patients with persistent symptoms, the mainstay of medical therapy is acid suppression. High-dosage regimens of hydrogen potassium PPIs, such as omeprazole (up to 40 mg/d), can reduce gastric acidity by as much as 80% to 90%. This usually heals mild esophagitis. In severe esophagitis, healing may occur in only one-half of the patients. In patients who reflux a combination of gastric and duodenal juice, acid-suppression therapy may give relief of symptoms, while still allowing mixed reflux to occur. This can allow persistent mucosal damage in an asymptomatic patient. Unfortunately, within 6 months of discontinuation of any form of medical therapy for GERD, 80% of patients have a recurrence of symptoms, and 40% of individuals with daily GERD eventually develop symptoms that “breakthrough” adequately dosed PPIs. Once initiated, most patients with GERD will require lifelong treatment with PPIs, both to relieve symptoms and to control any coexistent esophagitis or stricture. Although control of symptoms has historically served as the endpoint of therapy, the wisdom of this approach has recently been questioned, particularly in patients with BE. Evidence suggesting that reflux control may prevent the development of adenocarcinoma and lead to regression of dysplastic and nondysplastic Barrett’s segments has led many to consider control of reflux, and not symptom control, a better therapeutic endpoint. However, this hypothesis remains controversial. It should be noted that complete control of reflux using PPIs can be difficult, as has been highlighted by studies of acid breakthrough while on PPI therapy and of persistent reflux following antireflux surgery. Castell, Triadafilopoulos, and others have shown that 40% to 80% of patients with BE continue to have abnormal esophageal acid exposure despite up to 20 mg twice daily of PPIs. Ablation trials have shown that mean doses of 56 mg of omeprazole were necessary to normalize 24-hour esophageal pH studies. It is likely that antireflux surgery results in more reproducible and reliable elimination of reflux of both acid and duodenal contents, although long-term outcome studies suggest that as many as 25% of postfundoplication patients will have persistent pathologic esophageal acid exposure confirmed by positive 24-hour pH studies.

**Suggested Therapeutic Approach.** Traditionally a stepwise approach is used for the treatment of GERD. First-line therapy entails antisecretory medication, usually PPIs, in most patients. Failure of medication to adequately control GERD symptoms suggests either that the patient may have relatively severe disease or a non-GERD cause for his or her symptoms. Endoscopic examination at this stage of the patient’s evaluation is recommended and will provide the opportunity to assess the degree of mucosal injury and presence of BE. Treatment options for these patients entails either long term PPI use vs. antireflux surgery. Laparoscopic antireflux surgery in these patients achieves long-term control of symptoms in 85% to 90%. The measurement
of esophageal acid exposure via 24-hour pH should be undertaken when patients are considered for surgery. The status of the LES and esophageal body function with esophageal manometry should also be performed at this stage. These studies will serve to establish the diagnosis and assess esophageal body dysfunction.

### Surgical Therapy for Gastroesophageal Reflux Disease

#### Selection of Patients for Surgery

Studies of the natural history of GERD indicate that most patients have a relatively benign form of the disease that is responsive to lifestyle changes and dietary and medical therapy and do not need surgical treatment. Approximately 25% to 50% of the patients with GERD have persistent or progressive disease, and it is this patient population that is best suited to surgical therapy. In the past, the presence of esophagitis and a structurally defective LES were the primary indications for surgical treatment, and many internists and surgeons were reluctant to recommend operative procedures in their absence. However, one should not be deterred from considering antireflux surgery in a symptomatic patient with or without esophagitis or a defective sphincter, provided the disease process has been objectively documented by 24-hour pH monitoring. This is particularly true in patients who have become dependent upon therapy with PPIs, or require increasing doses to control their symptoms. It is important to note that a good response to medical therapy in this group of patients predicts an excellent outcome following antireflux surgery.

In general, the key indications for antireflux surgery are (a) objectively proven gastroesophageal reflux disease, and (b) typical symptoms of gastroesophageal reflux disease (heartburn and/or regurgitation) despite adequate medical management, or (c) a younger patient unwilling to take lifelong medication. In addition, a structurally defective LES can also predict which patients are more likely to fail with medical therapy. Patients with normal sphincter pressures tend to remain well controlled with medical therapy, whereas patients with a structurally defective LES may not respond as well to medical therapy, and often develop recurrent symptoms within 1 to 2 years of beginning therapy. Such patients should be considered for an antireflux operation, regardless of the presence or absence of endoscopic esophagitis.

Young patients with documented reflux disease with or without a defective LES are also excellent candidates for antireflux surgery. They usually will require long-term medical therapy for control of their symptoms, and some will go on to develop complications of the disease. An analysis of the cost of therapy based on data from the Veterans Administration Cooperative trial indicates that surgery has a cost advantage over medical therapy in patients <49 years of age.

Severe endoscopic esophagitis in a symptomatic patient with a structurally defective LES is also an indication for early surgical therapy. These patients are prone to breakthrough of their symptoms while receiving medical therapy. Symptoms and mucosal injury can be controlled in such patients, but careful monitoring is required, and increasing dosages of PPIs are necessary. In everyday clinical practice, however, such treatment can be both difficult and impractical, and, in such cases, antireflux surgery can be considered early, especially if PPI therapy is problematic.

The development of a stricture in a patient represents a failure of medical therapy, and it is also an indication for a surgical antireflux procedure. In addition, strictures are often associated with a structurally defective sphincter and loss of esophageal contractility. Before proceeding with surgical treatment, malignancy and a drug-related etiology of the stricture should be excluded, and the stricture should be progressively dilated up to a 50 to 60F bougie. When the stricture is fully dilated, the relief of dysphagia is evaluated, and esophageal manometry is performed to determine the adequacy of peristalsis in the distal esophagus. If dysphagia is relieved and the amplitude of esophageal contractions is adequate, an antireflux procedure should be performed; if there is a global loss of esophageal contractility, caution should be exercised in performing an antireflux procedure with a complete fundoplication, and a partial fundoplication should be considered.

Barrett’s CLE is commonly associated with a severe structural defect of the LES and often poor contractility of the esophageal body. Patients with BE are at risk of the development of an adenocarcinoma. Whilst surgeons would like to think that an antireflux procedure can reduce the risk of progression to cancer, the evidence supporting this is relatively weak, and for now Barrett’s esophagus should be considered to be evidence that the patient has gastroesophageal reflux, and progression to antireflux surgery is indicated for the treatment of reflux symptoms, not cancer progression. If, however, high grade dysplasia or intramucosal carcinoma is found on mucosal biopsy specimens, treatment should then be directed at the BE and the lesion, using either evaluation endoscopic ablation, endoscopic resection, or esophageal resection.

The majority of patients requiring treatment for reflux have a relatively mild form of disease and will respond to antisecretory medications. Patients with more severe forms of disease, particularly those who develop persistent or progressive disease, should be considered for definitive therapy. Laparoscopic fundoplication will provide a long-term cure in the majority of these patients, with minimal discomfort and an early return to normal activity.

#### Preoperative Evaluation

Before proceeding with an antireflux operation, several factors should be evaluated. The clinical symptoms should be consistent with the diagnosis of gastroesophageal reflux. Patients presenting with the typical symptoms of heartburn and/or regurgitation which have responded, at least partly, to PPI therapy, will generally do well following surgery, whereas patients with atypical symptoms have a less predictable response. Reflux should also be objectively confirmed by either the presence of ulcerative esophagitis or an abnormal 24-hour pH study.

The propulsive force of the body of the esophagus should be evaluated by esophageal manometry to determine if it has sufficient power to propel a bolus of food through a newly reconstructed valve. Patients with normal peristaltic contractions can be considered for a 360° Nissen fundoplication or a partial fundoplication, depending on patient and surgeon preferences. When peristalsis is absent, a partial fundoplication is probably the procedure of choice, but only if achalasia has been ruled out.

Hiatal anatomy should also be assessed. In patients with smaller hiatal hernias, endoscopy evaluation usually provides sufficient information. However, when patients present with a very large hiatus hernia or for revision surgery after previous antireflux surgery, contrast radiology provides better anatomical information. The concept of anatomic shortening of the esophagus is controversial, with divergent opinions held about how
common this problem is. Believers claim that anatomic shortening of the esophagus compromises the ability of the surgeon to perform an adequate repair without tension and that this can lead to an increased incidence of breakdown or thoracic displacement of the repair. Some of those who hold this view claim that esophageal shortening is present when a barium swallow X-ray identifies a sliding hiatal hernia that will not reduce in the upright position or that measures more than 5 cm in length at endoscopy. When such identification is made, these surgeons usually add a gastroplasty to the antireflux procedure. Others claim that esophageal shortening is overdiagnosed and rarely seen, and that the morbidity of adding a gastroplasty outweighs any benefits. These surgeons would recommend a standard antireflux procedure in all patients undergoing primary surgery.

**Principles of Surgical Therapy.** The primary goal of antireflux surgery is to safely create a new antireflux valve at the gastroesophageal junction, while preserving the patient’s ability to swallow normally and to belch to relieve gaseous distention. Regardless of the choice of the procedure, this goal can be achieved if attention is paid to some basic principles when reconstructing the antireflux mechanism. First, the operation should create a flap valve which prevents regurgitation of gastric contents into the esophagus. This will result in an increase in the pressure of the distal esophageal sphincter region. Following a Nissen fundoplication the expected increase is to a level twice the resting gastric pressure (i.e., 12 mmHg for a gastric pressure of 6 mmHg). The extent of the pressure rise is often less following a partial fundoplication, although with all types of fundoplication the length of the reconstructed valve should be at least 3 cm. This not only augments sphincter characteristics in patients in whom they are reduced before surgery but also prevents unfolding of a normal sphincter in response to gastric distention (Fig. 25-32). Preoperative and postoperative esophageal manometry measurements have shown that the resting sphincter pressure and the overall sphincter length can be surgically augmented over preoperative values, and that the change in the former is a function of the degree of gastric wrap around the esophagus (Fig. 25-33). However, the aim of any fundoplication is to create a loose wrap and to maintain the position of the gastric fundus close to the distal intra-abdominal esophagus, in a flap valve arrangement. The efficacy of this relies on the close relationship between the fundus and the esophagus, not the “tightness” of the wrap.

Second, the operation should place an adequate length of the distal esophageal sphincter in the positive-pressure environment of the abdomen by a method that ensures its response to changes in intra-abdominal pressure. The permanent restoration of 2 or more cm of abdominal esophagus ensures the preservation of the relationship between the fundus and the esophagus. All of the popular antireflux procedures increase the length of the sphincter exposed to abdominal pressure by an average of at least 1 cm.

Third, the operation should allow the reconstructed cardia to relax on deglutition. In normal swallowing, a vagally mediated relaxation of the distal esophageal sphincter and the gastric fundus occurs. The relaxation lasts for approximately 10 seconds and is followed by a rapid recovery to the former tonicity. To ensure relaxation of the sphincter, three factors are important: (a) Only the fundus of the stomach should be used to buttress the sphincter, because it is known to relax in concert with the sphincter; (b) the gastric wrap should be properly placed around the sphincter and not incorporate a portion of the stomach or be placed around the stomach itself, because the body of the stomach does not relax with swallowing; and (c) damage to the vagal nerves during dissection of the thoracic esophagus should be avoided because it may result in failure of the sphincter to relax.

Fourth, the fundoplication should not increase the resistance of the relaxed sphincter to a level that exceeds the peristaltic power of the body of the esophagus. The resistance of the relaxed sphincter depends on the degree, length, and diameter of the gastric fundic wrap, and on the variation in intra-abdominal pressure. A 360° gastric wrap should be no longer than 2 cm and constructed over a large (50 to 60F) bougie. This will ensure that the relaxed sphincter will have an adequate diameter with minimal resistance. A bougie is not necessary when constructing a partial wrap.

Fifth, the operation should ensure that the fundoplication can be placed in the abdomen without undue tension and maintained there by approximating the crura of the diaphragm above the repair. Leaving the fundoplication in the thorax converts a sliding hernia into a PEH, with all the complications associated with that condition. Maintaining the repair in the abdomen...

---

**Figure 25-32.** A graphic illustration of the shortening of the lower esophageal sphincter that occurs as the sphincter is “taken up” by the cardia as the stomach distends.

**Figure 25-33.** The relationship between the augmentation of sphincter pressure over preoperative pressure (ΔP) and the degree of gastric fundic wrap in three different antireflux procedures. (Reproduced with permission from O’Sullivan GC, DeMeester TR, Joelson BE, et al: Interaction of lower esophageal sphincter pressure and length of sphincter in the abdomen as determinants of gastroesophageal competence, Am J Surg. 1982 Jan;143(1):40-47.)
under tension predisposes to an increased incidence of recurrence. How common this problem is encountered is disputed, with some surgeons advocating lengthening the esophagus by gastroplasty and constructing a partial fundoplication, and others claiming that this issue is now rarely encountered.

**Procedure Selection.** A laparoscopic approach is now used routinely in all patients undergoing primary antireflux surgery. Some surgeons advocate the use of a single antireflux procedure for all patients, whereas others advocate a tailored approach. Advocates of the laparoscopic Nissen fundoplication as the procedure of choice for a primary antireflux repair would generally apply this procedure in all patients with normal or near normal esophageal motility, and they would reserve a partial fundoplication for use in individuals with poor esophageal body motility. Others, based on the good longer-term outcomes now reported following partial fundoplication procedures, advocate the routine application of a partial fundoplication procedure, thereby avoiding any concerns about constructing a fundoplication in individuals with poor esophageal motility.

Experience and randomized studies have shown that both the Nissen fundoplication and various partial fundoplication procedures are all effective and durable antireflux repairs that generate an excellent outcome in approximately 90% of patients at longer-term follow-up.

**Primary Antireflux Repairs**

**Nissen Fundoplication.** The most common antireflux procedure is the Nissen fundoplication. In the past, this procedure has been performed through an open abdominal or a chest incision, but with the development of laparoscopic approaches primary antireflux surgery is now routinely undertaken using the laparoscope. Rudolph Nissen described this procedure as a 360° fundoplication around the lower esophagus for a distance of 4 to 5 cm, without division of the short gastric blood vessels. Although this provided good control of reflux, it was associated with a number of side effects that have encouraged modifications of the procedure as originally described. These include using only the gastric fundus to envelop the esophagus in a fashion analogous to a Witzel jejunostomy, sizing the fundoplication with a large (50 to 60F) bougie, limiting the length of the fundoplication to 1 to 2 cm, and dividing the short gastric vessels. The essential elements necessary for the performance of a transabdominal fundoplication are common to both the laparoscopic and open procedures and include the following:

1. Hiatal dissection and preservation of both vagi along their entire length
2. Circumferential esophageal mobilization
3. Hiatal closure, usually posterior to the esophagus
4. Creation of a short and floppy fundoplication over an esophageal dilator

In addition, many surgeons also routinely divide the short gastric blood vessels, although this step is not universally applied, and the results of several randomized trials have failed to show that this step yields any benefit.

The laparoscopic approach to fundoplication has now replaced the open abdominal Nissen fundoplication as the procedure of choice. Five ports are usually used (Fig. 25-34), and dissection is begun by incising the gastrohepatic omentum above and below the hepatic branch of the anterior vagus nerve, which is usually preserved. The circumference of the diaphragmatic hiatus is dissected and the esophagus is mobilized by careful dissection of the anterior and posterior soft tissues within the hiatus. The esophagus is held anterior and to the left and the hiatal pillars are approximated with interrupted nonabsorbable sutures, starting posteriorly and working anteriorly. A tension-free fundoplication should be constructed. This can usually be achieved either with or without division of the short gastric blood vessels, according to surgeon preference. If the vessels are divided, the upper one-third of the greater curvature is mobilized by sequentially dissecting and dividing these vessels, commencing distally and working proximally. Following complete fundal mobilization, the posterior wall of the fundus is brought behind the esophagus to the right side, and the anterior wall of the fundus is brought anterior to the esophagus. The fundic lips are manipulated to allow the fundus to envelop the esophagus without twisting. A 50 to 60F bougie is passed to properly size the fundoplication, and it is sutured using nonabsorbable sutures. Some surgeons use a single U-stitch of 2-0 polypropylene buttressed with felt pledgets (Fig. 25-35), and others use 2-4 interrupted sutures.
**Posterior Partial Fundoplication.** Partial fundoplications were developed as an alternative to the Nissen procedure in an attempt to minimize the risk of postfundoplication side effects, such as dysphagia, inability to belch, and flatulence. The commonest approach has been a posterior partial or Toupet fundoplication. Some surgeons use this type of procedure for all patients presenting for antireflux surgery, whereas others apply a tailored approach in which a partial fundoplication is constructed in patients with impaired esophageal motility, in which the propulsive force of the esophagus is thought to be insufficient to overcome the outflow obstruction of a complete fundoplication. The Toupet posterior partial fundoplication consists of a 270° gastric fundoplication around the distal 4 cm of esophagus (Fig. 25-36). It is usually stabilized by anchoring the wrap posteriorly to the hiatal rim.

**Anterior Partial Fundoplication.** An alternative approach to partial fundoplication is to construct an anterior partial fundoplication. Following posterior hiatal repair, the anterior fundus is rolled over the front of the esophagus and sutured to the hiatal rim and the esophageal wall. Division of the short gastric vessels

---

**Figure 25-35.** A. Laparoscopic Nissen fundoplication is performed with a five-trocar technique. B. The liver retractor is affixed to a mechanical arm to hold it in place throughout the operation. C. After division of the gastrohepatic omentum above the hepatic branch of the vagus (pars flaccida), the surgeon places a blunt atraumatic grasper beneath the phrenoesophageal ligament. D. After completion of the crural closure, an atraumatic grasper is placed right to left behind the gastroesophageal junction. The grasper is withdrawn, pulling the posterior aspect of the gastric fundus behind the esophagus. E. Once the suture positions are chosen, the first stitch (2-0 silk, 20 cm long) is introduced through the 10-mm trocar, and the needle is passed first through the left limb of the fundus, then the esophagus (2.5 cm above the gastroesophageal junction), then through the right limb of the fundus. F. Final position of the fundoplication.
is never needed when constructing this type of fundoplication. Various degrees of anterior partial fundoplication have been described—90°, 120°, 180°. The anterior 180° partial fundoplication (Fig. 25-37) provides a more robust fundoplication and achieves an excellent longer-term outcome in approximately 90% of patients at follow-up of at least 10 years. With this procedure, the fundus and esophagus are sutured to the right side of the hiatal rim to create a flap valve at the gastroesophageal junction and to stabilize a 3 to 4 cm length of intra-abdominal esophagus.

Collis Gastroplasty. When a shortened esophagus is encountered, many surgeons choose to add an esophageal lengthening procedure before fundoplication, to reduce the tension on the gastroesophageal junction, believing this will minimize the risk of failure due to postoperative hiatus hernia. The commonest approach to this is the Collis gastroplasty. This entails using a stapler to divide the cardia and upper stomach, parallel to the lesser curvature of the stomach, thereby creating a gastric tube in continuity with the esophagus, and effectively lengthening the esophagus by several centimeters. Laparoscopic techniques for Collis gastroplasty have been described (Fig. 25-38). Following gastroplasty a fundoplication is constructed, with the highest suture placed on the native esophagus when constructing a Nissen fundoplication. Not all surgeons choose to undertake a Collis procedure, however, as there is controversy about the actual incidence of the shortened esophagus and widely divergent views are held about how often this problem is encountered. In addition, some surgeons have questioned the wisdom of creating an motile tube of gastric wall, which can secrete acid, and then placing a Nissen fundoplication below this.

Outcome After Fundoplication. Studies of long-term outcome following both open and laparoscopic fundoplication document the ability of laparoscopic fundoplication to relieve typical reflux symptoms (heartburn, regurgitation, and dysphagia) in more than

Figure 25-35. (Continued)
After removal of the fat pad and release of tension on the Penrose drain, the gastroesophageal junction (GES) retracts to the level of the hiatus. The interior end of the staple line is marked 2 cm below the angle of His. The first horizontal firing of the stapler occurs by maximally articulating the stapler to the left, aiming toward the previously marked spot adjacent to the dilator. The vertical staple line is created by a single firing of the GIA placed parallel and flush against the 48F dilator. The highest Nissen fundoplication suture is placed on the native esophagus, and the second suture tucks in the apex of the staple line.

90% of patients at follow-up intervals averaging 2 to 3 years and 80% to 90% of patients 5 years or more following surgery. This includes evidence-based reviews of antireflux surgery, prospective randomized trials comparing antireflux surgery to PPI therapy and open to laparoscopic fundoplication and analysis of U.S. national trends in use and outcomes. Postoperative pH studies indicate that more than 90% of patients will normalize their pH tracings. The results of laparoscopic fundoplication compare favorably with those of the “modern” era of open fundoplication. They also indicate the less predictable outcome of atypical reflux symptoms (cough, asthma, laryngitis) after surgery, being relieved in only two-thirds of patients.

The goal of surgical treatment for GERD is to relieve the symptoms of reflux by reestablishing the gastroesophageal barrier. The challenge is to accomplish this without inducing dysphagia or other untoward side effects. Dysphagia, existing before surgery, usually improves following laparoscopic fundoplication. Temporary dysphagia is common after surgery and generally resolves within 3 months, but it can take up to 12 months in some individuals, and dysphagia sufficient to require ongoing dietary modification persists in up to 5% of individuals following Nissen fundoplication. Other side effects common to antireflux surgery include the inability to belch and vomit and increased flatulence. Most patients cannot vomit through an intact wrap, though this is rarely clinically relevant. Most patients are unable to belch gas from the stomach in the first 3 to 6 months after fundoplication, but 80% to 90% regain the ability to belch normally beyond the first 12 months of follow-up. Hyperflatulence is a common and noticeable problem, likely related to increased air swallowing that is present in most patients with reflux disease, aggravated by the inability to belch in some patients.
Randomized Controlled Trials Addressing Surgical Technique

Division of the Short Gastric Blood Vessels

Originally, Nissen’s description of a total fundoplication entailed a 360° fundoplication during which the short gastric blood vessels were left intact. However, with reports of troublesome postoperative dysphagia, division of these vessels—to achieve full fundal mobilization and thereby ensure a loose fundoplication—was promoted and has entered common practice. The evidence supporting dividing these vessels has been based on the outcomes from uncontrolled case series of patients undergoing Nissen fundoplication either with vs. without division of the short gastric vessels. However, the results from these studies have been conflicting, with different proponents reporting good results irrespective of whether these vessels have been divided or not. To address this issue, six randomized trials that enrolled a total of 438 patients have been reported. None of these trials demonstrated any differences for the postoperative dysphagia or recurrent gastro-esophageal reflux. However, in the three largest of the six trials an increased incidence of flatulence and bloating symptoms, as well as greater difficulty with belching, was seen in patients in whom the short gastric vessels were divided.

A recent meta-analysis from Engstrom et al., generated by combining the raw data from Australian and Swedish trials, evaluated a larger cohort of 201 patients, with 12 years of follow-up in 170, and also confirmed equivalent reflux control but found more abdominal bloating after division of the short gastric vessels. Overall, these trials fail to support the belief that dividing the short gastric vessels improves any outcome following Nissen fundoplication. The trials actually suggest that dividing the vessels increases the complexity of the procedure and leads to a poorer outcome due to the increase in bloating symptoms.

Nissen vs. Posterior Partial Fundoplication

Eleven randomized trials have compared Nissen vs. posterior partial fundoplication. Some of the trials contributed little to the pool of evidence, as they are either small or underpowered, and failed to show significant outcome differences. The larger trials, however, have consistently demonstrated equivalent reflux control, but they also show a reduced incidence of wind-related side-effects (flatulence, bloating, and inability to belch) following posterior partial fundoplication procedures, although less dysphagia following a posterior fundoplication was only demonstrated in 2 of the 11 trials. Lundell et al. reported the outcomes of Nissen vs. Toupet partial fundoplication in a trial that enrolled 137 patients with reported follow-up to 18 years. Reflux control and dysphagia symptoms were similar, but flatulence was commoner after Nissen fundoplication at some medium-term follow-up time points, and revision surgery was more common following Nissen fundoplication, mainly to correct postoperative paraoesophageal herniation. At 18 years follow-up, success rates of more than 80% were reported for both procedures, as well as no significant differences in the incidence of side effects. The data from this trial suggested that the mechanical side effects following Nissen fundoplication progressively improve with very long-term follow-up. Strate et al. reported 2-year follow-up in a trial that enrolled 200 patients. Approximately 85% of each group was satisfied with the clinical outcome, but dysphagia was significantly more common following Nissen fundoplication (19 vs. 8 patients).

Other trials (Guérin et al–140 patients, Booth et al–127, Khan et al–121, Shaw et al–100) also report similar reflux control within the first few years of follow-up. Only Booth et al. demonstrated less dysphagia following posterior fundoplication. Subgroup analysis in 3 trials (Booth, Shaw, Zornig) did not reveal differences between patients with vs. without pre-operative oesophageal motility. Overall these trials suggest that some side-effects, mainly wind-related issues, are less common following posterior partial fundoplication. However, the hypothesis that dysphagia is less of a problem following posterior partial fundoplication has only been substantiated in 2 of 11 trials.

Nissen vs. Anterior Fundoplication

Six trials have evaluated Nissen vs. anterior partial fundoplication variants. Four have assessed Nissen vs. anterior 180° partial fundoplication (Watson et al–107 patients, Baigrie et al–161, Cao et al–100, Raue et al–64). These trials all demonstrated equivalent reflux control, but less dysphagia and less wind-related side effects after anterior 180° partial fundoplication at up to 5 years follow-up. Only the study from Watson et al. has reported follow-up to 10 years, and at late follow-up in their trial there were no significant outcome differences for the two procedures, with equivalent control of reflux, and no differences for side effects due to a progressive decline in dysphagia as follow-up extended beyond 5 years.

Two trials compared laparoscopic anterior 90° partial fundoplication vs. Nissen fundoplication (Watson et al–112 patients, Spence et al–79). In both of these trials, side-effects were less common following anterior 90° fundoplication, but this was offset by a slightly higher incidence of recurrent reflux at up to 5 years follow-up. Satisfaction with the overall outcome was similar for both fundoplication variants.

Anterior vs. Posterior Partial Fundoplication

Two randomized trials have directly compared anterior vs. posterior partial fundoplication. Hagedorn et al randomized 95 patients to undergo either Toupet vs. anterior 120° partial fundoplication, and Khaf et al enrolled 103 patients to anterior 180° vs. posterior partial fundoplication. Both studies demonstrated better reflux control, offset by more side effects following posterior partial fundoplication. The anterior 120° partial fundoplication performed by Hagedorn et al was similar to the anterior 90° variant described above. However, the outcomes following this procedure were much worse in this trial than the outcomes in other studies, with the average exposure time to acid (pH <4%–5.6%) following anterior fundoplication in their study unusually high compared to other studies. Khan et al only reported 6 months follow-up, and longer-term outcomes are awaited before drawing firm conclusions. The overall results from all eight trials that included an anterior fundoplication variant suggest that this type of fundoplication achieves satisfactory reflux control, with less dysphagia and other side-effects, yielding a good overall outcome. However, the reduced incidence of troublesome side-effects is traded off against a higher risk of recurrent reflux.

Outcome of Antireflux Surgery in Patients With Barrett’s Esophagus

Few studies have focused on the alleviation of symptoms after antireflux surgery in patients with BE (Table 25–7). Those that are available document excellent to good results in 72% to 95% of patients at 5 years following surgery. Several nonrandomized studies have compared medical and surgical therapy and report better outcomes after antireflux surgery. Parrella and colleagues reported the only randomized trial to evaluate this issue. They enrolled 101 patients over 18 years, and median follow-up was 6 years. Medical therapy consisted of 20 mg of omeprazole (PPI) twice daily since 1992 in all medically treated patients, and surgical therapy consisted of an open Nissen
fundoplication. The symptomatic outcome in the two groups was nearly identical, although esophagitis and/or stricture persisted in 20% of the medically treated patients, compared to only 3% to 7% of patients following antireflux surgery. About 15% of patients had abnormal acid exposure after surgery. Although pH data were not routinely collected in patients on PPI therapy, in the subgroup of 12 patients that did have 24-hour monitoring on treatment, 3 of 12 (25%) had persistently high esophageal acid exposure, and most (75%) had persistently high bilirubin exposure.

The common belief that Barrett’s epithelium cannot be reversed by antireflux surgery may not be correct. Within the control arm of a randomized trial of ablation vs. surveillance, Bright and associates identified approximately 50% regression in the length of Barrett’s esophagus in 20 patients within the control arm of a randomized trial of ablation vs. surveillance.

Current data indicate that patients with BE should remain in an endoscopic surveillance program following antireflux surgery. Biopsy specimens should be reviewed by a pathologist with expertise in the field. If low-grade dysplasia is confirmed, biopsy specimens should be repeated after 12 weeks of high-dose acid suppression therapy. If high-grade dysplasia or intramucosal cancer is evident on more than one biopsy specimen, then treatment is escalated. Treatment options include endoscopic mucosal resection, endoscopic ablation of the BE, or esophageal resection. Esophageal resection is advisable when an invasive cancer (stage T1b or deeper) is present, or for multifocal long segment BE in younger and fit patients in whom endoscopic treatments are unlikely to be adequate. Endoscopic mucosal resection allows smaller intramucosal tumors to be removed with clear pathology margins, and it can be used as a “big biopsy” to obtain better pathological staging, and even to excise shorter segments of BE in a piecemeal fashion. Ablation, commonly using radiofrequency ablation, has been shown at short-term follow-up in a randomized trial to reduce the rate of progression from high grade dysplasia to invasive cancer by approximately 50%. However, following any endoscopic treatment, patients need to continue with close endoscopic surveillance as recurrence can occur and the longer-term outcome following these treatments remains uncertain. Early detection and treatment have been shown to decrease the mortality rate from esophageal cancer in these patients.

If the dysplasia is reported as lower grade or indeterminate, then inflammatory change that is often confused with dysplasia should be suppressed by a course of acid suppression therapy in high doses for 2 to 3 months, followed by rebiopsy of the Barrett’s segment.

### Table 25-7

<table>
<thead>
<tr>
<th>AUTHOR</th>
<th>YEAR</th>
<th>NO. OF PATIENTS</th>
<th>% EXCELLENT TO GOOD RESPONSE</th>
<th>MEAN FOLLOW-UP, YEARS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starnes</td>
<td>1984</td>
<td>8</td>
<td>75</td>
<td>2</td>
</tr>
<tr>
<td>Williamson</td>
<td>1990</td>
<td>37</td>
<td>92</td>
<td>3</td>
</tr>
<tr>
<td>DeMeester</td>
<td>1990</td>
<td>35</td>
<td>77</td>
<td>3</td>
</tr>
<tr>
<td>McDonald</td>
<td>1996</td>
<td>113</td>
<td>82.2</td>
<td>6.5</td>
</tr>
<tr>
<td>Ortiz</td>
<td>1996</td>
<td>32</td>
<td>90.6</td>
<td>5</td>
</tr>
</tbody>
</table>

Reoperation for Failed Antireflux Repairs. Failure of an antireflux procedure occurs when, after the repair, the patient is unable to swallow normally, experiences upper abdominal discomfort during and after meals, or has recurrence or persistence of reflux symptoms. The assessment of these symptoms and the selection of patients who need further surgery are challenging problems. Functional assessment of patients who have recurrent, persistent, or emergent new symptoms following a primary antireflux repair is critical to identifying the cause of the failure. Analysis of patients requiring reoperation after a previous antireflux procedure shows that placement of the wrap around the stomach is the most frequent cause for failure after open procedures, while herniation of the repair into the chest is the most frequent cause of failure after a laparoscopic procedure. Partial or complete breakdown of the fundoplication and construction of a too-tight a fundoplication or overnarrowing the esophageal hiatus occurs with both open and closed procedures.

Patients who have recurrence of heartburn and regurgitation without dysphagia and have good esophageal motility are most amenable to reoperation, and they can be expected to have an excellent outcome. When dysphagia is the cause of failure, the situation can be more difficult to manage. If the dysphagia occurred immediately following the repair, it is usually due to a technical failure, most commonly a misplaced fundoplication around the upper stomach, or overnarrowing of the esophageal diaphragmatic hiatus and reoperation is usually satisfactory. When dysphagia is associated with poor motility and multiple previous repairs, further revision fundoplication is unlikely to be successful, and in otherwise fit patients it is appropriate to setoriously consider esophageal resection. With each reoperation, the esophagus is damaged further, and the chance of preserving function is decreased. Also, blood supply is reduced, and ischemic necrosis of the esophagus can occur after several previous mobilizations.

### GIANT DIAPHRAGMATIC (HIATAL) HERNIAS

With the advent of clinical radiology, it became evident that a diaphragmatic hernia was a relatively common abnormality and was not always accompanied by symptoms. Three types of esophageal hiatal hernia were identified: (a) the sliding hernia, type I, characterized by an upward dislocation of the cardia in the posterior mediastinum (Fig. 25-39A); (b) the rolling or PEH, type II, characterized by an upward dislocation of the gastric fundus alongside a normally positioned cardia (Fig. 25-39B); and (c) the combined sliding-rolling or mixed hernia, type III, characterized by an upward dislocation of both the cardia and the gastric fundus (Fig. 25-39C). The end stage of type I and type II hernias occurs when the whole stomach migrates up into the chest by rotating 180° around its longitudinal axis, with the cardia and pylorus as fixed points. In this situation, the abnormality is usually referred to as an intrathoracic stomach (Fig. 25-39D). In some taxonomies, a type IV hiatal hernia is declared when an additional organ, usually the colon, herniates as well. Types II–IV hiatal hernias are also referred to as paraesophageal hernia (PEH), as a portion of the stomach is situated adjacent to the esophagus, above the gastroesophageal junction.

### Incidence and Etiology

The true incidence of a hiatal hernia is difficult to determine because of the absence of symptoms in a large number of patients who are subsequently shown to have a hernia. When radiographic examinations are done in response to GI symptoms,
Figure 25-39. A. Radiogram of a type I (sliding) hiatal hernia. B. Radiogram of a type II (rolling or paraesophageal) hernia. C. Radiogram of a type III (combined sliding-rolling or mixed) hernia. D. Radiogram of an intrathoracic stomach. This is the end stage of a large hiatal hernia regardless of its initial classification. Note that the stomach has rotated 180° around its longitudinal axis, with the cardia and pylorus as fixed points. (Reproduced with permission from Nyhus LM, Condon RE: Hernia, 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 1989.)
the incidence of a sliding hiatal hernia is seven times higher than that of a PEH. The PEH is also known as the giant hiatal hernia. Over time the pressure gradient between the abdomen and chest enlarges the hiatal hernia. In many cases the type 1 sliding hernia will evolve into a type III mixed hernia. Type II hernias are quite rare. The age distribution of patients with PEHs is significantly different from that observed in sliding hiatal hernias. The median age of the former is 61 years old; of the latter, 48 years old. PEHs are more likely to occur in women by a ratio of 4:1.

Structural deterioration of the phrenoesophageal membrane over time may explain the higher incidence of hiatal hernias in the older age group. These changes involve thinning of the upper fascial layer of the phrenoesophageal membrane (i.e., the supradiaphragmatic continuation of the endothoracic fascia) and loss of elasticity in the lower fascial layer (i.e., the infra-diaphragmatic continuation of the transversalis fascia). Consequently, the phrenoesophageal membrane yields to stretching in the cranial direction due to the persistent intra-abdominal pressure and the tug of esophageal shortening on swallowing. Interestingly, the stretching and thinning occurs more anteriorly and posteriorly, with fixation of the left crus of the diaphragm to the stomach at the 3 o’clock position, as viewed from the foot. This creates an anterior and posterior hernia sac, the latter of which is often filled with epiphrenic and retroperitoneal fat. These observations point to the conclusion that the development of a hiatal hernia is an age-related phenomenon secondary to repetitive upward stretching of the phrenoesophageal membrane.

Clinical Manifestations

The clinical presentation of a giant hiatal (paraesophageal) hernia differs from that of a sliding hernia. There is usually a higher prevalence of symptoms of dysphagia and postprandial fullness with PEHs, but the typical symptoms of heartburn and regurgitation present in sliding hiatal hernias can also occur. Both are caused by gastroesophageal reflux secondary to an underlying mechanical deficiency of the cardia. The symptoms of dysphagia and postprandial fullness in patients with a PEH are explained by the compression of the adjacent esophagus by a distended cardia, or twisting of the GEJ by the torsion of the stomach that occurs as it becomes progressively displaced in the chest. The postprandial fullness or retrosternal chest pain is a thought to be a result of distension of the stomach with gas or food in the hiatal hernia. Many patients with sliding hernias and reflux symptoms will lose the reflux symptoms when the hernia evolves into the paraesophageal variety. This can be explained by the recreation of the cardiopheric angle when the stomach herniates alongside the GEJ or becomes twisted in the sac. Repair of the hernia without addressing the reflux can create extremely bothersome heartburn. Respiratory complications are frequently associated with a PEH and consist of dyspnea and recurrent pneumonia from aspiration. New research demonstrates that the cause of dyspnea in the presence of a giant PEH is more likely to be left atrial compression, decreasing cardiac output, than a restrictive pulmonary effect, as has been hypothesized for many years.

Approximately one-third of patients with a PEH are found to be anemic, which is due to recurrent bleeding from ulceration of the gastric mucosa in the herniated portion of the stomach, even if ulcerations are not detected at the time of endoscopy. The association of anemia and PEH is best proven by fixing the hernia. Anemia is corrected in >90% of patients with this condition. With time, more and more stomach migrates into the chest and can cause intermittent foregut obstruction due to the rotation that has occurred. In contrast, many patients with PEH are asymptomatic or complain of minor symptoms. However, the presence of a PEH can be life-threatening in that the hernia can lead to sudden catastrophic events, such as excessive bleeding or volvulus with acute gastric obstruction or infarction. With mild dilatation of the stomach, the gastric blood supply can be markedly reduced, causing gastric ischemia, ulceration, perforation, and sepsis. The probability of incarceration/strangulation is not well known, although recent studies suggest that the lifetime risk is less than 5%, making this concern an insufficient concern for routine repair of the asymptomatic PEH.

The symptoms of sliding hiatal hernias are usually due to functional abnormalities associated with gastroesophageal reflux and include heartburn, regurgitation, and dysphagia. These patients have a mechanically defective LES, giving rise to the reflux of gastric juice into the esophagus and the symptoms of heartburn and regurgitation. The symptom of dysphagia occurs from the presence of mucosal edema, Schatzki’s ring, stricture, or the inability to organize peristaltic activity in the body of the esophagus as a consequence of the disease.

There is a group of patients with sliding hiatal hernias not associated with reflux disease who have dysphagia without any obvious endoscopic or manometric explanation. Video barium radiograms have shown that the cause of dysphagia in these patients is an obstruction of the swallowed bolus by diaphragmatic impingement on the herniated stomach. Manometrically, this is reflected by a double-humped high-pressure zone at the GEJ. The first pressure rise is due to diaphragmatic impingement on the herniated stomach, and the second is due to the true distal esophageal sphincter. These patients usually have a mechanically competent sphincter, but the impingement of the diaphragm on the stomach can result in propelling the contents of the supradiaphragmatic portion of the stomach up into the esophagus and pharynx, resulting in complaints of pharyngeal regurgitation and aspiration. Consequently, this abnormality is often confused with typical GERD. Surgical reduction of the hernia results in relief of the dysphagia in 91% of patients.

Diagnosis

A chest X-ray with the patient in the upright position can diagnose a hiatal hernia if it shows an air-fluid level behind the cardiac shadow. This is usually caused by a PEH or an intrathoracic stomach. The accuracy of the upper GI barium study in detecting a paraesophageal hiatal hernia is greater than for a sliding hernia because the latter can often spontaneously reduce. The paraesophageal hiatal hernia is a permanent herniation of the stomach into the thoracic cavity, so a barium swallow provides the diagnosis in virtually every case. Attention should be focused on the position of the GEJ, when seen, to differentiate it from a type II hernia (see Fig. 25-39B and C). Fiber-optic esophagoscopy is useful in the diagnosis and classification of a hiatal hernia because the scope can be refluxed. In this position, a sliding hiatal hernia can be identified by noting a gastric pouch lined with rugal folds extending above the impression caused by the crura of the diaphragm, or measuring at least 2 cm between the crura, identified by having the patient sniff, and the squamo-columnar junction on withdrawal of the scope (Fig. 25-40). A PEH is identified on retroversion of the scope by noting a separate orifice adjacent to the GEJ into which gastric rugal folds ascend. A sliding-rolling or mixed hernia can be identified by noting a gastric pouch lined with rugal folds above the diaphragm, with the GEJ entering about midway up the side of the pouch.
Pathophysiology

Physiologic testing with 24-hour esophageal pH monitoring has shown increased esophageal exposure to acid gastric juice in 60% of the patients with a paraesophageal hiatal hernia, compared with the observed 71% incidence in patients with a sliding hiatal hernia. It is now recognized that paraesophageal hiatal hernia can be associated with pathologic gastroesophageal reflux.

Physiologic studies have also shown that the competency of the cardia depends on an interrelationship between distal esophageal sphincter pressure, the length of the sphincter that is exposed to the positive-pressure environment of the abdomen, and the overall length of the sphincter. A deficiency in any one of these manometric characteristics of the sphincter is associated with incompetency of the cardia regardless of whether a hernia is present. Patients with a PEH who have an incompetent cardia have been shown to have a distal esophageal sphincter with normal pressure, but a shortened overall length and displacement outside the positive-pressure environment of the abdomen. One might expect esophageal body function to be diminished with the esophagus “accordioned” up into the chest. Surprisingly, esophageal peristalsis in patients with PEH is normal in 88%.

Treatment

The treatment of paraesophageal hiatal hernia is largely surgical. Controversial aspects include: (a) indications for repair, (b) diaphragmatic repair, (c) role of fundoplication, and (d) existence and treatment of the short esophagus.

Indications and Surgical Approach. The presence of a paraesophageal hiatal hernia has traditionally been considered an indication for surgical repair. This recommendation is largely based upon two clinical observations. First, retrospective studies have shown a significant incidence of catastrophic, life-threatening complications of bleeding, infarction, and perforation in patients being followed with known paraesophageal herniation. Second, emergency repair carries a high mortality. In the classic report of Skinner and Belsey, six of 21 patients with a PEH, treated medically because of minimal symptoms, died from the complications of strangulation, perforation, exsanguinating hemorrhage, or acute dilatation of the herniated intrathoracic stomach. For the most part, these catastrophes occurred without warning. Others have reported similar findings.

Recent studies suggest that catastrophic complications may be somewhat less common. Allen and colleagues followed 23 patients for a median of 78 months with only four patients progressively worsening. There was a single mortality secondary to aspiration that occurred during a barium swallow examination to investigate progressive symptoms. Although emergency repairs had a median hospital stay of 48 days compared to a stay of 9 days in those having elective repair, there were only three cases of gastric strangulation in 735 patient-years of follow-up.

If surgery is delayed and repair is done on an emergency basis, operative mortality is high, compared to <1% for an elective repair. With this in mind, patients with a PEH are generally counseled to have elective repair of their hernia, particularly if they are symptomatic. Watchful waiting of asymptomatic PEHs may be an acceptable option.

The surgical approach to repair of a paraesophageal hiatal hernia may be either transabdominal (laparoscopic or open) or transthoracic. Each has its advantages and disadvantages. A transthoracic approach facilitates complete esophageal mobilization but is rarely used because the access trauma and postoperative pain are significantly greater than a laparoscopic approach.

The transabdominal approach facilitates reduction of the volvulus that is often associated with PEHs. Although some degree of esophageal mobilization can be accomplished transthoracally, complete mobilization to the aortic arch is difficult or impossible without risk of injury to the vagal nerves.

Laparoscopic repair of PEH would appear to have become the standard approach. Laparoscopic repair of a pure type II, or mixed type III PEH is an order of magnitude more difficult than a standard laparoscopic Nissen fundoplication. Most would recommend that these procedures are best avoided until the surgeon has accumulated considerable experience with laparoscopic antireflux surgery. There are several reasons for this. First, the vertical and horizontal volvulus of the stomach often associated with PEHs makes identification of the anatomy, in particular the location of the esophagus, difficult. Second, dissection of a large PEH sac may result in significant bleeding if the surgeon deviates from the correct plane of dissection between the peritoneal sac and the endothoracic fascia. Finally, redundant tissue present at the GEJ following dissection of the sac frustrates the creation of a fundoplication. This tissue, which includes the epiphenic fat pad and hernia sac should be removed at the time of PEH repair. Mindful of these difficulties, and given appropriate experience, patients with PEH may be approached laparoscopically, with expectation of success in the majority.

Diaphragmatic Repair

It has been shown that PEH repair has a relatively high incidence of recurrence (10–40%) when the crura is closed primarily with permanent suture. Techniques to reduce hernia recurrence continue to evolve. Most surgeons believe that recurrence may be reduced with the use of synthetic or biologic mesh to reinforce the standard crural closure. Randomized controlled studies have
demonstrated a reduction in PEH recurrence rate when mesh was used. Nonabsorbable synthetic mesh must be used carefully and not in a keyhole fashion at the hiatus because of a potential risk of esophagus or gastric erosion and mesh infection. Biologic mesh (acellular porcine dermis, acellular human dermis, porcine small intestinal submucosa) has become more widely used, but these meshes are significantly more expensive than synthetic mesh, and the only randomized study supporting biologic mesh usage failed to demonstrate superiority over suture alone after 5 years of rigorous follow-up.

**Role of Fundoplication in Giant Hiatal Hernia Repair.** Controversy remains as to whether to perform an antireflux procedure at all, in selected cases only, or in all patients. Most advocate the routine addition of an antireflux procedure following repair of the hernia defect. There are several reasons for this. Physiologic testing with 24-hour esophageal pH monitoring has shown increased esophageal exposure to acid gastric juice in 60% to 70% of patients with a paraesophageal hiatal hernia, nearly identical to the observed 71% incidence in patients with a sliding hiatal hernia. Furthermore, there is no relation between the symptoms experienced by the patient with a PEH and the competency of the cardia. Finally, dissection of the gastroesophageal esophagus may lead to postoperative reflux despite a negative preoperative pH score.

**The Short Esophagus and PEH**

Giant PEH can be associated with a short esophagus in up to 5% to 20% of patients as a result of chronic cephalad displacement of the GEJ. The presence of a short esophagus increases the difficulty of laparoscopic PEH repair. Approximately 10% to 20% of surgical failures with PEH repair is due to the lack of recognition of a short esophagus. Preoperative results of barium swallow and esophagogastroduodenoscopy may provide an indication of short esophagus, but no combination of preoperative clinical variables reliably predict the presence of short esophagus, defined as the failure to achieve 2.5 cm of intra-abdominal esophagus with standard mediastinal dissection techniques. Hence, the diagnosis of this entity continues to be made definitively only in the operating room. Collis gastroplasty achieves esophageal lengthening by creation of a neoesophagus using the gastric cardia. The totally laparoscopic approach to the short esophagus has evolved from a method using an end-to-end anastomosis circular stapler to the current approach that uses a linear stapler creating a stapled wedge gastroplasty. Elements of importance in fashioning the fundoplication after Collis gastroplasty include placement of the initial suture of the fundoplication on the esophagus, immediately above the GEJ to ensure that acid-secreting (gastric) mucosa does not reside above the fundoplication. A second element that ensures safety and avoids wrap deformation is to place the gastric portion of the staple line against the neoesophagus, such that the tip of the gastric staple line sits adjacent to the middle suture of the fundoplication on the right side of the esophagus.

**Results**

Most outcome studies report relief of symptoms following surgical repair of PEHs in more than 90% of patients. The current literature suggests that laparoscopic repair of a paraesophageal hiatal hernia can be successful. Most authors report symptomatic improvement in 80% to 90% of patients, and <10% to 15% prevalence of recurrent symptomatic hernia. However, the problem of recurrent asymptomatic or minimally symptomatic hernia following PEH repair, open or laparoscopic, is becoming increasingly appreciated. Recurrent hiatal hernia is the most common cause of anatomic failure following laparoscopic Nissen fundoplication done for GERD (5–10%), but this risk is compounded for the giant hernia where radiologic recurrence is detected in 25% to 40% of patients. It appears that optimal results with open or laparoscopic giant hiatal hernia repair should include options for mesh buttressing of hiatal closure and selective esophageal lengthening with one of the many techniques developed for the creation of a Collis gastroplasty. Despite this high incidence of radiologic recurrence, and the surgical pursuit of a remedy, it must be reinforced that asymptomatic recurrent hernias, like primary PEH, do not need to be repaired. The risk of incarceration, strangulation, or obstruction is minimal.

**SCHATZKI’S RING**

Schatzki’s ring is a thin submucosal circumferential ring in the lower esophagus at the squamocolumnar junction, often associated with a hiatal hernia. Its significance and pathogenesis are unclear (Fig. 25-41). The ring was first noted by Templeton, but Schatzki and Gary defined it as a distinct entity in 1953. Its prevalence varies from 0.2% to 14% in the general population, depending on the technique of diagnosis and the criteria used. Stennon believed the ring to be a pleat of mucosa formed by infolding of redundant esophageal mucosa due to shortening of the esophagus. Others believe the ring to be congenital, and still others suggest it is an early stricture resulting from inflammation of the esophageal mucosa caused by chronic reflux. Schatzki’s ring is a distinct clinical entity having different symptoms, upper GI function studies, and response to treatment compared with patients with a hiatal hernia, but without a ring. Twenty-four-hour esophageal pH monitoring has shown that patients with a Schatzki’s ring have a lower incidence of reflux than hiatal hernia controls. They also have better LES function. This, together with the presence of a ring, could represent a protective mechanism to prevent gastroesophageal reflux.

![Figure 25-41. Barium esophagogram showing Schatzki’s ring (i.e., a thin circumferential ring in the distal esophagus at the squamocolumnar junction). Below the ring is a hiatal hernia.](image-url)
Symptoms associated with Schatzki’s ring are brief episodes of dysphagia during hurried ingestion of solid foods. Its treatment has varied from dilation alone to dilation with antireflux measures, antireflux procedure alone, incision, and even excision of the ring. Little is known about the natural progression of Schatzki’s rings. Using radiologic techniques, Chen and colleagues showed progressive stenosis of rings in 59% of patients, whereas Schatzki found that the rings decreased in diameter in 29% of patients and remained unchanged in the rest.

Symptoms in patients with a ring are caused more by the presence of the ring than by gastroesophageal reflux. Most patients with a ring but without proven reflux respond to one dilation, while most patients with proven reflux require repeated dilations. In this regard, the majority of Schatzki’s ring patients without proven reflux have a history of ingestion of drugs known to be damaging to the esophageal mucosa. Bonavina and associates have suggested drug-induced injury as the cause of stenosis in patients with a ring, but without a history of reflux. Because rings also occur in patients with proven reflux, it is likely that gastroesophageal reflux also plays a part. This is supported by the fact that there is less drug ingestion in the history of these patients. Schatzki’s ring is probably an acquired lesion that can lead to stenosis from chemical-induced injury by pill lodgment in the distal esophagus, or from reflux-induced injury to the lower esophageal mucosa.

The best form of treatment of a symptomatic Schatzki’s ring in patients who do not have reflux consists of esophageal dilation for relief of the obstructive symptoms. In patients with a ring who have proven reflux and a mechanically defective sphincter, an antireflux procedure is necessary to obtain relief and avoid repeated dilation.

**SCLERODERMA**

Scleroderma is a systemic disease accompanied by esophageal abnormalities in approximately 80% of patients. In most, the disease follows a prolonged course. Renal involvement occurs in a small percentage of patients and signals a poor prognosis. The onset of the disease is usually in the third or fourth decade of life, occurring twice as frequently in women as in men.

Small vessel inflammation appears to be an initiating event, with subsequent perivascular deposition of normal collagen, which may lead to vascular compromise. In the GI tract, the predominant feature is smooth muscle atrophy. Whether the atrophy in the esophageal musculature is a primary effect or occurs secondary to a neurogenic disorder is unknown. The results of pharmacologic and hormonal manipulation, with agents that act either indirectly via neural mechanisms or directly on the muscle, suggest that scleroderma is a primary neurogenic disorder. Methacholine, which acts directly on smooth muscle receptors, causes a similar increase in LES pressure in normal controls and in patients with scleroderma. Erophonium, a cholinesterase inhibitor that enhances the effect of acetylcholine when given to patients with scleroderma, causes an increase in LES pressure that is less marked in these patients than in normal controls, suggesting a neurogenic rather than myogenic etiology. Muscle ischemia due to perivascular compression has been suggested as a possible mechanism for the motility abnormality in scleroderma. Others have observed that in the early stage of the disease, the manometric abnormalities may be reversed by reserpine, an agent that depletes catecholamines from the adrenergic system. This suggests that, in early scleroderma, an adrenergic overactivity may be present that causes a parasympathetic inhibition, supporting a neurogenic mechanism for the disease. In advanced disease manifested by smooth muscle atrophy and collagen deposition, reserpine no longer produces this reversal. Consequently, from a clinical perspective, the patient can be described as having a poor esophageal pump and a poor valve.

The diagnosis of scleroderma can be made manometrically by the observation of normal peristalsis in the proximal striated esophagus, with absent peristalsis in the distal smooth muscle portion (Fig. 25-42). The LES pressure is progressively weakened as the disease advances. Because many of the systemic sequelae of the disease may be nondiagnostic, the motility pattern is frequently used as a specific diagnostic indicator. Gastroesophageal reflux commonly occurs in patients with scleroderma because they have both hypertensive sphincters and poor esophageal clearance. This combined defect can lead to severe esophagitis and stricture formation. The typical barium swallow shows a dilated, barium-filled esophagus, stomach, and duodenum, or a hiatal hernia with distal esophageal stricture and proximal dilatation (Fig. 25-43).

Traditionally, esophageal symptoms have been treated with PPIs, antacids, elevation of the head of the bed, and multiple dilations for strictures, with generally unsatisfactory results. The degree of esophagitis is usually severe and may lead to marked esophageal shortening as well as stricture. Scleroderma patients have frequently had numerous dilations before they are referred to the surgeon. The surgical management is somewhat controversial, but the majority of opinion suggests that a partial fundoplication (anterior or posterior) performed laparoscopically is the procedure of choice. The need for a partial fundoplication is dictated by the likelihood of severe dysphagia if a total fundoplication is performed in the presence of aperistalsis. Esophageal shortening may require a Collis gastroplasty in combination with a partial fundoplication. Surgery reduces esophageal acid exposure but does not return it to normal because of the poor
EOSINOPHILIC ESOPHAGITIS

Eosinophilic esophagitis (EE) was first described in 1977, but it has become well known only in the last two decades. The condition is characterized by a constellation of symptoms, endoscopic and radiologic findings, and distinctive pathology. The etiology of eosinophilic esophagitis is not entirely known but its similarities, immunologically, to asthma suggest that it is a form of “allergic esophagitis.”

Symptoms

The presentation of eosinophilic esophagitis is chest pain (often postprandial) and dysphagia. Dysphagia may occur with liquids or solids, but solid food dysphagia is most common. Because dysphagia and chest pain are characteristic of GERD, EE is often confused with GERD; however, EE does not respond to proton pump inhibitors. The evaluation of the patient with EE and dysphagia and chest pain with esophagram and endoscopy usually reveals the diagnosis.

Signs

A barium swallow should be the first test obtained in the patient with dysphagia. EE has a characteristic finding often called the “ringed esophagus” or the “feline esophagus,” as the esophageal rings are felt to look like the stripes on a housecat (Fig. 25-44). The endoscopic appearance of EE is also characteristic, and also appears as a series of rings (Fig. 25-45).

Pathology

Endoscopic biopsy specimens should be taken when eosinophilic esophagitis is suspected. To make the diagnosis of EE, the pathologist should see a minimum of 15 eosinophils per high powered field, usually at the base of the epithelium (Fig. 25-46).

Treatment

The treatment of EE is largely symptomatic and includes testing for food allergies and elimination of identified items from the diet. Second-line therapy includes inhaled or ingested corticosteroids, as would be used to treat asthma. If dysphagia is not relieved with steroids, it may be necessary to dilate the clearance function of the body of the esophagus. Only 50% of the patients have a good-to-excellent result. If the esophagitis is severe, or there has been a previous failed antireflux procedure and the disease is associated with delayed gastric emptying, a gastric resection with Roux-en-Y gastrojejunostomy has proved the best option.

Figure 25-43. Barium esophagogram of a patient with scleroderma and stricture. Note the markedly dilated esophagus and retained food material. (Reproduced with permission from Waters PF, DeMeester TR: Foregut motor disorders and their surgical management, Med Clin North Am. 1981 Nov;65(6):1235-1268.)

Figure 25-44. The esophagus on the left shows a stacking of rings, demonstrating eosinophilic esophagus. The esophagus on the right is a normal barium swallow.
esophagus. Because of the length of esophageal involvement, rigid dilators (Maloney or Savary) are often used. Great care must be exercised, as the inflamed EE is quite friable. The mucosal tears easily, and esophageal perforation (full thickness laceration) has been reported with EE dilation.

MOTILITY DISORDERS OF THE PHARYNX AND ESOPHAGUS

Clinical Manifestations
Dysphagia (i.e., difficulty in swallowing) is the primary symptom of esophageal motor disorders. Its perception by the patient is a balance between the severity of the underlying abnormality causing the dysphagia and the adjustment made by the patient in altering eating habits. Consequently, any complaint of dysphagia must include an assessment of the patient’s dietary history. It must be known whether the patient experiences pain, chokes, or vomits with eating; whether the patient requires liquids with the meal, is the last to finish, or is forced to interrupt or avoid a social meal; and whether he or she has been admitted to the hospital for food impaction. These assessments, plus an evaluation of the patient’s nutritional status, help to determine how severe the dysphagia is and judge the need for surgical intervention, rather than more conservative methods of treating dysphagia.

Motility Disorders of the Pharynx and Upper Esophagus—Transit Dysphagia
Disorders of the pharyngeal phase of swallowing result from a discoordination of the neuromuscular events involved in chewing, initiation of swallowing, and propulsion of the material from the oropharynx into the cervical esophagus. They can be categorized into one or a combination of the following abnormalities: (a) inadequate oropharyngeal bolus transport; (b) inability to pressurize the pharynx; (c) inability to elevate the larynx; (d) discoordination of pharyngeal contraction and cricopharyngeal relaxation; and (e) decreased compliance of the pharyngo-esophageal segment secondary to neuromuscular disease. The latter may result in incomplete relaxation of the cricopharyngeus and cervical esophagus during swallowing. Taken together, these disorders are termed transit dysphagia by many.

Transit dysphagia is usually congenital or results from acquired disease involving the central and peripheral nervous system. This includes cerebrovascular accidents, brain stem tumors, poliomyelitis, multiple sclerosis, Parkinson’s disease, pseudobulbar palsy, peripheral neuropathy, and operative damage to the cranial nerves involved in swallowing. Pure muscular diseases such as radiation-induced myopathy, dermatomyositis, myotonic dystrophy, and myasthenia gravis are less common causes. Rarely, extrinsic compression of the cervical esophagus by thyromegaly, lymphadenopathy, or hyperostosis of the cervical spine can cause transit dysphagia.

Diagnostic Assessment of the Cricopharyngeal Segment
Transit dysphagia difficult to assess with standard manometric techniques because of the rapidity of the oropharyngeal phase of swallowing, the elevation of the larynx, and the asymmetry of the cricopharyngeus. Video- or cineradiography is currently the
The most objective test to evaluate oropharyngeal bolus transport, pharyngeal compression, relaxation of the pharyngoesophageal segment, and the dynamics of airway protection during swallowing. It readily identifies a diverticulum (Fig. 25-47), stasis of the contrast medium in the valleculae, a cricopharyngeal bar, and/or narrowing of the pharyngoesophageal segment. These are anatomic manifestations of neuromuscular disease, and they result from the loss of muscle compliance in portions of the pharynx and esophagus composed of skeletal muscle.

Careful analysis of video- or cineradiographic studies combined with manometry using specially designed catheters can identify the cause of a pharyngoesophageal dysfunction in most situations (Fig. 25-48). Motility studies may demonstrate inadequate pharyngeal pressurization, insufficient or lack of cricopharyngeal relaxation, marked discoordination of pharyngeal pressurization, cricopharyngeal relaxation and cervical esophageal contraction, or a hypopharyngeal bolus pressure suggesting decreased compliance of the skeletal portion of the cervical esophagus.

In many patients with cricopharyngeal dysfunction, including those with Zenker’s diverticulum, it has been difficult to consistently demonstrate a motility abnormality or discoordination of pharyngoesophageal events. The abnormality most apt to be present is a loss of compliance in the pharyngoesophageal segment manifested by an increased bolus pressure. Cook and colleagues have demonstrated an increased resistance to the movement of a bolus through what appears on manometry to be a completely relaxed cricopharyngeal sphincter. Using simultaneous manometry and videofluoroscopy, they showed that, in these patients, the cricopharyngeus is only partially relaxed; that is, the sphincter is relaxed enough to allow a drop of its pressure to esophageal baseline on manometry, but insufficiently relaxed to allow unimpaired passage of the bolus into the esophagus. This incomplete relaxation is due to a loss of compliance of the muscle in the pharyngoesophageal segment, and may be associated with a cricopharyngeal bar or Zenker’s diverticulum. This decreased compliance of the cricopharyngeal sphincter can be recognized on esophageal manometry by a “shoulder” on the pharyngeal pressure wave, the amplitude of which correlates directly with the degree of outflow obstruction (Fig. 25-49). Increasing the diameter of this noncompliant segment reduces the resistance imposed on the passage of a bolus. Consequently, patients with low pharyngeal pressure (i.e., poor piston function of the pharynx), or patients with increased resistance of the pharyngocervical esophageal segment from loss of skeletal muscle compliance, are improved by a cricopharyngeal myotomy. This enlarges the pharyngoesophageal segment and reduces outflow resistance. Esophageal muscle biopsy specimens from patients with Zenker’s diverticulum have shown histologic evidence of the restrictive myopathy in the cricopharyngeal muscle. These findings correlate well with the observation of a decreased compliance of the upper esophagus demonstrated by videoradiography and the findings on detailed manometric studies of the pharynx and cervical esophagus. They suggest that the diverticulum develops as a consequence of the outflow resistance to bolus transport through the noncompliant muscle of the pharyngoesophageal segment.

The requirements for a successful pharyngoesophageal myotomy are (a) adequate oropharyngeal bolus transport; (b) the presence of an intact swallowing reflex; (c) reasonable coordination of pharyngeal pressurization with cricopharyngeal relaxation; and (d) a cricopharyngeal bar, Zenker’s diverticulum, or a narrowed pharyngoesophageal segment on videofluorogram and/or the presence of excessive pharyngoesophageal shoulder pressure on motility study.

**Zenker’s Diverticulum.** In the past, the most common recognized sign of cricopharyngeal dysfunction was the presence of a
Zenker’s diverticulum, originally described by Ludlow in 1769. The eponym resulted from Zenker’s classic clinicopathologic descriptions of 34 cases published in 1878. Pharyngoesophageal diverticula have been reported to occur in 1 of 1000 routine barium examinations, and classically occur in elderly, white males. Zenker’s diverticula tend to enlarge progressively with time due to the decreased compliance of the skeletal portion of the cervical esophagus that occurs with aging.

Presenting symptoms include dysphagia associated with the spontaneous regurgitation of undigested, bland material, often interrupting eating or drinking. On occasion, the dysphagia can be severe enough to cause debilitation and significant weight loss. Chronic aspiration and repetitive respiratory infection are common associated complaints. Once suspected, the diagnosis is established by a barium swallow. Endoscopy is usually difficult in the presence of a cricopharyngeal diverticulum, and potentially dangerous, owing to obstruction of the true esophageal lumen by the diverticulum and the attendant risk of diverticular perforation.

**Cricopharyngeal Myotomy.** The low morbidity and mortality associated with cricopharyngeal and upper esophageal myotomy have encouraged a liberal approach toward its use for almost any problem in the oropharyngeal phase of swallowing. This attitude has resulted in an overall success rate in the relief of symptoms of only 64%. When patients are selected for surgery using radiographic or motility markers of disease, a much higher proportion will benefit. Two methods of cricopharyngeal myotomy are in common use, one using traditional surgical approaches, and one using rigid laryngoscopy and a linear cutting stapler.

**Open Cricopharyngeal Myotomy, Diverticulopexy, and Diverticulectomy.** The myotomy can be performed under local or general anesthesia through an incision along the anterior border of the left sternocleidomastoid muscle. The pharynx and cervical esophagus are exposed by retracting the sternocleidomastoid muscle and carotid sheath laterally and the thyroid, trachea, and larynx medially (Fig. 25-50). When a pharyngoesophageal diverticulum is present, localization of the pharyngoesophageal segment is easy. The diverticulum is carefully freed from the overlying areolar tissue to expose its neck, just below the inferior pharyngeal constrictor and above the cricopharyngeus muscle. It can be difficult to identify the cricopharyngeus muscle in the absence of a diverticulum. A benefit of local anesthesia is that the patient can swallow and demonstrate an area of persistent narrowing at the pharyngoesophageal junction. Furthermore, before closing the incision, gelatin can be fed to the patient to ascertain whether the symptoms have been relieved, and to inspect the opening of the previously narrowed pharyngoesophageal segment. Under general anesthesia, and in the absence of a diverticulum, the placement of a nasogastric tube to the level of the manometrically determined cricopharyngeal sphincter helps in localization of the structures. The myotomy is extended cephalad by dividing 1 to 2 cm of inferior constrictor muscle of the pharynx, and caudad by dividing the cricopharyngeal muscle and the cervical esophagus for a length of 4 to 5 cm. The cervical wound is closed only when all oozing of blood has ceased because a hematoma after this procedure is common and is often associated with temporary dysphagia while the hematoma absorbs. Oral alimentation is started the day after surgery. The patient is usually discharged on the first or second postoperative day.
Esophagus and Diaphragmatic Hernia

If a diverticulum is present and is large enough to persist after a myotomy, it may be sutured in the inverted position to the prevertebral fascia using a permanent suture (i.e., diverticulopexy) (Fig. 25-51). If the diverticulum is excessively large so that it would be redundant if suspended, or if its walls are thickened, a diverticulectomy should be performed. This is best performed under general anesthesia by placing a Maloney dilator (48F) in the esophagus, after controlling the neck of the diverticulum and after myotomy. A linear stapler is placed across the neck of the diverticulum, and the diverticulum is excised distal to the staple line. The security of this staple line and effectiveness of the myotomy may be tested before hospital discharge with a water-soluble contrast esophagogram. Postoperative complications include fistula formation, abscess, hematoma, recurrent nerve paralysis, difficulties in phonation, and Horner’s syndrome. The incidence of the first two can be reduced by performing a diverticulopexy rather than diverticulectomy.

Endoscopic Cricopharyngotomy. Endoscopic stapled cricopharyngotomy and diverticulectomy recently has been described. This procedure is most effective for larger diverticula (>2 cm) and may be impossible to perform for the small diverticulum. The procedure uses a specialized “diverticuloscope” with two retractable valves passed into the hypopharynx. The lips of the diverticuloscope are positioned so that one lip lies in the esophageal lumen and the other in the diverticular lumen. The valves of the diverticuloscope are retracted appropriately so as to visualize the septum interposed between the diverticulum and the esophagus. An endoscopic linear stapler is introduced into the diverticuloscope and positioned against the common septum with the anvil in the diverticulum and the cartridge in the esophageal lumen. Firing of the stapler divides the common septum between the posterior esophageal and the diverticular wall over a length of 30 mm, placing three rows of staples on each side. More than one stapler application may be needed, depending on the size of the diverticulum (Fig. 25-52). The patient is allowed to resume liquid feeds immediately and is usually discharged the day after surgery. Complications are rare and may include perforation at the apex of the diverticulum and failure to relieve dysphagia resulting from incomplete myotomy. The former complication can usually be treated with antibiotics, but it may, rarely, require neck drainage.

Recurrence of a Zenker’s diverticulum may occur with long follow-up and is more common after diverticulectomy without myotomy, presumably due to persistence of the underlying loss of compliance of the cervical esophagus when a myotomy is not performed. After endoscopic cricopharyngotomy lateral residual “pouches” may be seen on radiographs, but they are rarely responsible for residual or recurrent symptoms if the myotomy has been complete.

Postoperative motility studies have shown that the peak pharyngeal pressure generated on swallowing is not affected, the resting cricopharyngeal pressure is reduced but not eliminated, and the cricopharyngeal sphincter length is shortened. Consequently, after myotomy, there is protection against esophagopharyngeal regurgitation.

Motility Disorders of the Esophageal Body and Lower Esophageal Sphincter
Disorders of the esophageal phase of swallowing result from abnormalities in the propulsive pump action of the esophageal body or the relaxation of the LES. These disorders result from either primary esophageal abnormalities, or from generalized neural, muscular, or collagen vascular disease (Table 25-8). The use of standard and high-resolution esophageal manometry techniques has allowed specific primary esophageal motility disorders to be identified out of a pool of nonspecific motility abnormalities. Primary esophageal motor disorders include achalasia, DES, nutcracker esophagus, and the hypertensive LES. The manometric characteristics of these disorders are shown in Table 25-9.

The boundaries between the primary esophageal motor disorders are vague, and intermediate types exist, some of which may combine more than one type of motility pattern. These findings indicate that esophageal motility disorders should be looked at as a spectrum of abnormalities that reflects various stages of destruction of esophageal motor function.

Achalasia. The best known and best understood primary motility disorder of the esophagus is achalasia, with an incidence of six
Table 25-8

**Esophageal motility disorders**

<table>
<thead>
<tr>
<th>Primary esophageal motility disorders</th>
<th>Secondary esophageal motility disorders</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achalasia, “vigorous” achalasia</td>
<td>Collagen vascular diseases: progressive systemic sclerosis, polymyositis and dermatomyositis, mixed connective tissue disease, systemic lupus erythematosus, etc.</td>
</tr>
<tr>
<td>Diffuse and segmental esophageal spasm</td>
<td>Chronic idiopathic intestinal pseudo-obstruction</td>
</tr>
<tr>
<td>Nutcracker esophagus</td>
<td>Neuroromuscular diseases</td>
</tr>
<tr>
<td>Hypertensive lower esophageal sphincter</td>
<td>Endocrine and metastatic disorders</td>
</tr>
<tr>
<td>Non-specific esophageal motility disorders</td>
<td></td>
</tr>
</tbody>
</table>

Table 25-9

**Manometric characteristics of the primary esophageal motility disorders**

<table>
<thead>
<tr>
<th>Achalasia</th>
<th>Diffuse esophageal spasm (DES)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incomplete lower esophageal sphincter (LES) relaxation (&lt;75% relaxation)</td>
<td>Simultaneous (nonperistaltic contractions) (&gt;20% of wet swallows)</td>
</tr>
<tr>
<td>Aperistalsis in the esophageal body</td>
<td>Repetitive and multipeaked contractions</td>
</tr>
<tr>
<td>Elevated LES pressure ≥26 mmHg</td>
<td>Spontaneous contractions</td>
</tr>
<tr>
<td>Increased intraesophageal baseline pressures relative to</td>
<td>Intermittent normal peristalsis</td>
</tr>
<tr>
<td>gastric baseline</td>
<td>Contractions may be of increased amplitude and duration</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nutcracker esophagus</th>
<th>Hypertensive lower esophageal sphincter</th>
<th>Ineffective esophageal motility disorders</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean peristaltic amplitude (10 wet swallows) in distal esophagus ≥180 mmHg</td>
<td>Elevated LES pressure (≥26 mmHg)</td>
<td>Decreased or absent amplitude of esophageal peristalsis (&lt;30 mmHg)</td>
</tr>
<tr>
<td>Increased mean duration of contractions (&gt;7.0 s)</td>
<td>Normal LES relaxation</td>
<td>Increased number of nontransmitted contractions</td>
</tr>
<tr>
<td>Normal peristaltic sequence</td>
<td>Normal peristalsis in the esophageal body</td>
<td></td>
</tr>
</tbody>
</table>

Simultaneous esophageal waves develop as a result of the increased resistance to esophageal emptying caused by the nonrelaxing LES. This conclusion is supported by experimental studies in which a band placed loosely around the GEJ in experimental models did not change sphincter pressures but resulted in impaired relaxation of the LES and outflow resistance. This led to a markedly increased frequency of simultaneous waveforms and a decrease in contraction amplitude. The changes were associated with radiographic dilation of the esophagus and were reversible after removal of the band. Observations in patients with pseudochaclasia due to tumor infiltration, a tight stricture in the distal esophagus, or an antireflux procedure that is too tight also provide evidence that dysfunction of the esophageal body can be caused by the increased outflow obstruction of a nonrelaxing LES. The observation that esophageal peristalsis can return in patients with classic achalasia following dilation or myotomy provides further support that achalasia is a primary disease of the LES.

The pathogenesis of achalasia is presumed to be a neurogenic degeneration, which is either idiopathic or due to infection. In experimental animals, the disease has been reproduced by destruction of the nucleus ambiguus and the dorsal motor nucleus of the vagus nerve. In patients with the disease, degenerative changes have been shown in the vagus nerve and in the ganglia in the myenteric plexus of the esophagus itself. This degeneration results in hypertension of the LES, a failure of the sphincter to relax on swallowing, elevation of intraluminal esophageal pressure, esophageal dilatation, and a subsequent loss of progressive peristalsis in the body of the esophagus. The esophageal dilatation results from the combination of a nonrelaxing sphincter, which causes a functional retention of ingested material in the esophagus, and elevation of intraluminal pressure from repetitive pharyngeal air swallowing (Fig. 25-53). With time, the functional disorder results in anatomic alterations seen on radiographic studies, such as a dilated esophagus with a tapering, “bird’s beak”-like narrowing of the distal end (Fig. 25-54). There is usually an air-fluid level in the esophagus from the retained food and saliva, the height of which reflects the degree of resistance imposed by the nonrelaxing sphincter. As the disease progresses, the esophagus becomes massively dilated and tortuous.

A subgroup of patients with otherwise typical features of classic achalasia has simultaneous contractions of their esophageal body that can be of high amplitude. This manometric pattern has been termed vigorous achalasia, and chest pain episodes are a common finding in these patients. Since the development of high resolution esophageal manometry technology, the term vigorous achalasia has been replaced with Chicago type 3 achalasia. Differentiation of type 3 achalasia from DES can be difficult. In both diseases, videoradiographic examination may show a corkscrew deformity of the esophagus and diverticulum formation.

**Diffuse and Segmental Esophageal Spasm.** DES is characterized by substernal chest pain and/or dysphagia. DES differs from classic achalasia in that it is primarily a disease of the esophageal body, produces a lesser degree of dysphagia, causes more chest pain, and has less effect on the patient’s general condition. Nonetheless, it is impossible to differentiate achalasia from DES on the basis of symptoms alone. Esophagogram and esophageal manometry are required to distinguish these two entities. True symptomatic DES is a rare condition, occurring about five times less frequently than achalasia.

The causation and neuromuscular pathophysiology of DES are unclear. The basic motor abnormality is rapid wave progression down the esophagus secondary to an abnormality in
the latency gradient. Hypertrophy of the muscular layer of the esophageal wall and degeneration of the esophageal branches of the vagus nerve have been observed in this disease, although these are not constant findings. Manometric abnormalities in DES may be present over the total length of the esophageal body but usually are confined to the distal two-thirds. In segmental esophageal spasm, the manometric abnormalities are confined to a short segment of the esophagus.

The classic manometric findings in these patients are characterized by the frequent occurrence of simultaneous waveforms and multipeaked esophageal contractions, which may be of abnormally high amplitude or long duration. Key to the diagnosis of DES is that there remain some peristaltic waveforms in excess of those seen in achalasia. A criterion of 30% or more peristaltic waveforms out of 10 wet swallows has been used to differentiate DES from vigorous achalasia. However, this figure is arbitrary and often debated.

The LES in patients with DES usually shows a normal resting pressure and relaxation on swallowing. A hypertensive sphincter with poor relaxation may also be present. In patients with advanced disease, the radiographic appearance of tertiary contractions appears helical and has been termed corkscrew esophagus or pseudodiverticulosis (Fig. 25-55). Patients with segmental or diffuse esophageal spasm can compartmentalize the esophagus and develop an epiphrenic or midesophageal diverticulum between two areas of high pressure occurring simultaneously (Fig. 25-56).

**Nutcracker Esophagus.** The disorder, termed nutcracker or supersqueezeresophagus, was recognized in the late 1970s. Other terms used to describe this entity are hypertensive peristalsis or high-amplitude peristaltic contractions. It is the most common of the primary esophageal motility disorders. By definition the so-called nutcracker esophagus is a manometric abnormality in patients who are characterized by peristaltic esophageal contractions with peak amplitudes greater than two SDs above the normal values in individual laboratories. Contraction amplitudes in these patients can easily be above 400 mmHg. At the lower end of peak pressure, it is unclear whether nutcracker esophagus causes any symptoms. In fact, chest pain symptoms in nutcracker esophagus patients may be related to GERD rather than intraluminal hypertension. Treatment in these patients should be aimed at the treatment of GERD. At the high end (peak pressures >300 mmHg) chest pain may be the result of the nutcracker physiology, as treatment directed at reducing intraluminal pressure is more effective than when used for those with lower peak pressures.

**Hypertensive Lower Esophageal Sphincter.** Hypertensive lower esophageal sphincter (LES) in patients with chest pain or dysphagia was first described as a separate entity by Code and associates. This disorder is characterized by an elevated basal pressure of the LES with normal relaxation and
normal propulsion in the esophageal body. About one-half of these patients, however, have associated motility disorders of the esophageal body, particularly hypertensive peristalsis and simultaneous waveforms. In the remainder, the disorder exists as an isolated abnormality. Dysphagia in these patients may be caused by a lack of compliance of the sphincter, even in its relaxed state. Myotomy of the LES may be indicated in patients not responding to medical therapy or dilation. When the symptom contribution of the hypertensive sphincter is in doubt, it is possible to inject the LES with botulinum toxin, endoscopically. If symptoms are relieved (temporarily) with this technique, then it is likely that myotomy will provide more permanent benefit.

**Secondary Esophageal Motility Disorders.** Connective tissue disease, particularly scleroderma and the CREST syndrome, exhibits severe esophageal motility disorders. Additionally, patients treated as infants for esophageal atresia will often develop secondary motility disorders manifest later in life. Symptoms of these disorders are heartburn and dysphagia. The latter may be a result of a peptic stricture rather than the esophageal dysmotility. An esophageal motility study will usually show severely reduced or absent peristalsis with severely reduced or absent LES pressure. The role of antireflux surgery under these conditions is controversial but, if performed, should be limited to partial fundoplication, as full (Nissen) fundoplication may result in severe dysphagia.

**Nonspecific Esophageal Motor Disorders and Ineffective Esophageal Motility.** Many patients complaining of dysphagia or chest pain of noncardiac origin demonstrate a variety of wave patterns and contraction amplitudes on esophageal manometry that are clearly out of the normal range, but do not meet the criteria of a primary esophageal motility disorder. Esophageal motility in these patients frequently shows an increased number of multiphasic or repetitive contractions, contractions of prolonged duration, nontransmitted contractions, an interruption of a peristaltic wave at various levels of the esophagus, or contractions of low amplitude. These motility abnormalities have been termed *nonspecific esophageal motility disorders*. Their significance in the causation of chest pain or dysphagia is still unclear. Surgery plays no role in the treatment of these disorders unless there is an associated diverticulum.

A clear distinction between primary esophageal motility disorders and nonspecific esophageal motility disorders is often not possible. Patients diagnosed as having nonspecific esophageal motility abnormalities on repeated studies will occasionally show abnormalities consistent with nutcracker esophagus. Similarly, progression from a nonspecific esophageal motility disorder to classic DES has been demonstrated. Therefore, the finding of a nonspecific esophageal motility disorder may represent only a manometric marker of an intermittent, more severe esophageal motor abnormality. Combined ambulatory 24-hour esophageal pH and motility monitoring has shown that an increased esophageal exposure to gastric juice is common in patients diagnosed as having a nonspecific esophageal motility disorder. In some situations, the motor abnormalities may be induced by the irritation of refluxed gastric juice; in other situations, it may be a primary event unrelated to the presence of reflux. High-amplitude peristalsis (nutcracker esophagus) and low-amplitude peristalsis (ineffective esophageal motility) are frequently associated with GERD.

**Diverticula of the Esophageal Body.** Diverticula of the esophagus may be characterized by their location in the esophagus (proximal, mid-, or distal esophagus), or by the nature of
concomitant pathology. Diverticula associated with motor disorders are termed *pulsion diverticula* and those associated with inflammatory conditions are termed *traction diverticula*. Pulsion diverticula occur most commonly with nonspecific motility disorders, but they can occur with all of the primary motility disorders. In the latter situation, the motility disorder is usually diagnosed before the development of the diverticulum. When associated with achalasia, the development of a diverticulum may temporarily alleviate the symptom of dysphagia by becoming a receptacle for ingested food and substitute the symptom of dysphagia for postprandial pain and regurgitation of undigested food. If a motility abnormality of the esophageal body or LES cannot be identified, a traction or congenital cause for the diverticulum should be considered.

Because development in radiology preceded development in motility monitoring, diverticula of the esophagus were considered historically to be a primary abnormality, the cause, rather than the consequence, of motility disorders. Consequently, earlier texts focused on them as specific entities based upon their location.

Epiphrenic diverticula arise from the terminal third of the thoracic esophagus and are usually found adjacent to the diaphragm. They have been associated with distal esophageal muscular hypertrophy, esophageal motility abnormalities, and increased luminal pressure. They are “pulsion” diverticula, and they are associated with diffuse spasm, achalasia, or nonspecific motor abnormalities in the body of the esophagus.

Whether the diverticulum should be surgically resected or suspended depends on its size and proximity to the vertebral body. When diverticula are associated with esophageal motility disorders, esophageal myotomy from the proximal extent of the diverticulum to the stomach should be combined with diverticulectomy. If diverticulectomy alone is performed, one can expect a high incidence of suture line rupture due to the same intraluminal pressure that initially gave rise to the diverticulum. If the diverticulum is suspended to the prevertebral fascia of the thoracic vertebra, a myotomy is begun at the neck of the diverticulum and extended across the LES. If the diverticulum is excised by dividing the neck, the muscle is closed over the excision site, and a myotomy is performed on the opposite esophageal wall, starting just above the level of the diverticulum or at the proximal extent of the spastic segment of the esophagus if high resolution motility is used. If complete, the myotomy will cross the LES, reducing distal esophageal peak pressure, and it will increase the likelihood that dysphagia will be replaced with GERD symptoms. Increasingly, partial fundoplication (anterior or posterior) is performed after LES myotomy to decrease the frequency of disabling GERD developing after myotomy and diverticulectomy. When a large diverticulum is associated with a hiatal hernia, then hiatal hernia repair is added. All these procedures may be performed with traditional or minimally invasive techniques.

Midesophageal or traction diverticula were first described in the 19th century (Fig. 25-57). At that time, they were frequently noted in patients who had mediastinal LN involvement with tuberculosis. It was theorized that adhesions formed between the inflamed mediastinal nodes and the esophagus. By contraction, the adhesions exerted traction on the esophageal wall and led to a localized diverticulum (Fig. 25-58). This theory was based on the findings of early dissections, where adhesions between diverticula and LNs were commonly found. Other conditions associated with mediastinal lymphadenopathy, such as pulmonary fungal infections (e.g., aspergillosis), lymphoma, or sarcoid, may create traction esophageal diverticula after successful treatment. Rarely, when no underlying inflammatory pathology is identified, a motility disorder may be identified.

Most midesophageal diverticula are asymptomatic and incidentally discovered during investigation for nonesophageal complaints. In such patients, the radiologic abnormality may...
be ignored. Patients with symptoms of dysphagia, regurgitation, chest pain, or aspiration, in whom a diverticulum is discovered, should be thoroughly investigated for an esophageal motor abnormality. Occasionally, a patient will present with a bronchoesophageal fistula manifested by a chronic cough on ingestion of meals. The diverticulum in such patients is most likely to have an inflammatory etiology.

The indication for surgical intervention is dictated by the degree of symptomatic disability. Usually, midesophageal diverticula can be suspended due to their proximity to the spine. If a motor abnormality is documented, a myotomy should be performed as described for an epiphrenic diverticulum.

**OPERATIONS FOR ESOPHAGEAL MOTOR DISORDERS AND DIVERTICULA**

**Long Esophageal Myotomy for Motor Disorders of the Esophageal Body**

A long esophageal myotomy is indicated for dysphagia caused by any motor disorder characterized by segmental or generalized simultaneous waveforms in a patient whose symptoms are not relieved by medical therapy. Such disorders include diffuse and segmental esophageal spasm, vigorous or type 3 achalasia, and nonspecific motility disorders associated with a mid- or midesophageal esophageal diverticulum. However, the decision to operate must be made by a balanced evaluation of the patient’s symptoms, diet, lifestyle adjustments, and nutritional status, with the most important factor being the possibility of improving the patient’s swallowing disability. The symptom of chest pain alone is not an indication for a surgical procedure.

The identification of patients with symptoms of dysphagia and chest pain who might benefit from a surgical myotomy is difficult. Ambulatory motility studies have shown that when the prevalence of “effective contractions” (i.e., peristaltic waveforms consisting of contractions with an amplitude above 30 mmHg) drops below 50% during meals, the patient is likely to experience dysphagia (Fig. 25-59). This would suggest that relief from the symptom can be expected with an improvement of esophageal contraction amplitude or amelioration of nonperistaltic waveforms. Prokinetic agents may increase esophageal contraction amplitude, but they do not alter the prevalence of simultaneous waveforms. Patients in whom the efficacy of esophageal propulsion is severely compromised because of a high prevalence of simultaneous waveforms usually receive little benefit from medical therapy. In these patients, a surgical myotomy of the esophageal body can improve the patients’ dysphagia, provided the loss of contraction amplitude in the remaining peristaltic waveforms, caused by the myotomy, has less effect on swallowing function than the presence of the excessive simultaneous contractions. This situation is reached when the prevalence of effective waveforms during meals drops below 30% (i.e., 70% of esophageal waveforms are ineffective).

In patients selected for surgery, preoperative high-resolution manometry is essential to determine the proximal extent of the esophageal myotomy. Most surgeons extend the myotomy distally across the LES to reduce outflow resistance. Consequently, some form of antireflux protection is needed to avoid gastroesophageal reflux if there has been extensive dissection of the cardia. In this situation, most authors prefer a partial, rather than a full, fundoplication, in order not to add back-resistance that will further interfere with the ability of the myotomized esophagus to empty (Fig. 25-60). If the symptoms of reflux are present preoperatively, 24-hour pH monitoring is required to confirm its presence.

The procedure may be performed either open or via thoracoscopy. The open technique is performed through a left thoracotomy in the sixth intercostal space (Fig. 25-61). An incision is made in the posterior mediastinal pleura over the esophagus, and the left lateral wall of the esophagus is exposed. The esophagus is not circumferentially dissected unless necessary. A 2-cm incision is made into the abdomen through the parietal peritoneum at the midpoint of the left crus. A tongue of gastric fundus is pulled into the chest. This exposes the GEJ and its associated fat pad. The latter is excised to give a clear view of the junction. A myotomy is performed through all muscle layers, extending distally over the stomach 1 to 2 cm below the GEJ, and proximally on the esophagus over the distance of the manometric abnormality. The muscle layer is dissected from the mucosa laterally for a distance of 1 cm. Care is taken to divide all minute muscle bands, particularly in the area of the GEJ. The gastric fundic tongue is sutured to the margins of the myotomy over a distance of 3 to 4 cm and replaced into the abdomen. This maintains separation of the muscle and acts as a partial fundoplication to prevent reflux.
Figure 25-61. Technique of long myotomy: A. Exposure of the lower esophagus through the left sixth intercostal space and incision of the mediastinal pleura in preparation for surgical myotomy. B. Location of a 2-cm incision made through the phrenoesophageal membrane into the abdomen along the midlateral border of the left crus. C. Retraction of tongue of gastric fundus into the chest through the previously made incision. D. Removal of the gastroesophageal fat pad to expose the gastroesophageal junction. E. A myotomy down to the mucosa is started on the esophageal body. F. Completed myotomy extending over the stomach for 1 cm. G. Reconstruction of the cardia after a myotomy, illustrating the position of the sutures used to stitch the gastric fundic flap to the margins of the myotomy. H. Reconstruction of the cardia after a myotomy, illustrating the intra-abdominal position of the gastric tongue covering the distal 4 cm of the myotomy.
If an epiphrenic diverticulum is present, it is excised by dividing the neck with a stapler sized for the thickness of the diverticulum (2.0–4.8-mm staple leg length) followed by a closure of the muscle over the staple line, when possible. The myotomy is then performed on the opposite esophageal wall. If a midesophageal diverticulum is present, the myotomy is made so that it includes the muscle around the neck, and the diverticulum is suspended by attaching it to the paravertebral fascia of the thoracic vertebra above the level of the diverticular neck. Before performing any operation for an esophageal diverticulum, it is wise to endoscope the patient to wash all food and other debris from the diverticulum.

The results of myotomy for motor disorders of the esophageal body have improved in parallel with the improved preoperative diagnosis afforded by manometry. Previous published series report between 40% and 92% improvement of symptoms, but interpretation is difficult due to the small number of patients involved and the varying criteria for diagnosis of the primary motor abnormality. When myotomy is accurately done, 93% of the patients have effective palliation of dysphagia after a mean follow-up of 5 years, and 89% would have the procedure again, if it was necessary. Most patients gain or maintain rather than lose weight after the operation. Postoperative motility studies show that the myotomy reduces the amplitude of esophageal contractions to near zero and eliminates simultaneous peristaltic waves. If the benefit of obliterating the simultaneous waves exceeds the adverse effect on bolus propulsion caused by the loss of peristaltic waveforms, the patient’s dysphagia is likely to be improved by the procedure. If not, the patient is likely to continue to complain of dysphagia and to have little improvement as a result of the operation.

The thoracoscopic technique may be performed through the left or right chest. There has been little experience gained with doing adequate operations (as described previously with the open exposure) through left thoracoscopy, so most surgeons will combine a right thoracoscopic long myotomy with an abdominal approach for Heller myotomy and partial fundoplication. These two procedures may be done at the same setting, by double positioning the patient, or they may be done at two operations. If this is the case, it is best to do the abdominal component first, as the esophageal outflow obstruction is the source of most of the symptoms. Performing abdominal myotomy (and diverticulectomy, if present) may be all that is required.

A new procedure, peroral endoscopic myotomy (POEM) allows a long myotomy to be performed from the lumen of the esophagus with an endoscope. This procedure is attractive for, at a minimum, those with type 3 achalasia (vigorous achalasia), where it is necessary to divide esophagogastroduodenal circular muscle on both sides of the diaphragm to the extent that might not be possible with laparoscopy or thoracoscopy alone. The POEM procedure is started by opening the esophageal mucosa several centimeters above the spastic segment with a needle–knife electrosurgery device passed through an endoscope. A long submucosal plane is developed with the endoscope, down to and below the LES. The circular muscle of the LES and the esophagus is divided with endoscopic electrosurgery all the way back until normal (nonspastic) esophagus is reached. The submucosal entry site in the esophagus is then closed with endoscopic clips. While the results of POEM are still accumulating, the procedure is attractive because it is extremely minimally invasive and can be done on an outpatient basis.

Epiphrenic diverticula cannot be treated with POEM and are most frequently addressed with laparoscopic access, in combination with a laparoscopic division of the LES (Heller myotomy) (Fig. 25-62). If the diverticulum can be completely mobilized through the hiatus, it may be safely excised from below. The neck of the diverticulum is transected with a GIA stapler after passage of a 48F dilator. Not infrequently, the diverticulum is sufficiently large that access to the neck of the diverticulum across the hiatus is quite difficult. Additionally, the inflammatory reaction to the diverticulum may further make the transhiatal dissection difficult. Under these circumstances, it is safer to perform the diverticulectomy through a right thoracoscopic approach either at the time of the initial procedure or at a later date, depending upon the frailty of the patient. Following diverticulectomy, it is critical that the esophageal staple line be treated with a great deal of care. Closure of the muscle over the staple line is preferable. Additionally, the patient is kept NPO or on clear liquids for 5 to 7 days, and a contrast study is obtained before advancing to a full liquid or “mushy food” diet. Solid foods are withheld for 2 weeks to decrease the likelihood of staple line leak. But-tressing or sealing the staple line with fibrin glue is also an attractive option.
Figure 25-62. A. Epiphrenic diverticula are situated above the lower esophageal sphincter on right side of esophagus. B. Stapler amputates neck of diverticulum. C. Muscle reapproximated over staple line, and Heller myotomy is performed.

Myotomy of the Lower Esophageal Sphincter (Heller Myotomy)
Second only to reflux disease, achalasia is the most common functional disorder of the esophagus to require surgical intervention. The goal of treatment is to relieve the functional outflow obstruction secondary to the loss of relaxation and compliance of the LES. This requires disrupting the LES muscle. When performed adequately (i.e., reducing sphincter pressure to <10 mmHg), and done early in the course of disease, LES myotomy results in symptomatic improvement with the occasional return of esophageal peristalsis. Reduction in LES resistance can be accomplished intraluminally by hydrostatic balloon dilation, which ruptures the sphincter muscle, by botulinum toxin injection, or by a surgical myotomy that cuts the sphincter. The difference between these three methods appears to be the greater likelihood of reducing sphincter pressure to <10 mmHg by surgical myotomy compared with hydrostatic balloon dilation. However, patients whose sphincter pressure has been reduced by hydrostatic balloon dilation to <10 mmHg have an outcome similar to those after surgical myotomy (Fig. 25-63). Botulinum toxin injection may achieve similar results, but it has a longer duration of action that may be measured in weeks or months, rather than years. Botulinum toxin injection may best be used as a diagnostic tool, when it is not clear whether a hypertensive LES is the primary cause of dysphagia. Responsiveness to botulinum toxin injection may predict a good response to Heller myotomy.

The therapeutic decisions regarding the treatment of patients with achalasia center on four issues. The first issue is the question of whether newly diagnosed patients should be treated with pneumatic dilation or a surgical myotomy. Long-term follow-up studies have shown that pneumatic dilation...
achieves adequate relief of dysphagia and pharyngeal regurgitation in 50% to 60% of patients (Fig. 25-64). Close follow-up is required, and if dilatation fails, myotomy is indicated. For those patients who have a dilated and tortuous esophagus or an associated hiatal hernia, balloon dilation is dangerous and surgery is the better option. The outcome of the one controlled randomized study (38 patients) comparing the two modes of therapy suggests that surgical myotomy as a primary treatment gives better long-term results. Several randomized trials comparing laparoscopic cardiomyotomy with balloon dilation or botulinum toxin injection have favored the surgical approach as well.

Figure 25-63. Prevalence of clinical remission in 122 patients stratified according to postdilatation lower esophageal sphincter (LES) pressures greater than or <10 mmHg. (Reproduced with permission from Ponce J, Garrigues V, Pertejo V, et al: Individual prediction of response to pneumatic dilation in patients with achalasia, Dig Dis Sci. 1996 Nov;41(11):2135-2141.)


Although it has been reported that a myotomy after previous balloon dilation is more difficult, this has not been the experience of these authors unless the cardia has been ruptured in a sawtooth manner. In this situation, operative intervention, either immediately or after healing has occurred, can be difficult. Similarly, myotomy after botulinum toxin injection has reported to be more difficult, but this is largely a function of the submucosal inflammatory response, which may be a bit unpredictable, and is most intense in the first 6 to 12 weeks after injection. It is important to wait at least 3 months after botulinum toxin injection to perform cardiomyotomy to minimize the risk of encountering dense inflammation.

The second issue is the question of whether a surgical myotomy should be performed through the abdomen or the chest. Myotomy of the LES can be accomplished via either an abdominal or thoracic approach. In the absence of a previous upper abdominal surgery, most surgeons prefer the abdominal approach to LES myotomy as laparoscopy results in less pain and a shorter length of stay than thoracoscopy. In addition, it is a bit easier to ensure a long gastric myotomy when the approach is transabdominal.

The third issue—and one that has been long debated—is the question of whether an antireflux procedure should be added to a surgical myotomy. Excellent results have been reported following meticulously performed myotomy without an antireflux component. Retrospective studies, with long-term follow-up of large cohorts of patients undergoing Heller myotomy demonstrated that, after 10 years, more than 50% of patients had reflux symptoms without a fundoplication. In a recent randomized clinical trial, 7% of patients undergoing Dor fundoplication following LES myotomy had abnormal 24-hour pH probes, and 42% of patients with a myotomy only had abnormal reflux profiles. If an antireflux procedure is used as an adjunct to esophageal myotomy, a complete 360° fundoplication should be avoided. Rather, a 270° Belsey fundoplication, a Toupet posterior 180° fundoplication, or a Dor anterior 180° fundoplication should be used to avoid the long-term esophageal dysfunction secondary to the outflow obstruction afforded by the fundoplication itself.

The fourth issue centers on whether or not a cure of this disease is achievable. Long-term follow-up studies after surgical myotomy have shown that late deterioration in results occurs after this procedure, regardless of whether an antireflux procedure is done, and also after balloon dilation, even when the sphincter pressure is reduced to below 10 mmHg. It may be that, even though a myotomy or balloon rupture of the LES muscle reduces the outflow obstruction at the cardia, the underlying motor disorder in the body of the esophagus persists and deteriorates further with the passage of time, leading to increased impairment of esophageal emptying. The earlier an effective reduction in outflow resistance can be accomplished, the better the outcome will be, and the more likely some esophageal body function can be restored.

In performing a surgical myotomy of the LES, there are four important principles: (a) complete division of all circular and collar-sling muscle fibers, (b) adequate distal myotomy to reduce outflow resistance, (c) “undermining” of the muscularis to allow wide separation of the esophageal muscle, and (d) prevention of postoperative reflux. In the past, the drawback of a surgical myotomy was the need for an open procedure, which often deterred patients from choosing the best treatment option for achalasia. With the advent of minimally invasive surgical techniques two decades ago, laparoscopic cardiomyotomy...
(Heller myotomy) has become the treatment of choice for most patients with achalasia.

Open Esophageal Myotomy

Open techniques of distal esophageal myotomy are rarely used outside reoperations. In fact, primary procedures can almost always be successfully completed via laparoscopy. A modified Heller myotomy can be performed through a left thoracotomy incision in the sixth intercostal space along the upper border of the seventh rib. The esophagus and a tongue of gastric fundus are exposed as described for a long myotomy. A myotomy through all muscle layers is performed, extending distally over the stomach to 1 to 2 cm below the junction, and proximally on the esophagus for 4 to 5 cm. The cardia is reconstructed by suturing the tongue of gastric fundus to the margins of the myotomy to prevent healing of the myotomy site and to provide reflux protection in the area of the divided sphincter. If an extensive dissection of the cardia has been done, a more formal Belsey repair is performed. The tongue of gastric fundus is allowed to retract into the abdomen. Traditionally, nasogastric drainage is maintained for 6 days to prevent distention of the stomach during healing. An oral diet is resumed on the seventh day, after a barium swallow study shows unobstructed passage of the bolus into the stomach without extravasation.

In a randomized, long-term follow-up by Csendes and colleagues of 81 patients treated for achalasia, either by forceful dilation or by surgical myotomy, myotomy was associated with a significant increase in the diameter at the GEJ and a decrease in the diameter at the middle third of the esophagus on follow-up radiographic studies. There was a greater reduction in sphincter pressure and improvement in the amplitude of esophageal contractions after myotomy. After dilation, 13% of patients regained some peristalsis, compared with 28% after surgery. These findings were shown to persist over a 5-year follow-up period, at which time 95% of those treated with surgical myotomy were doing well. Of those who were treated with dilation, only 54% were doing well, while 16% required redilation, and 22% eventually required surgical myotomy to obtain relief.

If simultaneous esophageal contractions are associated with the sphincter abnormality, the so-called vigorous achalasia, then the myotomy should extend over the distance of the abnormal motility as mapped by the preoperative motility study. Failure to do this will result in continuing dysphagia and a dissatisfied patient. The best objective evaluation of improvement in the patient following either balloon dilation or myotomy is a scintigraphic measurement of esophageal emptying time. A good therapeutic response improves esophageal emptying toward normal. However, some degree of dysphagia may persist despite improved esophageal emptying, due to disturbances in esophageal body function. When an antireflux procedure is added to the myotomy, it should be a partial fundoplication. A 360° fundoplication is associated with progressive retention of swallowed food, regurgitation, and aspiration to a degree that exceeds the patient’s preoperative symptoms.

Laparoscopic Cardiomyotomy

More commonly known as a laparoscopic Heller myotomy, after Ernst Heller, a German surgeon who described a “double myotomy” in 1913, the laparoscopic approach is similar to the Nissen fundoplication in terms of the trocar placement and exposure and dissection of the esophageal hiatus (Fig. 25-65). The procedure begins by division of the short gastric vessels in preparation for fundoplication. Exposure of the GEJ via removal of the gastroesophageal fat pad follows. The anterior vagus nerve is swept right laterally along with the fat pad. Once completed, the GEJ and distal 4 to 5 cm of esophagus should be bared of any overlying tissue, and generally follows dissection of the GEJ. A distal esophageal myotomy is performed. It is generally easiest to begin myotomy 1 to 2 cm above the GEJ, in an area above that of previous botulinum toxin injections or balloon dilation. Either scissors or a hook-type electrocautery can be used to initiate the incision in the longitudinal and circular muscle. Distally, the myotomy is carried across the GEJ and onto the proximal stomach for approximately 2 to 3 cm. After completion, the muscle edges are separated bluntly from the esophageal mucosa for approximately 50% of the esophageal circumference. An antireflux procedure follows completion of the myotomy. Either an anterior hemifundoplication augmenting the angle of His (Dor) or posterior partial fundoplication (Toupet) can be performed. The Dor type fundoplication is slightly easier to perform, and it does not require disruption of the normal posterior gastroesophageal attachments (a theoretical advantage in preventing postoperative reflux).

Per Oral Endoscopic Myotomy (POEM)

The POEM procedure was developed in Japan. It is the ultimate minimally invasive myotomy as it requires no incisions through the skin. With the POEM procedure, a very effective myotomy is performed entirely from the lumen of the esophagus. The POEM procedure is started by opening the esophageal mucosa 10 cm above the lower esophageal sphincter with a needle–knife electrosurgery device passed through an endoscope. A long submucosal plane is developed with the endoscope, down to and below the LES. The circular muscle of the LES, above and below the gastroesophageal junction, is divided with endoscopic electrosurgery. The submucosal entry site in the esophagus is then closed with endoscopic clips. While the results of POEM are still accumulating, the procedure is attractive because it is extremely minimally invasive, and can be done on an outpatient basis. The major downside of POEM is that an effective antireflux valve cannot be created, exposing the patient to a 40% to 50% risk of GERD post procedure.

Outcome Assessment of the Therapy for Achalasia

Critical analysis of the results of therapy for motor disorders of the esophagus requires objective measurement. The use of symptoms alone as an endpoint to evaluate therapy for achalasia may be misleading. The propensity for patients to unconsciously modify their diet to avoid difficulty swallowing is underestimated, making an assessment of results based on symptoms unreliable. Insufficient reduction in outflow resistance may allow progressive esophageal dilation to develop slowly, giving the impression of improvement because the volume of food able to be ingested with comfort increases. A variety of objective measurements may be used to assess success, including LES pressure, esophageal baseline pressure, and scintigraphic assessment of esophageal emptying time. Esophageal baseline pressure is usually negative compared to gastric pressure. Given that the goal of therapy is to eliminate the outflow resistance of a nonrelaxing sphincter, measurements of improvements in esophageal baseline pressure and scintigraphic transit time may be better indicators of success, but these are rarely reported.
Eckardt and associates investigated whether the outcome of pneumatic dilation in patients with achalasia could be predicted on the basis of objective measurements. Postdilation LES pressure was the most valuable measurement for predicting long-term clinical response. A postdilation sphincter pressure <10 mmHg predicted a good response. Approximately 50% of the patients studied had postdilation sphincter pressures between 10 and 20 mmHg, with a 2-year remission rate of 71%. More important, 16 of 46 patients were left with a postdilation sphincter pressure of >20 mmHg and had an unacceptable outcome. Overall, only 30% of patients dilated remained in symptomatic remission at 5 years.

Bonavina and colleagues reported good to excellent results with transabdominal myotomy and Dor fundoplication in 94% of patients after a mean follow-up of 5.4 years. No operative mortality occurred in either of these series, attesting to the safety of the procedure. Malthaner and Pearson reported the long-term clinical results in 35 patients with achalasia, having a minimum follow-up of 10 years (Table 25-10). Twenty-two of these patients underwent primary esophageal myotomy and Belsey hemifundoplication at the Toronto General Hospital. Excellent to good results were noted in 95% of patients at 1 year, declining to 68%, 69%, and 67% at 10, 15, and 20 years, respectively. Two patients underwent early reoperation for an incomplete myotomy, and three underwent an esophagectomy for progressive disease. They concluded that there was a deterioration of the initially good results after surgical myotomy and hiatal repair for achalasia, which is due to late complications of gastroesophageal reflux.

Ellis reported his lifetime experience with transthoracic short esophageal myotomy without an antireflux procedure. One hundred seventy-nine patients were analyzed at a mean follow-up of 9 years, ranging from 6 months to 20 years. Overall, 89% of patients were improved at the 9-year mark. He also observed that the level of improvement deteriorated with time, with excellent results (patients continuing to be symptom free) decreasing from 54% at 10 years to 32% at 20 years. He concluded that a short transthoracic myotomy without an antireflux procedure provides excellent long-term relief of dysphagia, and, contrary to Malthaner and Pearson’s experience, does not result in complications of gastroesophageal reflux. Both studies document nearly identical results 10 to 15 years following the procedure, and both report deterioration over time, probably due to progression of the underlying disease. The addition of an antireflux procedure if the operation is performed transthoracically has no significant effect on the outcome.

**Figure 25-65.** A. Longitudinal muscle is divided. B. Mechanical disruption of lower esophageal sphincter muscle fibers. C. Myotomy must be carried across gastroesophageal junction. D. Gastric extension should equal 2 to 3 cm. E. Anterior (Dor) fundoplication is sutured to the diaphragmatic arch. F. Posterior (Toupet) fundoplication is sutured to cut edges of myotomy. EG jct = esophagogastric junction.
Table 25-10
Reasons for failure of esophageal myotomy

<table>
<thead>
<tr>
<th>REASON</th>
<th>ELLIS, MYOTOMY ONLY (N = 81)</th>
<th>GOULBOURNE, MYOTOMY ONLY (N = 65)</th>
<th>MALTHANER, MYOTOMY + ANTIREFLUX (N = 22)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflux</td>
<td>4%</td>
<td>5%</td>
<td>18%</td>
</tr>
<tr>
<td>Inadequate myotomy</td>
<td>2%</td>
<td>—</td>
<td>9%</td>
</tr>
<tr>
<td>Megaesophagus</td>
<td>2%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Poor emptying</td>
<td>4%</td>
<td>3%</td>
<td>—</td>
</tr>
<tr>
<td>Persistent chest pain</td>
<td>1%</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

The outcome of laparoscopic myotomy and hemifundoplication has been well documented. Two reports of over 100 patients have documented relief of dysphagia in 93% of patients. Richter and coworkers reviewed published reports to date, including 254 patients with an average success rate of 93% at 2.5 years. Conversion to an open procedure occurs in 0% to 5% of patients. Complications are uncommon, occurring in <5% of patients. Intraoperative complications consist largely of mucosal perforation, and have been more likely to occur after botulinum toxin injection. The incidence of objective reflux disease as evidenced by abnormal acid output is <10%.

A number of randomized clinical trials in the past decade have compared the outcomes of laparoscopic Heller myotomy to pneumatic dilation and to botulinum toxin injection. In each of these trials, laparoscopic Heller myotomy and partial fundoplication was superior to the alternative treatment. Lastly, a randomized clinical trial examining the need for fundoplication following Heller myotomy demonstrated a great deal more reflux in patients without fundoplication, and no better swallowing in the Heller-only group. The best treatment for achalasia is a laparoscopic Heller myotomy and partial fundoplication. The role of POEM in the management of classic (nonspastic) achalasia is yet to be established.

**Esophageal Resection for End-Stage Motor Disorders of the Esophagus**

Patients with dysphagia and long-standing benign disease, whose esophageal function has been destroyed by the disease process or multiple previous surgical procedures, are best managed by esophagectomy. Fibrosis of the esophagus and cardia can result in weak contractions and failure of the distal esophageal sphincter to relax. The loss of esophageal contractions can result in the stasis of food, esophageal dilatation, regurgitation, and aspiration. The presence of these abnormalities signals end-stage motor disease. In these situations, esophageal replacement is usually required to establish normal alimentation. Before proceeding with esophageal resection for patients with end-stage benign disease, the choice of the organ to substitute for the esophagus (i.e., stomach, jejunum, or colon) should be considered. The choice of replacement is affected by a number of factors, as described later in “Techniques of Esophageal Reconstruction.” If minimally invasive esophagectomy is to be performed, thorascoposcopic dissection should be combined with abdominal dissection. Attempts at MIS transhiatal esophagectomy for the massively dilated esophagus may result in large volume bleeding from mediastinal vessels that become enlarged with esophageal dilation, and such bleeding must be directly controlled for hemostasis to be adequate and the operation to be safe.

**CARCINOMA OF THE ESOPHAGUS**

Squamous carcinoma accounts for the majority of esophageal carcinomas worldwide. Its incidence is highly variable, ranging from approximately 20 per 100,000 in the United States and Britain, to 160 per 100,000 in certain parts of South Africa and the Henan Province of China, and even 540 per 100,000 in the Guriev district of Kazakhstan. The environmental factors responsible for these localized high-incidence areas have not been conclusively identified, though additives to local foodstuffs (nitroso compounds in pickled vegetables and smoked meats) and mineral deficiencies (zinc and molybdenum) have been suggested. In Western societies, smoking and alcohol consumption are strongly linked with squamous carcinoma. Other definite associations link squamous carcinoma with long-standing achalasia, lye strictures, tylosis (an autosomal dominant disorder characterized by hyperkeratosis of the palms and soles), and human papillomavirus.

Adenocarcinoma of the esophagus, once an unusual malignancy, is diagnosed with increasing frequency (Fig. 25-66) and now accounts for more than 50% of esophageal cancer in most Western countries. The shift in the epidemiology of esophageal cancer from predominantly squamous carcinoma seen in association with smoking and alcohol to adenocarcinoma in the setting of BE is one of the most dramatic changes that has occurred in the history of human neoplasia. Although esophageal carcinoma is a relatively uncommon malignancy, its prevalence is exploding, largely secondary to the well-established association among gastroesophageal reflux, BE, and esophageal adenocarcinoma. Although BE was once a nearly uniformly lethal disease, survival has improved slightly because of advances in the understanding of its molecular biology, screening and surveillance practices, improved staging, minimally invasive surgical techniques, and neoadjuvant therapy.

Furthermore, the clinical picture of esophageal adenocarcinoma is changing. It now occurs not only considerably more frequently but also in younger patients, and it is often detected at an earlier stage. These facts support rethinking the traditional approach of assuming palliation is appropriate in all patients. The historical focus on palliation of dysphagia in an elderly patient with comorbidities should change when dealing with a young patient with dependent children and a productive life ahead. The potential for cure becomes of paramount importance.

The gross appearance resembles that of squamous cell carcinoma. Microscopically, adenocarcinoma almost always originates in Barrett’s mucosa and resembles gastric cancer. Rarely, it arises in the submucosal glands and forms intramural growths that resemble the mucoid or adenoid cystic carcinomas of the salivary glands.

The most important etiologic factor in the development of primary adenocarcinoma of the esophagus is a metaplastic columnar-lined or Barrett’s esophagus, which occurs in approximately 10% to 15% of patients with GERD. When studied prospectively, the incidence of adenocarcinoma in a patient with BE is one in 100 to 200 patient-years of follow-up (i.e., for every 100 patients with BE followed for 1 year, one will develop adenocarcinoma). Although this risk appears to be small, it is at least 40 to 60 times that expected for a similar population without BE. This risk is similar to the risk for developing lung cancer in a person with a 20-pack-per-year history of smoking. Endoscopic surveillance for patients with BE is recommended for two reasons: (a) at present there is no reliable evidence that medical therapy removes the risk of neoplastic transformation, and (b) malignancy in BE is curable if detected at an early stage.

**Clinical Manifestations**

Esophageal cancer generally presents with dysphagia, although increasing numbers of relatively asymptomatic patients are now identified on surveillance endoscopy, or present with nonspecific upper GI symptoms and undergo screening endoscopy. Extension of the primary tumor into the tracheobronchial tree can occur primarily with squamous cell carcinoma and can cause stridor, tracheoesophageal fistula, and resultant coughing, choking, and aspiration.
Dysphagia usually presents late in the natural history of the disease because the lack of a serosal layer on the esophagus allows the smooth muscle to dilate with ease. As a result, the dysphagia becomes severe enough for the patient to seek medical advice only when more than 60% of the esophageal circumference is infiltrated with cancer. Consequently, the disease is usually advanced if symptoms herald its presence. Tracheoesophageal fistula may be present in some patients on their first visit to the hospital, and more than 40% will have evidence of distant metastases. With tumors of the cardia, anorexia and weight loss usually precede the onset of dysphagia. The physical signs of esophageal tumors are those associated with the presence of distant metastases.

**General Approach to Esophageal Cancer**

Therapy of esophageal cancer is dictated by the stage of the cancer at the time of diagnosis. Put simply, one needs to determine if the disease is confined to the esophagus, (T1–T2, N0), locally advanced (T1–3, N1), or disseminated (any T, any N, M1). If cancer is confined to the esophagus, removal of the tumor with adjacent lymph nodes may be curative. Very early tumors confined to the mucosa (T in situ, T1a, intramucosal cancer) may be addressed with endoscopic treatment. When the tumor is locally aggressive, modern therapy dictates a multimodality approach in a surgically fit patient. Multimodality therapy is either chemotherapy followed by surgery or radiation and chemotherapy followed by surgery. When given before surgery, these treatments are referred to as neoadjuvant or induction therapy. For disseminated cancer, treatment is aimed at palliation of symptoms. If the patient has dysphagia, as many do, the most rapid form of palliation is the endoscopic placement of an expandable esophageal stent. For palliation of GEJ cancer, radiation may be the first choice, as stents placed across the GEJ create a great deal of gastroesophageal reflux.

**Staging of Esophageal Cancer**

Choosing the best therapy for an individual patient requires accurate staging. Staging starts with the history and physical. LN disease remote from the tumor, particularly in the cervical region, may be palpable on neck examination and generally indicates cancer dissemination. This is often referred to as M1a disease, indicating that these patients should not be treated with therapy directed toward locally advanced cancer. Other metastatic LNs are rarely palpable but are equally ominous, especially the umbilical LN in GEJ cancer.

Computed tomographic (CT) scanning of the chest, abdomen, and pelvis provides information on local invasion of the primary cancer, LN involvement, or disseminated disease. The most common sites of esophageal cancer metastases are lung, liver, and peritoneal surfaces, including the omentum and small bowel mesentery. If masses are identified that are
not characteristic for cancer or are in a location that precludes resection with the cancer specimen, positron emission tomography (PET) scanning may be able to tell whether the masses are metabolically active (likely to be cancer) or not. A PET active focus corresponding to a mass on CT scan outside of the field of esophageal resection should be biopsied before resection is performed.

The introduction of endoscopic ultrasound (EUS) has made it possible to identify patients who are potentially curable before surgical therapy. Using an endoscope, the depth of the wall penetration by the tumor and the presence of LN metastases can be determined with 80% accuracy. A curative resection should be encouraged if EUS indicates that the tumor has not invaded adjacent organs (T4b), and/or fewer than six enlarged LNs are imaged. Thoracoscopic and laparoscopic staging of esophageal cancer may add benefit when the nature of enlarged LNs remote from the cancer cannot be determined or when advanced imaging systems (PET and high-resolution spiral CT) are not available.

Occasionally, diagnostic laparotomy and jejunostomy tube placement may precede induction chemoradiation in the patient with severe dysphagia and weight loss from a locally advanced cancer. In summary, esophageal cancer is diagnosed with endoscopic biopsy and is staged with CT scanning of the chest and abdomen, EUS, and PET scan for all patients with CT or EUS evidence of advanced disease (T2 or greater, N1-2 or NX). Experience with esophageal resection in patients with early stage disease has identified characteristics of esophageal cancer that are associated with improved survival. A number of studies suggest that only metastasis to LNs and tumor penetration of the esophageal wall have a significant and independent influence on prognosis. Factors known to be important in the survival of patients with advanced disease, such as cell type, degree of cellular differentiation, or location of tumor in the esophagus, have no effect on survival of patients who have undergone resection for early disease. Studies also showed that patients having five or fewer LN metastases have a better outcome. Using these data, Skinner developed the wall penetration, LN, and distant organ metastases system for staging.

The wall penetration, LN, and distant organ metastases system differed somewhat from the previous efforts to develop a satisfactory staging criteria for carcinoma of the esophagus. Most surgeons agreed that the 1983 tumor, nodes, and metastasis system left much to be desired. In the third edition of the manual for Staging of Cancer of the American Joint Committee on Cancer (AJCC) in 1988, an effort was made to provide a finer discrimination between stages than had been contained in the previous edition in 1983. In 2016, further refinements of the staging system of esophageal cancer were approved by the AJCC, recognizing the difference in survival afforded by resection of limited LN disease adjacent to the tumor, compared to multilevel LN disease and positive LNs remote from the primary. Table 25-11 shows the AJCC definitions for the primary tumor, lymph nodes, distant metastasis, and overall staging schema for both squamous cell carcinoma and adenocarcinoma.

Clinical Approach to Carcinoma of the Esophagus and Cardia

The selection of a curative vs. a palliative operation for cancer of the esophagus is based on the location of the tumor, the patient’s age and health, the extent of the disease, and preoperative staging. Figure 25-67 shows an algorithm of the clinical decisions important in the selection of curative or palliative therapy.

Tumor Location. The selection of surgical therapy for patients with carcinoma of the esophagus depends not only on the anatomic stage of the disease and an assessment of the swallowing capacity of the patient but also on the location of the primary tumor.

It is estimated that 8% of the primary malignant tumors of the esophagus occur in the cervical portion (Fig. 25-68). They are almost always squamous cell cancer, with a rare adenocarcinoma arising from a congenital inlet patch of columnar lining. These tumors, particularly those in the postcricoid area, represent a separate pathologic entity for two reasons: (a) they are more common in females and appear to be a unique entity in this regard; and (b) the efferent lymphatics from the cervical esophagus drain completely differently from those of the thoracic esophagus. The latter drain directly into the paratracheal and deep cervical or internal jugular LNs with minimal flow in a longitudinal direction. Except in advanced disease, it is unusual for intrathoracic LNs to be involved.

Cervical esophageal cancer is frequently unresectable because of early invasion of the larynx, great vessels, or trachea. Radical surgery, including esophagealryngectomy may occasionally be performed for these lesions, but the ensuing morbidity makes this a less than desirable approach in the face of uncertain cure. Thus, for most patients with cervical esophageal cancer, stereotactic radiation with concomitant chemotherapy is the most desirable treatment.

Tumors that arise within the middle third of the esophagus are squamous carcinomas most commonly and are frequently associated with LN metastasis, which are usually in the thorax but may be in the neck or abdomen, and may skip areas in between. Although it is generally felt that individuals with midthoracic cancer and abdominal LN metastases are incurable with surgery, there are some emerging data that suggest that cervical LN metastases, if isolated, can be resected with benefit. Generally, T1 and T2 cancers without LN metastases are treated with resection only, but there is more and more data to suggest that TN involvement or transmural cancer (T3) warrants treatment with neoadjuvant chemoradiation therapy followed by resection. Although some surgeons prefer a transhiatal esophagectomy for all tumor locations, most surgeons believe that resection of mid-esophageal cancer should be performed under direct vision with either thoracoscopy (video-assisted thoracic surgery [VATS]) or with thoracotomy.

Tumors of the lower esophagus and cardia are usually adenocarcinomas. Unless preoperative and intraoperative staging clearly demonstrate an incurable lesion, resection in continuity with a LN dissection should be performed. Because of the propensity of GI tumors to spread for long distances submucosally, long lengths of grossly normal GI tract should be resected. The longitudinal lymph flow in the esophagus can result in skip areas, with small foci of tumor above the primary lesion, which underscores the importance of a wide resection of esophageal tumors. Wong has shown that local recurrence at the anastomosis can be prevented by obtaining a 10-cm margin of normal esophagus above the tumor. Anatomic studies have also shown that there is no submucosal lymphatic barrier between the esophagus and the stomach at the cardia, and Wong has
### Table 25-11

**American Joint Committee on Cancer (AJCC) Staging Schema for Esophageal Cancer**

<table>
<thead>
<tr>
<th><strong>TX</strong></th>
<th>Primary tumor cannot be assessed.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T0</strong></td>
<td>No evidence of primary tumor.</td>
</tr>
<tr>
<td><strong>Tis</strong></td>
<td>High-grade dysplasia.</td>
</tr>
<tr>
<td><strong>T1</strong></td>
<td>Tumor invades lamina propria, muscularis mucosae, or submucosa.</td>
</tr>
<tr>
<td><strong>T1a</strong></td>
<td>Tumor invades lamina propria or muscularis mucosae.</td>
</tr>
<tr>
<td><strong>T1b</strong></td>
<td>Tumor invades submucosa.</td>
</tr>
<tr>
<td><strong>T2</strong></td>
<td>Tumor invades muscularis propria.</td>
</tr>
<tr>
<td><strong>T3</strong></td>
<td>Tumor invades adventitia.</td>
</tr>
<tr>
<td><strong>T4</strong></td>
<td>Tumor invades adjacent structures.</td>
</tr>
<tr>
<td><strong>T4a</strong></td>
<td>Resectable tumor invading pleura, pericardium, or diaphragm.</td>
</tr>
<tr>
<td><strong>T4b</strong></td>
<td>Unresectable tumor invading other adjacent structures, such as aorta, vertebral body, trachea, etc.</td>
</tr>
<tr>
<td><strong>NX</strong></td>
<td>Regional lymph nodes cannot be assessed.</td>
</tr>
<tr>
<td><strong>N0</strong></td>
<td>No regional lymph node metastasis.</td>
</tr>
<tr>
<td><strong>N1</strong></td>
<td>Metastases in 1–2 regional lymph nodes.</td>
</tr>
<tr>
<td><strong>N2</strong></td>
<td>Metastases in 3–6 regional lymph nodes.</td>
</tr>
<tr>
<td><strong>N3</strong></td>
<td>Metastases in ≥7 regional lymph nodes.</td>
</tr>
<tr>
<td><strong>M0</strong></td>
<td>No distant metastasis.</td>
</tr>
<tr>
<td><strong>M1</strong></td>
<td>Distant metastasis.</td>
</tr>
</tbody>
</table>

**SQUAMOUS CELL CARCINOMA**

**Clinical (cTNM)**

<table>
<thead>
<tr>
<th>When cT is...</th>
<th>And cN is...</th>
<th>And M is...</th>
<th>Then the stage group is...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tis</td>
<td>N0</td>
<td>M0</td>
<td>0</td>
</tr>
<tr>
<td>T1</td>
<td>N0–1</td>
<td>M0</td>
<td>I</td>
</tr>
<tr>
<td>T2</td>
<td>N0–1</td>
<td>M0</td>
<td>II</td>
</tr>
<tr>
<td>T3</td>
<td>N0</td>
<td>M0</td>
<td>III</td>
</tr>
<tr>
<td>T1–3</td>
<td>N2</td>
<td>M0</td>
<td>III</td>
</tr>
<tr>
<td>T4</td>
<td>N0–2</td>
<td>M0</td>
<td>IVA</td>
</tr>
<tr>
<td>Any T</td>
<td>N3</td>
<td>M0</td>
<td>IVA</td>
</tr>
<tr>
<td>Any T</td>
<td>Any N</td>
<td>M1</td>
<td>IVB</td>
</tr>
</tbody>
</table>

**Pathological (pTNM)**

<table>
<thead>
<tr>
<th>When pT is...</th>
<th>And pN is...</th>
<th>And M is...</th>
<th>And G is...</th>
<th>And location is...</th>
<th>Then the stage group is...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tis</td>
<td>N0</td>
<td>M0</td>
<td>N/A</td>
<td>Any</td>
<td>0</td>
</tr>
<tr>
<td>T1a</td>
<td>N0</td>
<td>M0</td>
<td>G1</td>
<td>Any</td>
<td>IA</td>
</tr>
<tr>
<td>T1a</td>
<td>N0</td>
<td>M0</td>
<td>G2–3</td>
<td>Any</td>
<td>IB</td>
</tr>
<tr>
<td>T1a</td>
<td>N0</td>
<td>M0</td>
<td>GX</td>
<td>Any</td>
<td>IA</td>
</tr>
<tr>
<td>T1b</td>
<td>N0</td>
<td>M0</td>
<td>G1–3</td>
<td>Any</td>
<td>IB</td>
</tr>
<tr>
<td>T1b</td>
<td>N0</td>
<td>M0</td>
<td>GX</td>
<td>Any</td>
<td>IB</td>
</tr>
<tr>
<td>T2</td>
<td>N0</td>
<td>M0</td>
<td>G1</td>
<td>Any</td>
<td>IA</td>
</tr>
<tr>
<td>T3</td>
<td>N0</td>
<td>M0</td>
<td>G1–3</td>
<td>Lower</td>
<td>IA</td>
</tr>
<tr>
<td>T3</td>
<td>N0</td>
<td>M0</td>
<td>G1</td>
<td>Upper/middle</td>
<td>IIA</td>
</tr>
<tr>
<td>T3</td>
<td>N0</td>
<td>M0</td>
<td>G2–3</td>
<td>Upper/middle</td>
<td>IIB</td>
</tr>
<tr>
<td>Any T</td>
<td>N3</td>
<td>M0</td>
<td>Location X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any T</td>
<td>Any N</td>
<td>M1</td>
<td>Group X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Postneoadjuvant Therapy (ypTNM)**

<table>
<thead>
<tr>
<th>When yp T is...</th>
<th>And yp N is...</th>
<th>And M is...</th>
<th>Then the stage group is...</th>
</tr>
</thead>
<tbody>
<tr>
<td>T0–2</td>
<td>N0</td>
<td>M0</td>
<td>I</td>
</tr>
<tr>
<td>T1</td>
<td>N0</td>
<td>M0</td>
<td>II</td>
</tr>
<tr>
<td>T0–2</td>
<td>N0</td>
<td>M0</td>
<td>III</td>
</tr>
<tr>
<td>T3</td>
<td>N0</td>
<td>M0</td>
<td>III</td>
</tr>
<tr>
<td>T0–3</td>
<td>N2</td>
<td>M0</td>
<td>IIB</td>
</tr>
<tr>
<td>T4a</td>
<td>N0–1</td>
<td>M0</td>
<td>IVA</td>
</tr>
<tr>
<td>T4a</td>
<td>N0–2</td>
<td>M0</td>
<td>IVA</td>
</tr>
<tr>
<td>Any T</td>
<td>N3</td>
<td>M0</td>
<td>IVA</td>
</tr>
<tr>
<td>Any T</td>
<td>Any N</td>
<td>M1</td>
<td>IVB</td>
</tr>
</tbody>
</table>

**ADENOCARCINOMA**

**Clinical (cTNM)**

<table>
<thead>
<tr>
<th>When cT is...</th>
<th>And cN is...</th>
<th>And M is...</th>
<th>Then the stage group is...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tis</td>
<td>N0</td>
<td>M0</td>
<td>0</td>
</tr>
<tr>
<td>T1</td>
<td>N0</td>
<td>M0</td>
<td>I</td>
</tr>
<tr>
<td>T1</td>
<td>N0</td>
<td>M0</td>
<td>IIA</td>
</tr>
<tr>
<td>T2</td>
<td>N0</td>
<td>M0</td>
<td>IIB</td>
</tr>
<tr>
<td>T2</td>
<td>N1</td>
<td>M0</td>
<td>III</td>
</tr>
<tr>
<td>T3</td>
<td>N1–2</td>
<td>M0</td>
<td>III</td>
</tr>
<tr>
<td>T4a</td>
<td>N0–1</td>
<td>M0</td>
<td>III</td>
</tr>
<tr>
<td>T4b</td>
<td>N0–2</td>
<td>M0</td>
<td>III</td>
</tr>
<tr>
<td>Any T</td>
<td>N3</td>
<td>M0</td>
<td>IVA</td>
</tr>
<tr>
<td>Any T</td>
<td>Any N</td>
<td>M1</td>
<td>IVB</td>
</tr>
</tbody>
</table>

(Continued)
shown that 50% of the local recurrences in patients with esophageal cancer who are resected for cure occur in the intrathoracic stomach along the line of the gastric resection. Considering that the length of the esophagus ranges from 17 to 25 cm, and the length of the lesser curvature of the stomach is approximately 12 cm, a curative resection requires a cervical division of the esophagus and a >50% proximal gastrectomy in most patients with carcinoma of the distal esophagus or cardia.

**Age.** Resection for cure of carcinoma of the esophagus in a patient older than 80 years is rarely indicated because of the additional operative risk and the shorter life expectancy. Despite this general guideline, octogenarians with a high-performance status and excellent cardiopulmonary reserve may be considered candidates for esophagectomy, and recent case series have established its success in highly selected patients. It is in this group of patients that the lesser physiologic impact of minimally

---

**Figure 25-67.** Algorithm for the evaluation of esophageal cancer patients to select the proper therapy: curative en bloc resection, palliative transhiatal resection, or nonsurgical palliation. CT = computed tomography; FEV₁ = forced expiratory volume in 1 second. (Reproduced with permission from DeMeester TR: Esophageal carcinoma: current controversies, Semin Surg Oncol. 1997 Jul-Aug;13(4):217-233.)
invasive surgery may reduce the morbidity and mortality associated with open two- or three-field esophagectomy.

**Cardiopulmonary Reserve.** Patients undergoing esophageal resection should have sufficient cardiopulmonary reserve to tolerate the proposed procedure. The respiratory function is best assessed with the forced expiratory volume in 1 second, which ideally should be 2 L or more. Any patient with a forced expiratory volume in 1 second of <1.25 L is a poor candidate for thoracotomy because he or she has a 40% risk of dying from respiratory insufficiency within 4 years. In patients with poor pulmonary reserve, the transhiatal esophagectomy should be considered, as the pulmonary morbidity of this operation is less than is seen following thoracotomy. Clinical evaluation and electrocardiogram are not sufficient indicators of cardiac reserve. Echocardiography and dipyridamole thallium imaging provide accurate information on wall motion, ejection fraction, and myocardial blood flow. A defect on thallium imaging may require further evaluation with preoperative coronary angiography. A resting ejection fraction of <40%, particularly if there is no increase with exercise, is an ominous sign. In the absence of invasive testing, observed stair-climbing is an economical (albeit not quantitative) method of assessing cardiopulmonary reserve. Most individuals who can climb three flights of stairs without stopping will do well with two-field open esophagectomy, especially if an epidural catheter is used for postoperative pain relief.

**Nutritional Status.** The factor most predictive of postoperative complication is the nutritional status of the patient. Profound weight loss, more than 20 lb, associated with hypalbuminemia (albumin <3.5 g/dL) is associated with a much higher rate of complications and mortality than patients who enter curative surgery in better nutritional condition. Because malnourished patients generally have locally advanced esophageal cancer, if not metastatic disease, one should consider the placement of a feeding tube before the beginning of induction chemoradiation therapy. Although mild amounts of dysphagia are improved by induction chemoradiation therapy, more pronounced dysphagia and associated malnutrition should be addressed before the initiation of chemoradiation. A laparoscopic jejunostomy tube can be placed prior to induction therapy or at the time of esophagectomy. There are emerging data that 5 days’ pretreatment with immune-enhancing nutrition, rich in fish oils, decreases cardiac and other complications, following esophagectomy.

**Clinical Staging.** Clinical factors that indicate an advanced stage of carcinoma and exclude surgery with curative intent are recurrent nerve paralysis, Horner’s syndrome, persistent spinal pain, paralysis of the diaphragm, fistula formation, and malignant pleural effusion. Factors that make surgical cure unlikely include a tumor >8 cm in length, abnormal axis of the esophagus on a barium radiogram, more than four enlarged LNs on CT, a weight loss more than 20%, and loss of appetite. Studies indicate that there are several favorable parameters associated with tumors <4 cm in length, there are fewer with tumors between 4 and 8 cm, and there are no favorable criteria for tumors >8 cm in length. Consequently, the finding of a tumor >8 cm in length should exclude curative resection; the finding of a smaller tumor should encourage an aggressive approach.

**Preoperative Staging With Advanced Imaging.** For years, clinical staging, contrast radiography, endoscopy, and CT scanning formed the backbone of esophageal cancer staging. More recently, preoperative decision making is guided by endoscopic ultrasonography and PET scanning.

EUS provides the most reliable method of determining depth of cancer invasion. In the absence of enlarged LNs, the degree of wall invasion dictates surgical therapy. If a small focus of esophageal cancer is confined to the mucosa, endoscopic mucosal resection (EMR) is a preferable option. If the tumor invades into the submucosa, without visible lymph node involvement, most individuals would suggest esophagectomy with LN dissection, as positive nodes can be found in 20% to 25% of those with cancer limited to the mucosa and submucosa. If EUS demonstrates spread through the wall of the esophagus, especially if LNs are enlarged, then induction chemoradiation therapy (neoadjuvant therapy) should be strongly considered. Lastly, when the EUS demonstrates invasion of the trachea, bronchus, aorta, or spine, then surgical resection is rarely indicated. If there is invasion into the pleura (T4a), then surgical resection can be considered in the absence of a malignant effusion. Thus, it can be seen that the therapy of esophageal cancer is largely driven by the findings of an endoscopic ultrasonography. It is difficult to provide modern treatment of esophageal cancer without access to this modality.

PET scanning, usually combined with an axial CT scan (CTPET), usually is performed on patients with locally advanced cancer or questionable lesions on CT scan to determine whether metastases are present. The PET scan uses the injection of radiolabeled deoxyglucose, which is taken up in metabolically active tissues such as cancer. PET-positive areas must be correlated with the CT scan findings to assess the significance of “hot spots.” CTPET scanning has been especially useful before the initiation of chemoradiation therapy. An early response to chemoradiation therapy, by PET scan, improves the prognosis whether or not resection is ultimately performed. Conversely, if a PET-avid tumor shows no change in metabolic activity after 2 weeks of induction chemoradiation therapy, it is unlikely that further chemo- or radiation therapy will be of

---

**Figure 25-68.** Incidence of carcinoma of the esophagus and cardia based on tumor location.
any benefit. These patients have a worse prognosis and may be referred for resection or palliation without incurring the morbidity or expense of a full course of chemo- and radiation therapy.

**Palliation of Esophageal Cancer**

Palliation of esophageal cancer is indicated for individuals with metastatic esophageal cancer or cancer invading adjacent organs (T4b) who are unable to swallow, or individuals with fistulae into the tracheobronchial tree. Aortic esophageal fistulas are extremely rare and nearly 100% lethal. Dysphagia as a result of esophageal cancer can be graded from grade I, eating normally, to grade VI, unable to swallow saliva (Table 25-12). Grades I to III often can be managed with radiation therapy, usually in combination with chemotherapy. When surgical resection is not anticipated in the future, this is termed *definitive chemoradiation therapy* and usually is palliative. Radiation dose is increased from 45 Gy to 60 Gy administered over 8 weeks, rather than the 4 weeks given for chemoradiation induction therapy. In 20% of patients, a complete response to chemoradiation therapy will not only palliate the symptoms but will also leave the patient with undetectable cancer of the esophagus. Although some of these patients are truly cured, cancer will recur in many either locally or systemically 1 to 5 years following definitive chemoradiation. In a few patients, definitive chemoradiation will be successful in all sites but the esophagus. After a 12-month wait from initial treatment and no other sites of tumor detectable except the esophagus, some of these patients may be candidates for salvage esophagectomy.

For individuals with dysphagia grades IV and higher, additional treatment generally is necessary. The mainstay of therapy is in-dwelling esophageal stents. Covered removable stents may be used to seal fistulae or when stent removal becomes desirable in the future. When large, locally invasive tumors or metastatic esophageal cancer precludes any future hope of resection, uncovered expandable metal stents are the treatment of choice. The major limitations to stenting exist in cancers at the GEJ. A stent placed across the GEJ will result in severe gastroesophageal reflux and heartburn that can be quite disabling. In cancers at this level, radiation therapy alone may be preferable. If feeding access is desirable, a laparoscopic jejunostomy is usually the procedure of choice.

### Table 25-12

**Functional grades of dysphagia**

<table>
<thead>
<tr>
<th>GRADE</th>
<th>DEFINITION</th>
<th>INCIDENCE AT DIAGNOSIS (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Eating normally</td>
<td>11</td>
</tr>
<tr>
<td>II</td>
<td>Requires liquids with meals</td>
<td>21</td>
</tr>
<tr>
<td>III</td>
<td>Able to take semisolids but unable to take any solid food</td>
<td>30</td>
</tr>
<tr>
<td>IV</td>
<td>Able to take liquids only</td>
<td>40</td>
</tr>
<tr>
<td>V</td>
<td>Unable to take liquids, but able to swallow saliva</td>
<td>7</td>
</tr>
<tr>
<td>VI</td>
<td>Unable to swallow saliva</td>
<td>12</td>
</tr>
</tbody>
</table>


### Surgical Treatment

The surgical treatment of esophageal cancer is dependent upon the location of the cancer, the depth of invasion, LN metastases, the fitness of the patient for operation, and the culture and beliefs of the individuals and institutions in which the treatment is performed. In an ideal world, there would be a single, stage-specific method of treating esophageal cancer because the evidence would be unassailable and noncontroversial. Randomized clinical trials and meta-analyses would prove beyond a shadow of a doubt the value of surgery vs. nonoperative therapy and would dictate the type and extent of surgery that would optimally balance immediate morbidity and mortality with duration and quality of life conferred by the procedure and the perioperative management of the esophagectomy patient. Despite many noble attempts to establish this high level of evidence, many questions relating to the appropriate therapy of esophageal cancer remain. About the only area of complete agreement is that esophagectomy should not be performed if an R0 resection is not possible. In other words, if the surgeon does not believe he or she can remove all LNs invaded by cancer and provide a tumor-free radial margin and esophagus and stomach margins that are tumor free, then a resection should not be performed.

**Mucosally Based Cancer.** In patients with BE, and especially those with high-grade dysplasia, subcentimeter nodules are frequently discovered. Nodules should be resected in entirety, as they often harbor adenocarcinoma. Five years ago, such resection was performed with a transhiatal esophagectomy, but more recently EMR offers another method for removing intramucosal cancer. In this clinical situation, EMR is typically combined with EUS to rule out more invasive disease. EUS, however, is unable to differentiate between cancer that is confined to the mucosa (T1a) and that which invades the submucosa (T1b). Tumors invading the submucosa are not amenable to endoscopic mucosal resection because of the high-frequency (20–25%) concurrent finding of positive LNs, which cannot be removed without esophagectomy. On the other hand, intramucosal cancers have little risk of spreading to regional LNs. The current approach used involves performing EMR on all nodules identified in a field of Barrett’s esophagus, and then T staging is performed by histologic analysis. This approach dictates the need for future therapy such as esophagectomy.

For this reason, small intramucosal carcinomas may be removed with EMR in the following manner. The area beneath the nodule is infiltrated with saline through a sclerotherapy needle. A specialized suction cap is mounted on the end of the endoscope, and the nodule is drawn up into the cap; a snare is then applied to resect the tissue. Alternatively, a rubber band can be delivered, and the snare can be used to resect above the level of the rubber band. This specimen is then removed and sent to pathology. As long as the tumor is found to be confined to the mucosa and all margins are negative, the resection is complete. A positive margin or involvement of the submucosa warrants esophagectomy. Most importantly, these patients are at high risk for developing small nodular carcinomas elsewhere in their Barrett’s segment, and routine surveillance on a 3- to 6-month basis must be continued indefinitely. Alternatively, one can consider radiofrequency ablation of the remainder of the high-grade dysplasia after careful surveillance biopsy specimens demonstrate no further sign of cancer. This approach to the early esophageal cancer
should not be used when there is any suspicion of mediastinal or abdominal lymphadenopathy. Although it is currently rare that EMR provides definitive therapy of small nodular esophageal cancers, this may become more of the norm as greater surveillance reveals earlier cancers and proficiency of the technique by surgeons and gastroenterologists increases.

**Minimally Invasive Transhiatal Esophagectomy.**

Minimally invasive transhiatal esophagectomy is an increasingly popular procedure; however, the number of these operations performed around the world remains small. Mini-invasive surgery (MIS) transhiatal esophagectomy was first performed by Aureo DePaula in Brazil and has been modified and adopted by many individuals around the world. This operation combines the advantages of transhiatal esophagectomy at minimizing pulmonary complications with the advantages of laparoscopy (less pain, quicker rehabilitation). Several variations of MIS transhiatal esophagectomy have been developed. For the earliest lesions, such as high-grade dysplasia or intramucosal carcinoma, a vagal sparing procedure can be entertained. In such a procedure, the vagal trunks are separated from the esophagus at the level of the diaphragm and the lesser curvature dissection of the stomach allows the vagus and left gastric pedicle to remain intact. Clearly, this dissection, which hugs the stomach and esophagus, provides no LN staging and is thus inadequate for all high-grade dysplasia and intramucosal cancer.

MIS transhiatal esophagectomy is usually performed through five or six small incisions in the upper abdomen and a transverse cervical incision for removing the specimen and performing the cervical esophagogastrotomy. To remove the esophagus from the posterior mediastinum, especially the area behind the pulmonary vessels and the tracheal bifurcation, which cannot be visualized even with a long laparoscope placed in the posterior mediastinum, it is preferred to use a vein stripping “inversion” technique (Fig. 25-69A). The details of this operation are too lengthy to include in this text, but include the laparoscopic creation of a neo-esophagus (gastric conduit) along the greater curvature of the stomach using the right gastroepiploic artery as the primary vascular pedicle. The conduit can be created through a mini-laparotomy or laparoscopically. A Kocher maneuver releases the duodenum, and a pyloroplasty may be performed (optional). Retrograde esophageal stripping is performed by dividing the esophagus below the GEJ and sliding a vein stripper from the neck down into the abdomen followed by an inversion of the esophagus in the posterior mediastinum and removal through the neck (Fig. 25-69B). This technique is reserved for patients with high-grade dysplasia. For small cancers at the GEJ, the esophagus can be stripped in an antegrade fashion by sliding the vein stripper down from the cervical incision and out the tail of the lesser curvature (Fig. 25-69C). The tail of the lesser curvature is pulled out a port site high in the epigastrium while the esophagus is inverted into itself. For GEJ cancers, a wide celiac access LN dissection, splenic artery, hepatic artery, and posterior mediastinal LN dissection can be performed as well or better than through a laparotomy. The gastric conduit is pulled up to the neck with a chest tube and anastomosed to the cervical esophagus in an end-to-side fashion using a surgical stapler or with a handsewn anastomosis. Complications of this technique are primarily limited to leak from the esophagogastric anastomosis, which is self-limited and usually heals within 1 to 3 weeks, spontaneously.

**Figure 25-69.** A. Laparoscopic retrograde inversion. B. Laparoscopic antegrade inversion. A silk suture holds the tunnel after the esophagus is removed. C. The esophageal conduit is returned to the neck after passing a chest tube down the tunnel and suturing the conduit to the chest tube.
Open Transhiatal Esophagectomy. Transhiatal esophagectomy, also known as blunt esophagectomy or esophagectomy without a thoracotomy, was first performed in 1933 by a British surgeon, but was popularized in the last quarter of the 20th century by Mark Orringer from the University of Michigan. Although this operation may violate many of the principles of cancer resection, including extended radical LN dissection, this operation has performed as well as any of the more radical procedures in randomized trials, and in large database analyses. With transhiatal esophagectomy, the elements of dissection are similar to that described in the section entitled Minimally Invasive Transhiatal Esophagectomy, including the creation of the gastric tube and the posterior mediastinal dissection through the hiatus. Because this dissection is performed with the fingertips rather than under direct vision with surgical instruments, it requires an enlargement of the diaphragmatic hiatus. The lower mediastinal LN basins can be resected as can the upper abdominal LNs, making this an attractive option for GEJ cancers. The mediastinal LNs above the inferior pulmonary vein are not removed with this technique, but they rarely result in a point of isolated cancer recurrence.

Of all procedures for esophageal cancer, this operation is the quickest to perform in experienced hands and lies in an intermediate position between minimally invasive esophagectomy and the Ivor Lewis procedure with respect to complications and recovery.

Minimally Invasive Two- and Three-Field Esophagectomy. After a rocky start, minimally invasive esophagectomy using a thoracic dissection through VATS has become reasonably popular. In general, this operation is performed with an anastomosis created in the neck (three-field), but it may be performed with the anastomosis stapled in the high thorax (two-field). Both procedures will be described.

With a minimally invasive three-field esophagectomy, the patient is placed in the left lateral decubitus position. Double lumen intubation is required. Videoscopic access to the thorax is obtained in the midaxillary line in the ninth intercostal space and an angled telescope illuminates the chest superiorly. A mini-thoracotomy at about the sixth intercostal space anteriorly allows introduction of conventional surgical instruments, and a high trocar allows retraction of the lung away from the esophagus. In a three-field approach, the esophagus is dissected along its length to include division of the azygos vein and harvesting of the LNs in the upper, middle, and lower posterior mediastinum. Hilar, and posterior mediastinal nodes are all removed and sent with the specimen individually. The thoracic duct is divided at the level of the diaphragm and removed with the specimen.

Following complete intrathoracic dissection, the patient is placed in the supine position and five laparoscopic ports are placed as with the MIS transhiatal esophagectomy. The abdominal portions of the operation are identical to those described previously in the section entitled “Minimally Invasive Transhiatal Esophagectomy,” and the gastric conduit is then sewn to the tip of the fully mobilized GEJ and lesser curvature sleeve. A feeding tube is placed, and the pyloroplasty may be performed laparoscopically. A transverse cervical incision and dissection between the sternocleidomastoid and the anterior strap muscles allows access to the cervical esophagus. Great care is made to avoid stretching the recurrent laryngeal nerve. The esophagus and proximal stomach is then pulled up into the neck with the gastric conduit following. Cervical anastomosis is then performed.

The MIS transthoracic two-field esophagectomy is slightly different. In this operation, the abdominal portions of the operation are done first, including placement of the feeding tube, the creation of the conduit, and the sewing of the tip of the conduit to the fully dissected GEJ. The patient is then rolled into the left lateral decubitus position and, through right thoracoscopy, the esophagus is dissected and divided 10 cm above the tumor. Once freed, the specimen is pulled out through the mini-thoracotomy, and an end-to-end anastomosis stapler is introduced through the high corner of the gastric conduit and out a stab wound along the greater curvature. The anvil of the stapler is placed in the proximal esophagus and held with a purse-string, the stapler is docked, the anastomosis is created, and a gastrostomy is then closed with another firing of the GIA stapler. The three-field esophagectomy has the advantage of placing the anastomosis in the neck where leakage is unlikely to create a severe systemic consequence. On the other hand, placement of the anastomosis in the high chest minimizes the risks of injury to structures in the neck, particularly the recurrent laryngeal nerve. Although the leak of the intrathoracic anastomosis may be more likely to bear septic consequences, the incidence of leak is diminished. Other complications of this approach relate to pulmonary and cardiac status. In many series, the most common complication is pneumonia, the second is atrial fibrillation, and the third is anastomotic leak.

Ivor Lewis (En Bloc) Esophagectomy. The theory behind radical transthoracic esophagectomy is that greater removal of LNs and periesophageal tissues diminishes the chance of a positive radial margin and LN recurrence. Although there are no randomized data demonstrating this to be superior to other forms of esophagectomy, there are many retrospective data demonstrating improved survival with greater numbers of LNs harvested. A recent study from Sloan-Kettering demonstrates a direct relationship between the number of negative nodes harvested and long-term survival. Although such a survival advantage may be related to the completeness of resection, extended radical resections may also be a surrogate for experienced surgeons working in great institutions. As a time-honored operation, there is no doubt that en bloc esophagectomy is the standard to which less radical techniques must be compared.

Generally, this operation is started in the abdomen with an upper midline laparotomy and extensive LN dissection in and about the celiac access and its branches, extending into the porta hepatitis and along the splenic artery to the tail of the pancreas. All LNs are removed en bloc with the lesser curvature of the stomach. Unless the tumor extends into the stomach, reconstruction is performed with a greater curvature gastric tube. For GEJ cancers extending significantly into the gastric cardia or fundus, the proximal stomach is removed, and reconstruction is performed with an isoperistaltic section of left colon between the upper esophagus and the remnant stomach, or the colon is connected to a Roux-en-Y limb of jejunum, if total gastrectomy is necessary. In the majority of cases, colon interposition is unnecessary, and a gastric conduit is used.

Following closure of the abdominal incision, the patient is placed in the left lateral decubitus position and an anterolateral thoracotomy is performed through the sixth intercostal space. The azygos vein is divided and the posterior mediastinum is entirely cleaned out to include the thoracic duct, all periaortic tissues, and all tissue in the upper mediastinum along the course of the current laryngeal nerves and in the peribronchial,
hilar, and tracheal LN stations. The proximal stomach is pulled up into the thorax where a conduit is created (if not performed previously) and a handsewn or stapled anastomosis is made between the upper thoracic esophagus and the gastric conduit or transverse colon. Chest tubes are placed, and the patient is taken to the intensive care unit.

Because this is the most radical of dissections, complications are most common, including pneumonia, respiratory failure, atrial fibrillation, chylothorax, anastomotic leak, conduit necrosis, gastrocutaneous fistula, and, if dissection is too near the recurrent laryngeal nerves, hoarseness will occur with an increased risk of aspiration. Tracheobronchial injury resulting in fistulas between the bronchus and conduit may also occur, however rarely. Although this procedure and three-field esophagectomy are fraught with the highest complication rate, the long-term outcome of this procedure provides the greatest survival in many single-center series and retrospective reviews.

**Three-Field Open Esophagectomy.** Three-field open esophagectomy is very similar to a minimally invasive three-field except that all access is through open incisions. This procedure is preferred by certain Japanese surgeons and LN counts achieved through this kind of operation may run from 45 to 60 LNs. Most Western surgeons question the benefit of such radical surgery when it is hard to define a survival advantage. Nonetheless, high intrathoracic cancers probably deserve such an aggressive approach if cure is the goal.

**Salvage Esophagectomy.** Salvage esophagectomy is the nomenclature applied to esophagectomy performed after failure of definitive radiation and chemotherapy. The most frequent scenario is one in which distant disease (bone, lung, brain, or wide LN metastases) renders the patient nonoperable at initial presentation. Then, systemic chemotherapy, usually with radiation of the primary tumor, destroys all foci of metastasis, as demonstrated by CT and CT-PET, but the primary remains present and symptomatic. Following a period of observation, to make sure no new disease will become evident, salvage esophagectomy is performed, usually with an open two-field approach. Surprisingly, the cure rate of salvage esophagectomy is not inconsequential. One in four patients undergoing this operation will be disease free 5 years later, despite the presence of residual cancer in the operative specimen. Because of the dense scarring created by radiation treatment, this procedure is the most technically challenging of all esophagectomy techniques.

**Comparative Studies of Esophagectomy Technique**

**Transthoracic vs. Transhiatal Esophagectomy.** There has been a great debate as to whether en bloc esophagectomy will provide a greater long-term benefit and cure rate in esophageal cancer than transhiatal esophagectomy. In a recent 7-year follow-up of a Dutch study addressing GEJ and lower esophageal cancers, there does not appear to be any benefit to the more extensive dissection despite higher morbidity and mortality. In a subgroup analysis of those with one to eight positive LNs, it did appear that the en bloc transthoracic resection may add to longevity. In another large database analysis of the Surveillance, Epidemiology, and End Results database, transthoracic and transhiatal esophagectomy were compared. In this study, the transhiatal esophagectomy had a greater long-term survival, but when adjusted by cancer stage, this survival benefit disappeared. The mortality and morbidity after transhiatal esophagectomy appeared to be less. Suffice it to say that this debate over the best procedure for esophagectomy remains an open question.

The role of the minimally invasive surgical procedures for a cancer cure will require further study and longer follow-up. It would appear from preliminary analysis that the transhiatal esophagectomy, like its open cousin, may be performed with less morbidity and mortality than the VATS procedure. Long-term survival analyses will require careful follow-up for at least 5 to 10 years after cancer treatment. A recent European multicenter randomized trial comparing open and minimally invasive approaches revealed a highly significant reduction in pulmonary complications in the patients who underwent the minimally invasive approach. There was no difference in procedure-related mortality between the approaches.

**Alternative Therapies**

**Radiation Therapy.** Primary treatment with radiation therapy does not produce results comparable with those obtained with surgery. Currently, the use of radiotherapy is restricted to patients who are not candidates for surgery, and it is usually combined with chemotherapy. Radiation alone is used for palliation of dysphagia, but the benefit is short lived, lasting only 2 to 3 months. Furthermore, the length and course of treatment are difficult to justify in patients with a limited life expectancy. Radiation is effective in patients who have hemorrhage from the primary tumor.

**Adjuvant Chemotherapy.** The proposal to use adjuvant chemotherapy in the treatment of esophageal cancer began when it became evident that most patients develop postoperative systemic metastasis without local recurrence. This observation led to the hypothesis that undetected systemic micrometastasis had been present at the time of diagnosis, and if effective systemic therapy was added to local regional therapy, survival should improve.

Recently, this hypothesis has been supported by the observation of epithelial tumor cells in the bone marrow in 37% of patients with esophageal cancer who were resected for cure. These patients had a greater prevalence of relapse at 9 months after surgery compared to those patients without such cells. Such studies emphasize that hematogenous dissemination of viable malignant cells occurs early in the disease, and that systemic chemotherapy may be helpful if the cells are sensitive to the agent. On the other hand, systemic chemotherapy may be a hindrance, because of its immunosuppressive properties, if the cells are resistant. Unfortunately, current technology is not able to test tumor cell sensitivity to chemotherapeutic drugs. This requires that the choice of drugs be made solely on the basis of their clinical effectiveness against grossly similar tumors.

The decision to use preoperative rather than postoperative chemotherapy was based on the ineffectiveness of chemotherapeutic agents when used after surgery, and animal studies suggesting that agents given before surgery were more effective. The claim that patients who receive chemotherapy before resection are less likely to develop resistance to the drugs is unsupported by hard evidence. The claim that drug delivery is enhanced because blood flow is more robust before patients undergo surgical dissection is similarly flawed, due to the fact that if enough blood reaches the operative site to heal the wound or anastomosis, then the flow should be sufficient to...
deliver chemotherapeutic drugs. There are, however, data supporting the claim that preoperative chemotherapy in patients with esophageal carcinoma can, if effective, facilitate surgical resection by reducing the size of the tumor. This is particularly beneficial in the case of squamous cell tumors above the level of the carina. Reducing the size of the tumor may provide a safer margin between the tumor and the trachea and allow an anastomosis to a tumor-free cervical esophagus just below the cricopharyngeus. Involved margin at this level usually requires a laryngectomy to prevent subsequent local recurrence.

**Preoperative Chemotherapy.** Eight randomized prospective studies of neoadjuvant chemotherapy vs. surgery alone have demonstrated mixed results. For adenocarcinomas of the distal esophagus and proximal stomach, preoperative neoadjuvant 5-fluorouracil (5-FU) and cisplatin chemotherapy has been shown to provide a survival advantage over surgery alone in a well-powered study from the United Kingdom (MRC trial). This trial is one of the few to include enough patients (800) to detect small differences. The trial had a 10% absolute survival benefit at 2 years for the neoadjuvant chemotherapy group. In a second trial from the United Kingdom (MAGIC trial) of distal esophageal and proximal gastric adenocarcinomas, the use of epirubicin in combination with cisplatin and 5-FU also demonstrated a survival advantage for the induction chemotherapy arm with 4 years median follow-up. As a result of these two trials, standard treatment of locally advanced adenocarcinoma in Europe calls for neoadjuvant chemotherapy with one of these two regimens. Most failures are due to distant metastatic disease, underscoring the need for improved systemic therapy. Postoperative septic and respiratory complications may be more common in patients receiving chemotherapy.

**Preoperative Combination Chemo- and Radiotherapy.** Preoperative chemoradiotherapy using cisplatin and 5-FU in combination with radiotherapy has been reported by several investigators to be beneficial in both adenocarcinoma and squamous cell carcinoma of the esophagus. There have been 10 randomized prospective studies (Table 25-13). A recent meta-analysis of these trials demonstrates a 13% survival advantage for neoadjuvant chemoradiation therapy, which is more pronounced for patients with adenocarcinoma than for those with squamous carcinoma (Table 25-14). It was also observed that the benefit for chemotherapy alone (7%) was not as dramatic as for chemoradiotherapy used in the neoadjuvant setting. Additionally, other work has demonstrated the importance of obtaining an R0 (tumor-free) resection as the most important variable determining long-term survival. Although there are no direct, randomized comparisons between chemotherapy and chemoradiation therapy, it appears that the addition of radiation may improve local control of the tumor and may allow a greater opportunity for the surgeon to obtain an R0 resection.

The timing of surgery after chemoradiation induction is generally felt to be optimal between 6 and 8 weeks following the completion of induction therapy. Earlier than this time, active inflammation may make the resection hazardous, and the patients have not had time to recover fully from the chemoradiation. After 8 weeks, edema in the periesophageal tissue starts to turn to scar tissue, making dissection more difficult.

With chemoradiation, the complete response rates for adenocarcinoma range from 17% to 24% (Table 25-15). No tumor is detected in the specimen after esophagectomy. Patients demonstrating a complete response to chemoradiation have a better survival rate than those without complete response, but distant failure remains common.

At present, the strongest predictors of outcome of patients with esophageal cancer are the anatomic extent of the tumor at diagnosis and the completeness of tumor removal by surgical resection. After incomplete resection of an esophageal cancer, the 5-year survival rates are 0% to 5%. In contrast, after complete resection, independent of stage of disease, 5-year survival ranges from 15% to 40%, according to selection criteria and stage distribution. The importance of early recognition and adequate surgical resection cannot be overemphasized. Figure 25-70 is a global algorithm for the management of esophageal carcinoma.

**SARCOMA OF THE ESOPHAGUS**

Sarcomas and carcinosarcomas are rare neoplasms, accounting for approximately 0.1% to 1.5% of all esophageal tumors. They present with the symptom of dysphagia, which does not differ from the dysphagia associated with the more common epithelial carcinoma. Tumors located within the cervical or high thoracic esophagus can cause symptoms of pulmonary aspiration secondary to esophageal obstruction. Large tumors originating at the level of the tracheal bifurcation can produce symptoms of airway obstruction and syncope by direct compression of the tracheobronchial tree and heart (Fig. 25-71). The duration of dysphagia and age of the patients affected with these tumors are similar to those with carcinoma of the esophagus.

A barium swallow usually shows a large polypoid intraluminal esophageal mass, causing partial obstruction and dilatation of the esophagus proximal to the tumor (Fig. 25-72). The smooth polypoid nature of the lesion, although not diagnostic, is distinctive enough to suggest the presence of a sarcoma rather than the more common ulcerating, stenosing carcinoma.

Esophagoscopy commonly shows an intraluminal necrotic mass. When biopsy is attempted, it is important to remove the necrotic tissue until bleeding is seen on the tumor’s surface. When this is not done, the biopsy specimen will show only tissue necrosis. Even when viable tumor is obtained on biopsy, it has been these authors’ experience that it cannot be definitively identified as carcinoma, sarcoma, or carcinosarcoma on the basis of the histology of the portion biopsied. Biopsy results cannot be totally relied on to identify the presence of sarcoma, and it is often the polypoid nature of the lesion that arouses suspicion that it may be something other than carcinoma.

Polypoid sarcomas of the esophagus, in contrast to infiltrating carcinomas, remain superficial to the muscularis propria and are less likely to metastasize to regional LNs. In one series of 14 patients, local extension or tumor metastasis would have prevented a potentially curative resection in only five. Thus, the presence of a large polypoid tumor should not deter the surgeon from resecting the lesion.

Sarcomatous lesions of the esophagus can be divided into epidermoid carcinomas with spindle cell features, such as carcinosarcoma, and true sarcomas that arise from mesenchymal tissue, such as leiomyosarcoma, fibrosarcoma, and rhabdomyosarcoma. Based on current histologic criteria for diagnosis, fibrosarcoma and rhabdomyosarcoma of the esophagus are extremely rare lesions.

Surgical resection of polypoid sarcoma of the esophagus is the treatment of choice because radiation therapy has little
<table>
<thead>
<tr>
<th>YEAR ACTIVATED</th>
<th>TREATMENT SCHEDULE (RADIOThERAPY)</th>
<th>TREATMENT SCHEDULE (CHEMOTHERAPY)</th>
<th>CONCURRENT OR SEQUENTIAL</th>
<th>TUMOR TYPE</th>
<th>SAMPLE SIZE</th>
<th>MEDIAN FOLLOW-UP (MO)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>35 Gy, 1.75 Gy/fraction over 4 wk</td>
<td>Two cycles: cisplatin 20 mg/m² d 1–5; bleomycin 5 mg/m² d 1–5</td>
<td>Sequential</td>
<td>SCC</td>
<td>78</td>
<td>18^a</td>
</tr>
<tr>
<td>1986</td>
<td>40 Gy, 2 Gy/fraction over 4 wk</td>
<td>Two cycles: cisplatin 100 mg/m² d 1; 5-fluorouracil 1000 mg/m² d 1–4</td>
<td>Concurrent</td>
<td>SCC</td>
<td>69</td>
<td>12^a</td>
</tr>
<tr>
<td>1988</td>
<td>20 Gy, 2 Gy/fraction over 12 d</td>
<td>Two cycles: cisplatin 100 mg/m² d 1; 5-fluorouracil 600 mg/m² d 2–5, 22–25</td>
<td>Sequential</td>
<td>SCC</td>
<td>86</td>
<td>12^a</td>
</tr>
<tr>
<td>1989</td>
<td>45 Gy, 1.5 Gy/fraction over 3 wk</td>
<td>Two cycles: cisplatin 20 mg/m² d 1–5; 5-fluorouracil 300 mg/m² d 1–21; vinblastine 1 mg/m² d 1–4</td>
<td>Concurrent</td>
<td>SCC and adenocarcinoma</td>
<td>100</td>
<td>98</td>
</tr>
<tr>
<td>1989</td>
<td>37 Gy, 3.7 Gy/fraction over 2 wk</td>
<td>Two cycles: cisplatin 80 mg/m² d 0–2</td>
<td>Sequential</td>
<td>SCC</td>
<td>293</td>
<td>55</td>
</tr>
<tr>
<td>1990</td>
<td>40 Gy, 2.7 Gy/fraction over 3 wk</td>
<td>Two cycles: cisplatin 75 mg/m² d 7; 5-fluorouracil 15 mg/kg d 1–5</td>
<td>Concurrent</td>
<td>Adenocarcinoma</td>
<td>113</td>
<td>24</td>
</tr>
<tr>
<td>1990</td>
<td>40 Gy, 2.7 Gy/fraction over 3 wk</td>
<td>Two cycles: cisplatin 75 mg/m² d 7; 5-fluorouracil 15 mg/kg d 1–5</td>
<td>Concurrent</td>
<td>SCC</td>
<td>61</td>
<td>10</td>
</tr>
<tr>
<td>1994</td>
<td>35 Gy, 2.3 Gy/fraction over 3 wk</td>
<td>One cycle: cisplatin 80 mg/m² d 1; 5-fluorouracil 800 mg/m² d 2–5</td>
<td>Concurrent</td>
<td>SCC and adenocarcinoma</td>
<td>256</td>
<td>65</td>
</tr>
<tr>
<td>2006</td>
<td>50.4 Gy, 1.8 Gy/fraction over 5.6 wk</td>
<td>Two cycles: cisplatin 60 mg/m² d 1; 5-fluorouracil 1000 mg/m² d 3–5</td>
<td>Concurrent</td>
<td>SCC and adenocarcinoma</td>
<td>56</td>
<td>60</td>
</tr>
<tr>
<td>1999</td>
<td>45.6 Gy, 1.2 Gy/fraction over 28 d</td>
<td>Two cycles: cisplatin 60 mg/m² d 1; 5-fluorouracil 1000 mg/m² d 3–5</td>
<td>Concurrent</td>
<td>SCC</td>
<td>101</td>
<td>25</td>
</tr>
</tbody>
</table>

**Chemotherapy**

<table>
<thead>
<tr>
<th>YEAR ACTIVATED</th>
<th>TREATMENT SCHEDULE (CHEMOTHERAPY)</th>
<th>CONCURRENT OR SEQUENTIAL</th>
<th>TUMOR TYPE</th>
<th>SAMPLE SIZE</th>
<th>MEDIAN FOLLOW-UP (MO)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1982</td>
<td>—</td>
<td>—</td>
<td>SCC</td>
<td>39</td>
<td>20</td>
</tr>
<tr>
<td>1983</td>
<td>—</td>
<td>—</td>
<td>SCC</td>
<td>106</td>
<td>18^a</td>
</tr>
<tr>
<td>1988^c</td>
<td>—</td>
<td>—</td>
<td>SCC</td>
<td>46</td>
<td>75</td>
</tr>
<tr>
<td>1988</td>
<td>—</td>
<td>—</td>
<td>SCC</td>
<td>46</td>
<td>17^a</td>
</tr>
<tr>
<td>1989</td>
<td>—</td>
<td>—</td>
<td>SCC</td>
<td>147</td>
<td>17</td>
</tr>
<tr>
<td>1990</td>
<td>—</td>
<td>—</td>
<td>SCC</td>
<td>160</td>
<td>19^a</td>
</tr>
<tr>
<td>1990</td>
<td>—</td>
<td>—</td>
<td>SCC and adenocarcinoma</td>
<td>467</td>
<td>56</td>
</tr>
<tr>
<td>1992</td>
<td>—</td>
<td>—</td>
<td>SCC</td>
<td>96</td>
<td>24</td>
</tr>
<tr>
<td>1992</td>
<td>—</td>
<td>—</td>
<td>SCC and adenocarcinoma</td>
<td>802</td>
<td>37</td>
</tr>
</tbody>
</table>

^aEstimated as median survival.
^bUnpublished thesis.
^cYear of activation not reported, but imputed.
^dOnly available as an abstract.
SCC = squamous cell carcinoma.

success and the tumors remain superficial, with local invasion or distant metastases occurring late in the course of the disease. As with carcinoma, the absence of both wall penetration and LN metastases is necessary for curative treatment, and surgical resection is consequently responsible for the majority of the reported 5-year survivals. Resection also provides an excellent means of palliating the patient’s symptoms. The surgical technique for resection and the subsequent restoration of the GI continuity is similar to that described for carcinoma.

In these authors’ experience, four of the eight patients with carcinosarcoma survived for 5 years or longer. Even though this number is small, it suggests that resection produces better results in epithelial carcinoma with spindle cell features than in squamous cell carcinoma of the esophagus. Similarly, with leiomyosarcoma of the esophagus, the same scattered reports exist with little information on survival. Of seven patients with leiomyosarcoma, two died from their disease—one in 3 months and the other 4 years and 7 months after resection. The other five patients were reported to have survived more than 5 years.

It is difficult to evaluate the benefits of resection for leiomyoblastoma of the esophagus because of the small number of reported patients with tumors in this location. Most leiomyoblastomas occur in the stomach, and 38% of these patients succumb to the cancer in 3 years. Fifty-five percent of patients with extragastric leiomyoblastoma also die from the disease, within an average of 3 years. Consequently, leiomyoblastoma should be considered a malignant lesion and apt to behave like a leiomyosarcoma. The presence of nuclear hyperchromatism, increased mitotic figures (more than one per high-power field), tumor size larger than 10 cm, and clinical symptoms of longer than 6 months’ duration are associated with a poor prognosis.

### Table 25-14

<table>
<thead>
<tr>
<th>RISK GROUP</th>
<th>2-Y SURVIVAL RATE (%)</th>
<th>EXPECTED 2-Y MORTALITY</th>
<th>CONTROL (%)</th>
<th>TREATED (%)</th>
<th>ARR (%)</th>
<th>NNT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chemoradiotherapy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>20</td>
<td>80</td>
<td>64.8</td>
<td>15.2</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>35</td>
<td>65</td>
<td>52.7</td>
<td>12.3</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>50</td>
<td>50</td>
<td>40.5</td>
<td>9.5</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td><strong>Chemotherapy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>20</td>
<td>80</td>
<td>72.0</td>
<td>12.0</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>35</td>
<td>65</td>
<td>58.5</td>
<td>6.5</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>50</td>
<td>50</td>
<td>45.0</td>
<td>5.0</td>
<td>20</td>
<td></td>
</tr>
</tbody>
</table>

*Based on a 19% relative mortality reduction for those receiving concurrent chemoradiotherapy and a 10% relative mortality reduction for those receiving chemotherapy.

ARR = absolute risk reduction; NNT = number needed to treat to prevent one death.


### Table 25-15

<table>
<thead>
<tr>
<th>INSTITUTION</th>
<th>YEAR</th>
<th>NO. OF PATIENTS</th>
<th>REGIMEN</th>
<th>COMPLETE PATHOLOGIC RESPONSE (%)</th>
<th>SURVIVAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD Anderson</td>
<td>1990</td>
<td>35</td>
<td>P, E, 5-FU</td>
<td>3</td>
<td>42% at 3 y</td>
</tr>
<tr>
<td>SLMC</td>
<td>1992</td>
<td>18</td>
<td>P, 5-FU, RT</td>
<td>17</td>
<td>40% at 3 y</td>
</tr>
<tr>
<td>Vanderbilt</td>
<td>1993</td>
<td>39</td>
<td>P, E, 5-FU, RT</td>
<td>19</td>
<td>47% at 4 y</td>
</tr>
<tr>
<td>Michigan</td>
<td>1993</td>
<td>21</td>
<td>P, VBL, 5-FU, RT</td>
<td>24</td>
<td>34% at 5 y</td>
</tr>
<tr>
<td>MGH</td>
<td>1994</td>
<td>16</td>
<td>P, 5-FU</td>
<td>0</td>
<td>42% at 4 y</td>
</tr>
<tr>
<td>MGH</td>
<td>1994</td>
<td>22</td>
<td>E, A, P</td>
<td>5</td>
<td>58% at 2 y</td>
</tr>
</tbody>
</table>

A = doxorubicin; E = etoposide; 5-FU = 5-fluouracil; MGH = Massachusetts General Hospital; P = cisplatin; RT = radiation therapy; SLMC = St. Louis University Medical Center; VBL = vinblastine.

polyps, or simply as polyps. Pedunculated intraluminal tumors should be removed. If the lesion is not too large, endoscopic removal with a snare is feasible.

**Leiomyoma**

Leiomyomas constitute more than 50% of benign esophageal tumors. The average age at presentation is 38, which is in sharp contrast to that seen with esophageal carcinoma. Leiomyomas are twice as common in males. Because they originate in smooth muscle, 90% are located in the lower two-thirds of the esophagus. They are usually solitary, but multiple tumors have been found on occasion. They vary greatly in size and shape. Actually, tumors as small as 1 cm in diameter and as large as 10 lb have been removed.

Typically, leiomyomas are oval. During their growth, they remain intramural, having the bulk of their mass protruding toward the outer wall of the esophagus. The overlying mucosa is freely movable and normal in appearance. Dysphagia and pain are the most common complaints, the two symptoms occurring more frequently together than separately. Bleeding directly related to the tumor is rare, and when hematemesis or melena occur in a patient with an esophageal leiomyoma, other causes should be investigated.

A barium swallow is the most useful method to demonstrate a leiomyoma of the esophagus (Fig. 25-73). In profile, the tumor appears as a smooth, semilunar, or crescent-shaped filling defect that moves with swallowing, is sharply demarcated, and is covered and surrounded by normal mucosa. Esophagoscopy should be performed to exclude the reported observation of a coexistence with carcinoma. The freely movable mass, which bulges into the lumen, should not be biopsied because of an increased chance of mucosal perforation at the time of surgical enucleation. Endoscopic ultrasound is also a useful adjunct in the workup of leiomyoma and provides detail related to the anatomic extent and relationship to surrounding structures.

Despite their slow growth and limited potential for malignant degeneration, leiomyomas should be removed unless there are specific contraindications. The majority can be removed by simple enucleation. If, during removal, the mucosa is inadvertently entered, the defect can be repaired primarily. After tumor removal, the outer esophageal wall should be reconstructed by closure of the muscle layer. The location of the lesion and the

---

**Figure 25-70.** Suggested global algorithm for the management of carcinoma of the esophagus. CT = computed tomography.
**Figure 25-71.** A. Computed tomographic scan of a leiomyosarcoma (*black arrow*) that caused compression of the heart and symptoms of syncope. B. Surgical specimen of leiomyosarcoma shown in A with a pedunculated luminal lesion (*white arrow*) and a large extraesophageal component (*black arrow*). There was no evidence of lymph node metastasis at the time of operation.

**Figure 25-72.** A. Barium swallow showing a large polypoid intraluminal esophageal mass causing partial obstruction and dilation of the proximal esophagus. B. Operative specimen showing 9-cm polypoid leiomyoblastoma.
extent of surgery required will dictate the approach. Lesions of the proximal and middle esophagus require a right thoracotomy, whereas distal esophageal lesions require a left thoracotomy. Videothoracoscopic and laparoscopic approaches are now frequently used. The mortality rate associated with enucleation is low, and success in relieving the dysphagia is near 100%. Large lesions or those involving the GEJ may require esophageal resection.

**Esophageal Cyst**

Cysts may be congenital or acquired. Congenital cysts are lined wholly or partly by columnar ciliated epithelium of the respiratory type, by glandular epithelium of the gastric type, by squamous epithelium, or by transitional epithelium. In some, epithelial lining cells may be absent. Confusion over the embryologic origin of congenital cysts has led to a variety of names, such as enteric, bronchogenic, duplication, and mediastinal cysts. Acquired retention cysts also occur, probably as a result of obstruction of the excretory ducts of the esophageal glands.

Enteric and bronchogenic cysts are the most common, and they arise as a result of developmental abnormalities during the formation and differentiation of the lower respiratory tract, esophagus, and stomach from the foregut. During its embryologic development, the esophagus is lined successively with simple columnar, pseudostratified ciliated columnar, and, finally, stratified squamous epithelium. This sequence probably accounts for the fact that the lining epithelium may be any or a combination of these; the presence of cilia does not necessarily indicate a respiratory origin.

Cysts vary in size from small to very large, and they are usually located intramurally in the middle- to lower-third of the esophagus. Their symptoms are similar to those of a leiomyoma. The diagnosis similarly depends on radiographic, endoscopic, and endosonographic findings. Surgical excision by enucleation is the preferred treatment. During removal, a fistulous tract connecting the cysts to the airways should be sought, particularly in patients who have had repetitive bronchopulmonary infections.

ESOPHAGEAL PERFORATION

Perforation of the esophagus constitutes a true emergency. It most commonly occurs following diagnostic or therapeutic procedures. Spontaneous perforation, referred to as Boerhaave’s syndrome, accounts for only 15% of cases of esophageal perforation, foreign bodies for 14%, and trauma for 10%. Pain is a striking and consistent symptom and strongly suggests that an esophageal rupture has occurred, particularly if located in the cervical area following instrumentation of the esophagus, or sub- sternally in a patient with a history of resisting vomiting. If subcutaneous emphysema is present, the diagnosis is almost certain.

Spontaneous rupture of the esophagus is associated with a high mortality rate because of the delay in recognition and treatment. Although there usually is a history of resisting vomiting, in a small number of patients, the injury occurs silently, without any antecedent history. When the chest radiogram of a patient with an esophageal perforation shows air or an effusion in the pleural space, the condition is often misdiagnosed as a pneumothorax or pancreatitis. An elevated pleural amylase caused by the extrusion of saliva through the perforation may fix the diagnosis of pancreatitis in the mind of an unwary physician. If the chest radiogram is normal, a mistaken diagnosis of myocardial infarction or dissecting aneurysm is often made.

Spontaneous rupture usually occurs into the left pleural cavity or just above the GEJ. About 50% of patients have concomitant GERD, suggesting that minimal resistance to the transmission of abdominal pressure into the thoracic esophagus is a factor in the pathophysiology of the lesion. During vomiting, high peaks of intragastric pressure can be recorded, frequently exceeding 200 mmHg, but because extragastric pressure remains almost equal to intragastric pressure, stretching of the gastric wall is minimal. The amount of pressure transmitted to the esophagus varies considerably, depending on the position of the GEJ. When it is in the abdomen and exposed to intra-abdominal pressure, the pressure transmitted to the esophagus is much less than when it is exposed to the negative thoracic pressure. In the latter situation, the pressure in the lower esophagus will frequently equal intragastric pressure if the glottis remains closed. Cadaver studies have shown that when this pressure exceeds 150 mmHg, rupture of the esophagus is apt to occur. When a hiatal hernia is present and the sphincter remains exposed to abdominal pressure, the lesion produced is usually a Mallory-Weiss mucosal tear, and bleeding rather than perforation is the problem. This is due to the stretching of the supradiaphragmatic portion of the gastric wall. In this situation, the hernia sac represents an extension of the abdominal cavity, and the GEJ remains exposed to abdominal pressure.

**Diagnosis**

Abnormalities on the chest radiogram can be variable and should not be depended upon to make the diagnosis. This is because the abnormalities are dependent on three factors: (a) the time interval between the perforation and the radiographic examination, (b) the site of perforation, and (c) the integrity of the mediastinal pleura. Mediastinal emphysema, a strong indicator of perforation, takes at least 1 hour to be demonstrated and is present in only 40% of patients. Mediastinal widening secondary to edema may not occur for several hours. The site of perforation also can influence the radiographic findings. In cervical perforation, cervical emphysema is common and mediastinal emphysema rare; the converse is true for thoracic perforations.
Frequently, air will be visible in the erector spinae muscles on a neck radiogram before it can be palpated or seen on a chest radiogram (Fig. 25-74). The integrity of the mediastinal pleura influences the radiographic abnormality in that rupture of the pleura results in a pneumothorax, a finding that is seen in 77% of patients. In two-thirds of patients, the perforation is on the left side; in one-fifth, it is on the right side; and in one-tenth, it is bilateral. If pleural integrity is maintained, mediastinal emphysema (rather than a pneumothorax) appears rapidly. A pleural effusion secondary to inflammation of the mediastinum occurs late. In 9% of patients, the chest radiogram is normal.

The diagnosis is confirmed with a contrast esophagram, which will demonstrate extravasation in 90% of patients. The use of a water-soluble medium such as Gastrografin is preferred. Of concern is that there is a 10% false-negative rate. This may be due to obtaining the radiographic study with the patient in the upright position. When the patient is upright, the passage of water-soluble contrast material can be too rapid to demonstrate a small perforation. The studies should be done with the patient in the right lateral decubitus position (Fig. 25-75). In this, the contrast material fills the entire length of the esophagus, allowing the actual site of perforation and its interconnecting cavities to be visualized in almost all patients.

Management
The key to optimum management is early diagnosis. The most favorable outcome is obtained following primary closure of the perforation within 24 hours, resulting in 80% to 90% survival. Figure 25-76 is an operative photograph taken through a left thoracotomy of an esophageal rupture following a pneumatic dilation for achalasia. The most common location for the injury is the left lateral wall of the esophagus, just above the GEJ.

To get adequate exposure of the injury, a dissection similar to that described for esophageal myotomy is performed. A flap of stomach is pulled up and the soiled fat pad at the GEJ is removed. The edges of the injury are trimmed and closed primarily (Fig. 25-77). The closure is reinforced with the use of a pleural patch or construction of a Nissen fundoplication.

Mortality associated with immediate closure varies between 8% and 20%. After 24 hours, survival decreases to <50%, and is not influenced by the type of operative therapy (i.e., drainage alone or drainage plus closure of the perforation). If the time delay before closing a perforation approaches 24 hours and the tissues are inflamed, division of the cardia and resection of the diseased portion of the esophagus are recommended. The remainder of the esophagus is mobilized, and as much normal esophagus as possible is saved and brought out as an end cervical esophagostomy. In some situations, the retained esophagus may be so long that
it loops down into the chest. The contaminated mediastinum is drained and a feeding jejunostomy tube is inserted. The recovery from sepsis is often immediate, dramatic, and reflected by a marked improvement in the patient’s condition over a 24-hour period. On recovery from the sepsis, the patient is discharged and returns on a subsequent date for reconstruction with a substernal colon interposition. Failure to apply this aggressive therapy can result in a mortality rate in excess of 50% in patients in whom the diagnosis has been delayed.

Nonoperative management of esophageal perforation has been advocated in select situations. The choice of conservative therapy requires skillful judgment and necessitates careful radiographic examination of the esophagus. This course of management usually follows an injury occurring during dilation of esophageal strictures or pneumatic dilations of achalasia. Conservative management should not be used in patients who have free perforations into the pleural space. Cameron proposed three criteria for the nonoperative management of esophageal perforation: (a) the esophagram must show the perforation to be contained within the mediastinum and drain well back into the esophagus (Fig. 25-78), (b) symptoms should be mild, and (c) there should be minimal evidence of clinical sepsis. If these conditions are met, it is reasonable to treat the patient with hyperalimentation, antibiotics, and cimetidine to decrease acid secretion and diminish pepsin activity. Oral intake is resumed in 7 to 14 days, dependent on subsequent radiographic examinations.

**MALLORY-WEISS SYNDROME**

In 1929, Mallory and Weiss described four patients with acute upper GI bleeding who were found at autopsy to have mucosal tears at the GEJ. This syndrome, characterized by acute upper GI bleeding following vomiting, is considered to be the cause of up to 15% of all severe upper GI bleeds. The mechanism is similar to spontaneous esophageal perforation: an acute increase in intra-abdominal pressure against a closed glottis in a patient with a hiatal hernia.

Mallory-Weiss tears are characterized by arterial bleeding, which may be massive. Vomiting is not an obligatory factor, as there may be other causes of an acute increase in intra-abdominal pressure, such as paroxysmal coughing, seizures, and retching. The diagnosis requires a high index of suspicion, particularly in the patient who develops upper GI bleeding following prolonged vomiting or retching. Upper endoscopy confirms the suspicion by identifying one or more longitudinal fissures in the mucosa of the herniated stomach as the source of bleeding.

In the majority of patients, the bleeding will stop spontaneously with nonoperative management. In addition to blood replacement, the stomach should be decompressed and antiemetics administered, as a distended stomach and continued vomiting aggravate further bleeding. A Sengstaken-Blakemore tube will not stop the bleeding, as the pressure in the balloon is not sufficient to overcome arterial pressure. Endoscopic injection of epinephrine may be therapeutic if bleeding does not stop spontaneously. Only occasionally will surgery be required to stop blood loss. The procedure consists of laparotomy and high gastrostomy with oversewing of the linear tear. Mortality is uncommon, and recurrence is rare.
CAUSTIC INJURY

Accidental caustic lesions occur mainly in children, and, in general, rather small quantities of caustics are taken. In adults or teenagers, the swallowing of caustic liquids is usually deliberate, during a suicide attempt, and greater quantities are swallowed. Alkalis are more frequently swallowed accidentally than acids, because strong acids cause an immediate burning pain in the mouth.

Pathology
The swallowing of caustic substances causes an acute and a chronic injury. During the acute phase, care focuses on controlling the immediate tissue injury and the potential for perforation. During the chronic phase, the focus is on treatment of strictures and disturbances in pharyngeal swallowing. In the acute phase, the degree and extent of the lesion are dependent on several factors: the nature of the caustic substance, its concentration, the quantity swallowed, and the time the substance is in contact with the tissues.

Acids and alkalis affect tissue in different ways. Alkalis dissolve tissue, and therefore penetrate more deeply, while acids cause a coagulative necrosis that limits their penetration. Animal experiments have shown that there is a correlation between the depth of the lesion and the concentration of sodium hydroxide solution. When a solution of 3.8% comes into contact with the esophagus for 10 seconds, it causes necrosis of the mucosa and the submucosa but spares the muscular layer. A concentration of 22.5% penetrates the whole esophageal wall and into the peri-esophageal tissues. Cleansing products can contain up to 90% sodium hydroxide. The strength of esophageal contractions varies according to the level of the esophagus, being weakest at the striated muscle–smooth muscle interface. Consequently, clearance from this area may be somewhat slower, allowing caustic substances to remain in contact with the mucosa longer. This explains why the esophagus is preferentially and more severely affected at this level than in the lower portions.

The lesions caused by lye injury occur in three phases. First is the acute necrotic phase, lasting 1 to 4 days after injury. During this period, coagulation of intracellular proteins results in cell necrosis, and the living tissue surrounding the area of necrosis develops an intense inflammatory reaction. Second is the ulceration and granulation phase, starting 3 to 5 days after injury. During this period, the superficial necrotic tissue sloughs, leaving an ulcerated, acutely inflamed base, and granulation tissue fills the defect left by the sloughed mucosa. This phase lasts 10 to 12 days, and it is during this period that the esophagus is the weakest. Third is the phase of cicatrization and scarring, which begins the third week following injury. During this period, the previously formed connective tissue begins to contract, resulting in narrowing of the esophagus. Of the patients who develop strictures, 60% do so within 1 month, and 80% within 2 months. If dysphagia does not develop within 8 months, it is unlikely that a stricture will occur. Serious systemic reactions such as hypovolemia and acidosis resulting in renal damage can occur in cases in which the burns have been caused by strong acids. Respiratory complications such as laryngospasm, laryngoedema, and occasionally pulmonary edema can occur, especially when strong acids are aspirated.

Inspection of the oral cavity and pharynx can indicate that caustic substances were swallowed, but does not reveal that the esophagus has been burned. Conversely, esophageal burns can be present without apparent oral injuries. Because of this poor correlation, early esophagoscopy is advocated to establish the presence of an esophageal injury. To lessen the chance of perforation, the scope should not be introduced beyond the proximal esophageal lesion. The degree of injury can be graded according to the criteria listed in Table 25-16. Even if the esophagoscopy is normal, strictures may appear later. Radiographic examination is not a reliable means to identify the presence of early esophageal injury, but it is important in later follow-up to identify strictures. The most common locations of caustic injuries are shown in Table 25-17.

Clinical Manifestations
The clinical picture of an esophageal burn is determined by the degree and extent of the lesion. In the initial phase, complaints consist of pain in the mouth and substernal region, hypersalivation, pain on swallowing, and dysphagia. The presence of fever is strongly correlated with the presence of an esophageal lesion. Bleeding can occur, and, frequently, the patient vomits. These initial complaints disappear during the quiescent period of ulceration and granulation. During the cicatrization and scarring phase, the complaint of dysphagia reappears and is due to fibrosis and retraction, resulting in narrowing of the esophagus. Of the patients who develop strictures, 60% do so within 1 month, and 80% within 2 months. If dysphagia does not develop within 8 months, it is unlikely that a stricture will occur. Serious systemic reactions such as hypovolemia and acidosis resulting in renal damage can occur in cases in which the burns have been caused by strong acids. Respiratory complications such as laryngospasm, laryngoedema, and occasionally pulmonary edema can occur, especially when strong acids are aspirated.

Treatment
Treatment of a caustic lesion of the esophagus is directed toward management of both the immediate and late consequences of the injury. The immediate treatment consists of limiting the burn by administering neutralizing agents. To be effective, this must be done within the first hour. Lye or other alkali can be neutralized with half-strength vinegar, lemon juice, or orange juice. Acid can be neutralized with milk, egg white, or antacids. Sodium bicarbonate is not used because it generates carbon dioxide,
which might increase the danger of perforation. Emetics are contraindicated because vomiting renews the contact of the caustic substance with the esophagus and can contribute to perforation if too forceful. Hypovolemia is corrected, and broad-spectrum antibiotics are administered to lessen the inflammatory reaction and prevent infectious complications. If necessary, a feeding jejunostomy tube is inserted to provide nutrition. Oral feeding can be started when the dysphagia of the initial phase has regressed.

In the past, surgeons waited until the appearance of a stricture before starting treatment. Currently, dilations are started the first day after the injury, with the aim of preserving the esophageal lumen by removing the adhesions that occurred in the injured segments. However, this approach is controversial in that dilations can traumatize the esophagus, causing bleeding, and perforation, and there are data indicating that excessive dilations cause increased fibrosis secondary to the added trauma. The use of steroids to limit fibrosis has been shown to be effective in animals, but their effectiveness in human beings has not been established.

Extensive necrosis of the esophagus frequently leads to perforation, and it is best managed by resection. When there is extensive gastric involvement, the esophagus is nearly always necrotic or severely burned, and total gastrectomy and near-total esophagectomy are necessary. The presence of air in the esophageal wall is a sign of muscle necrosis and impending perforation and is a strong indication for esophagectomy.

Management of acute injury is summarized in the algorithm in Fig. 25-79. Some authors have advocated the use of an intraluminal esophageal stent (Fig. 25-80) in patients who are operated on and found to have no evidence of extensive esophagogastric necrosis. In these patients, a biopsy of the posterior gastric wall should be performed to exclude occult injury. If, histologically, there is a question of viability, a second-look operation should be done within 36 hours. If a stent is inserted, it should be kept in position for 21 days, and removed after a satisfactory barium esophagogram. Esophagoscopy should be done, and if strictures are present, dilations initiated.

Once the acute phase has passed, attention is turned to the prevention and management of strictures. Both antegrade dilation with a Hurst or Maloney bougie and retrograde dilation with a Tucker bougie have been satisfactory. In a series of 1079 patients, early dilations started during the acute phase gave excellent results in 78%, good results in 13%, and poor results in 2%. During the treatment, 55 patients died. In contrast, of 333 patients whose strictures were dilated when they became symptomatic, only 21% had excellent results, 46% good, and 6% poor, with three dying during the process. The length of time the surgeon should persist with dilation before consideration of esophageal resection is problematic. An adequate lumen should be re-established within 6 months to 1 year, with progressively longer intervals between dilations. If, during the course of treatment, an adequate lumen cannot be established or maintained (i.e., smaller bougies must be used), operative intervention should be considered. Surgical intervention is indicated when there is (a) complete stenosis in which all attempts from above and below have failed to establish a lumen, (b) marked irregularity and pocketing on barium swallow, (c) the development of a severe periesophageal reaction or mediastinitis with dilatation, (d) a fistula, (e) the inability to dilate or maintain the lumen above a 40F bougie, or (f) a patient who is unwilling or unable to undergo prolonged periods of dilation.

Figure 25-79. Algorithm summarizing the management of acute caustic injury.

Figure 25-80. The use of an esophageal stent to prevent stricture. The stent is constructed from a chest tube and placed in the esophagus at the time of an exploratory laparotomy. A Penrose drain is placed over the distal end as a flap valve to prevent reflux. The stent is supported at its upper end by attaching it to a suction catheter that is secured to the nares. Continuous suction removes saliva and mucus trapped in the pharynx and upper esophagus.
The variety of abnormalities seen requires that creativity be used when considering esophageal reconstruction. Skin tube esophagogastoplasties are now used much less frequently than they were in the past, and are mainly of historical interest. Currently, the stomach, jejunum, and colon are the organs used to replace the esophagus, through either the posterior mediastinum or the retrosternal route. A retrosternal route is chosen when there has been a previous esophagectomy or there is extensive fibrosis in the posterior mediastinum. When all factors are considered, the order of preference for an esophageal substitute is (a) colon, (b) stomach, and (c) jejunum. Free jejunal grafts based on the superior thyroid artery have provided excellent results. Whatever method is selected, it must be emphasized that these procedures cannot be taken lightly; minor errors of judgment or technique may lead to serious or even fatal complications.

Critical in the planning of the operation is the selection of cervical esophagus, pyriform sinus, or posterior pharynx as the site for proximal anastomosis. The site of the upper anastomosis depends on the extent of the pharyngeal and cervical esophageal damage encountered. When the cervical esophagus is destroyed and a pyriform sinus remains open the anastomosis can be made to the hypopharynx (Fig. 25-81). When the pyriform sinuses are completely stenosed, a transglottic approach is used to perform an anastomosis to the posterior oropharyngeal wall (Fig. 25-82). This allows excision of supraglottic strictures and elevation and anterior tilting of the larynx. In both of these situations, the patient must relearn to swallow. Recovery is long and difficult and may require several endoscopic dilations—and often reoperations. Sleeve resections of short strictures are not successful because the extent of damage to the wall of the esophagus can be greater than realized, and almost invariably the anastomosis is carried out in a diseased area.

The management of a bypassed damaged esophagus after injury is problematic. If the esophagus is left in place, ulceration from gastroesophageal reflux or the development of carcinoma must be considered. The extensive dissection necessary to remove the esophagus, particularly in the presence of marked periesophagitis, is associated with significant morbidity. Leaving the esophagus in place preserves the function of the vagus nerves, and, in turn, the function of the stomach. On the other hand, leaving a damaged esophagus in place can result in multiple blind sacs and subsequent development of mediastinal abscesses years later. Most experienced surgeons recommend that the esophagus be removed unless the operative risk is unduly high.

**ACQUIRED FISTULA**

The esophagus lies in close contact with the membranous portion of the trachea and left bronchus, predisposing to the formation of fistula to these structures. Most acquired esophageal fistulas are to the tracheobronchial tree and secondary to either esophageal or pulmonary malignancy. Traumatic fistulas and those associated with esophageal diverticula account for the remainder. Fistulas associated with traction diverticula are usually due to mediastinal inflammatory disease, and traumatic fistulas usually occur secondary to penetrating wounds, lye ingestion, or iatrogenic injury.

These fistulas are characterized by paroxysmal coughing following the ingestion of liquids, and by recurrent or chronic pulmonary infections. The onset of cough immediately after swallowing suggests aspiration, whereas a brief delay (30–60 seconds) suggests a fistula.

Spontaneous closure is rare, owing to the presence of malignancy or a recurrent infectious process. Surgical treatment of benign fistulas consists of division of the fistulous tract, resection of irreversibly damaged lung tissue, and closure of the esophageal defect. To prevent recurrence, a pleural flap should be interposed. Treatment of malignant fistulas is difficult, particularly in the presence of prior irradiation. Generally, only palliative treatment is indicated. This can best be done by using a specially designed esophageal endoprosthesis that bridges and occludes the fistula, allowing the patient to eat. A salivary tube is also a good option for proximal esophageal fistulas. This tube has a proximal “lip” that rests on the cricopharyngeal muscle and thereby directs the saliva into the tube and past the fistula. Rarely, esophageal diversion, coupled with placement of a feeding jejunostomy, can be used as a last resort.

**Figure 25-81.** Anastomosis of the bowel to a preserved pyriform sinus. To identify the site, a finger is inserted into the free pyriform sinus through a suprahyoid incision (dotted line). This requires removing the lateral inferior portion of the thyroid cartilage as shown in cross-section.

**Figure 25-82.** Anastomosis of the bowel to the posterior oropharynx. The anastomosis is done through an inverted trapezoid incision above the thyroid cartilage (dotted line). A triangle-shaped piece of the upper half of the cartilage is resected. Closure of the oropharynx is done so that the larynx is pulled up (sagittal section).
TECHNIQUES OF ESOPHAGEAL RECONSTRUCTION

Options for esophageal substitution include gastric advancement, colonic interposition, and either jejunal free transfer or advancement into the chest. Rarely, combinations of these grafts will be the only possible option. The indications for esophageal resection and substitution include malignant and end-stage benign disease. The latter includes reflux- or drug-induced stricture formation that cannot be dilated without damage to the esophagus, a dilated and tortuous esophagus secondary to severe motility disorders, lye-induced strictures, and multiple previous antireflux procedures. The choice of esophageal substitution has significant impact upon the technical difficulty of the procedure and influences the long-term outcome.

Partial Esophageal Resection

Distal benign lesions, with preserved proximal esophageal function, are best treated with the interposition of a segment of proximal jejunum into the chest and primary anastomosis. A jejunal interposition can reach to the inferior border of the pulmonary hilum with ease, but the architecture of its blood supply rarely allows the use of the jejunum proximal to this point. Because the anastomosis is within the chest, a thoracotomy is necessary.

The jejunum is a dynamic graft and contributes to bolus transport, whereas the stomach and colon function more as a conduit. The stomach is a poor choice in this circumstance because of the propensity for the reflux of gastric contents into the proximal remaining esophagus following an intrathoracic esophagogastrectomy. It is now well recognized that this occurs and can lead to incapacitating symptoms and esophageal destruction in some patients. Short segments of colon, on the other hand, lack significant motility and have a propensity for the development of esophagitis proximal to the anastomosis.

Replacement of the cervical portion of the esophagus, while preserving the distal portion, is occasionally indicated in cervical esophageal or head and neck malignancy, and following the ingestion of lye. Free transfer of a portion of jejunum to the neck has become a viable option and is successful in the majority of cases. Revascularization is achieved via use of the internal mammary artery and the internal mammary or innominate vein. Removal of the sternoclavicular joint aids in performing the vascular and distal esophageal anastomosis (Fig. 25-83).

Reconstruction After Total Esophagectomy

Neither the intrathoracic stomach nor the intrathoracic colon functions as well as the native esophagus after an esophagogastrectomy. The choice between these organs will be influenced by several factors, such as the adequacy of their blood supply and the length of resected esophagus that they are capable of bridging. If the stomach shows evidence of disease, or has been contracted or reduced by previous gastric surgery, the length available for esophageal replacement may not be adequate. The presence of diverticulitis, unrecognized carcinoma, or colitis prohibits the use of the colon. The blood supply of the colon is more affected by vascular disease than the blood supply of the stomach, which may prevent its use. Of the two, the colon provides the longest graft. The stomach can usually reach to the neck if the amount of lesser curvature resected does not interfere with the blood supply to the fundus. Gastric interposition has the advantage that only one anastomosis is required. On the other hand, there is greater potential for aspiration of gastric juice or stricture of the cervical anastomosis from chronic reflux when stomach is used for replacement.

Following an esophagogastrectomy, patients may have discomfort during or shortly after eating. The most common symptom is a postprandial pressure sensation or a feeling of being full, which probably results from the loss of the gastric reservoir. This symptom is less common when the colon is used as an esophageal substitute, probably because the distal third of the stomach is retained in the abdomen and the interposed colon provides an additional reservoir function.

King and Hölscher have reported a 40% and 50% incidence of dysphagia after reestablishing GI continuity with the stomach following esophagogastrectomy. This incidence is similar to Orringer’s results after using the stomach to replace the esophagus in patients with benign disease. More than one-half of the patients experienced dysphagia postoperatively;
two-thirds of this group required postoperative dilation, and one-fourth had persistent dysphagia and required home dilation. In contrast, dysphagia is uncommon, and the need for dilation is rare following a colonic interposition. Isolauri reported on 248 patients with colonic interpositions and noted a 24% incidence of dysphagia 12 months after the operation. When it occurred, the most common cause was recurrent mediastinal tumor. The high incidence of dysphagia with the use of the stomach is probably related to the esophagogastric anastomosis in the neck and the resulting difficulty of passing a swallowed bolus.

Another consequence of the transposition of the stomach into the chest is the development of postoperative duodenogastric reflux, probably due to pyloric denervation, and adding a pyloroplasty may worsen this problem. Following gastric advancement, the pylorus lies at the level of the esophageal hiatus, and a distinct pressure differential develops between the intrathoracic gastric and intra-abdominal duodenal lumina. Unless the pyloric valve is extremely efficient, the pressure differential will encourage reflux of duodenal contents into the stomach. Duodenogastric reflux is less likely to occur following colonic interposition because there is sufficient intra-abdominal colon to be compressed by the abdominal pressure and the pylorus and duodenum remain in their normal intra-abdominal position.

Although there is general acceptance of the concept that an esophagogastric anastomosis in the neck results in less postoperative esophagitis and stricture than one at a lower level, reflux esophagitis following a cervical anastomosis does occur, albeit at a lower rate than when the anastomosis is at a lower level. Most patients undergo cervical esophagogastrectomy for malignancy; thus, the long-term sequelae of an esophagogastric anastomosis in the neck are not of concern. However, patients who have had a cervical esophagogastrectomy for benign disease may develop problems associated with the anastomosis in the fourth or fifth postoperative year that are severe enough to require anastomotic revision. This is less likely in patients who have had a colonic interposition for esophageal replacement. Consequently, in patients who have a benign process or a potentially curable carcinoma of the esophagus or cardia, a colonic interposition is used to obviate the late problems associated with a cervical esophagogastrectomy. Colonic interposition for esophageal substitution is a more complex procedure than gastric advancement, with the potential for greater perioperative morbidity, particularly in inexperienced hands.

**Composite Reconstruction**

Occasionally, a combination of colon, jejunum, and stomach is the only reconstructive option available. This situation may arise when there has been previous gastric or colonic resection, when dysphagia has recurred after a previous esophageal resection, or following postoperative complications such as ischemia of an esophageal substitute. Although not ideal, combinations of colon, jejunum, and stomach used to restore GI continuity function surprisingly well and allow alimentary reconstruction in an otherwise impossible situation.

**Vagal Sparing Esophagectomy With Colon Interposition**

Traditional esophagectomy typically results in bilateral vagotomy and its attendant consequences. It is likely that symptoms such as dumping, diarrhea, early satiety, and weight loss seen in 15% to 20% of patients postesophagectomy are at least in part, if not completely, due to vagal interruption. The technique of vagal sparing esophagectomy with colon interposition has been described in an effort to avoid the morbidities associated with standard esophagectomy.

Through an upper midline abdominal incision, the right and left vagal nerves are identified, circled with a tape, and retracted to the right. A limited, highly selective proximal gastric vagotomy is performed along the cephalad 4 cm of the lesser curvature. The stomach is divided with an Endo-GIA stapler just below the GEJ. The colon is prepared to provide an interposed segment as previously described. A neck incision is made along the anterior border of the left sternocleidomastoid muscle, and the strap muscles are exposed. The omohyoid muscle is divided at its pulley, and the sternohyoid and sternothyroid muscles are divided at their manubrial insertion. The left carotid sheath is retracted laterally and the thyroid and trachea medially. The left inferior thyroid artery is ligated laterally as it passes under the left common carotid artery. The left recurrent laryngeal nerve is identified and protected. The esophagus is dissected circumferentially in an inferior direction, from the left neck to the apex of the right chest, to avoid injury to the right recurrent laryngeal nerve. The esophagus is divided at the level of the thoracic inlet, leaving about 3 to 4 cm of cervical esophagus. The proximal esophagus is retracted anteriorly and to the right with the use of two sutures to keep saliva and oral contents from contaminating the neck wound.

Returning to the abdomen, the proximal staple line of the gastric division is opened, and the esophagus is flushed with povidone-iodine solution. A vein stripper is passed up the esophagus into the neck wound. The distal portion of the esophagus in the neck is secured tightly around the stripping cable with “endoloops” and an umbilical tape for a trailer. The tip of the stripper is exchanged for a mushroom head, and the stripper is pulled back into the abdomen, inverting the esophagus as it transverses the posterior mediastinum. This maneuver strips the branches of the esophageal plexus off the longitudinal muscle of the esophagus, preserving the esophageal plexus along with the proximal vagal nerves and the distal vagal nerve trunks. In patients with end-stage achalasia, only the mucosa is secured around the stripping cable, so that it alone is stripped and the dilated muscular wall of the esophagus, with its enriched blood supply, remains. The resulting mediastinal tunnel, or in the case of achalasia the muscular tube, is dilated with a Foley catheter containing 90 mL of fluid in the balloon. The previously prepared interposed portion of the transverse colon is passed behind the stomach and up through the mediastinal tunnel into the neck. An end-to-end anastomosis is performed to the cervical esophagus using a single layer technique. The colon is pulled taut and secured to the left crus with four or five interrupted sutures. Five centimeters below the crus an opening is made in the mesentery adjacent to the colon along its mesenteric border, through which an Endo-GIA stapler is passed and the colon is divided. The proximal end, which is the distal end of the interposed colon, is Anastomosed high on the posterior fundic wall of the stomach, using a triangular stapling anastomotic technique. This is done by stapling longitudinally the stomach and colon together with a 75-mm Endo-GIA stapler, spreading the base of the incision apart, and closing it with a T-55 stapler. Colonic continuity is reestablished by bringing the proximal right colon to the distal staple line in the left colon and performing an end-to-end anastomosis using a double-layer technique.
Although conceptually appealing, preservation of vagal nerve integrity or the gastric reservoir function after vagal sparing esophagectomy only recently has been validated. Banki and associates compared patients undergoing vagal sparing esophagectomy to those with conventional esophagectomy and colon or gastric interposition. This study showed that vagal sparing esophagectomy preserved gastric secretion, gastric emptying, meal capacity, and body mass index, compared to esophageagastrotectomy with colon interposition or standard esophagectomy with gastric pull-up. Vagal sparing esophagectomy patients functioned, for the most part, similarly to normal subjects, allowing them to eat a normal meal, free of dumping or diarrhea. These results indicate that the vagal-sparing esophagectomy procedure does indeed preserve the vagal nerves, and it may be considered in the treatment of benign and early malignant lesions requiring esophagectomy.

**BIBLIOGRAPHY**

Entries highlighted in bright blue are key references.

**General References**


**Surgical Anatomy**


**Physiology**


**Assessment of Esophageal Function**


Gastroesophageal Reflux Disease


Gastroesophageal Reflux Disease


Orlando RC. The pathogenesis of gastroesophageal reflux disease: the relationship between epithelial defense, dysmotility, and acid exposure. Am J Gastroenterol. 1997;92:3S-5S.


Diaphragmatic Hernias


Motility Disorders of the Pharynx and Esophagus


**Carcinoma of the Esophagus**


Frenken M. Best palliation in esophageal cancer; surgery, stenting, radiation, or what? Dis Esophagus. 2001;14:120-123.


**Benign Tumors and Cysts**


**Esophageal Perforation**


Caustic Injury


Techniques of Esophageal Reconstruction


The stomach stores and facilitates the digestion and absorption of ingested food and helps regulate appetite. Treatable diseases of the stomach are common, and it is accessible and relatively forgiving of surgical manipulation. To provide accurate diagnosis and rational treatment, the physician must understand gastric anatomy, physiology, and pathophysiology; this includes a sound understanding of the mechanical, secretory, and endocrine processes by which the stomach accomplishes its important functions and a familiarity with the common benign and malignant gastric disorders. Important historical milestones\(^1\)\(^-\)\(^6\) that influenced the contemporary understanding of gastric disease and surgical therapy are summarized in Table 26-1.

## ANATOMY

### Anatomic Relationships and Gross Morphology

The stomach is the most proximal abdominal organ of the digestive tract (Fig. 26-1).\(^7\) The part of the stomach attached to the esophagus is called the cardia. Just proximal to the cardia at the gastroesophageal (GE) junction is the anatomically indistinct but physiologically demonstrable lower esophageal sphincter. At the distal end, the readily apparent pyloric sphincter connects the stomach to the proximal duodenum. The stomach is relatively fixed at these points, but the majority of the stomach is quite mobile with the shorter lesser curvature on the right and the longer greater curvature on the left.

The superior-most part of the stomach is the distensible fundus, bounded superiorly by the diaphragm and laterally by the spleen. The angle of His is where the fundus meets the left side of the GE junction. Generally, the inferior extent of the fundus is considered to be the horizontal plane of the GE junction, where the body (corpus) of the stomach begins. The body of the stomach contains most of the parietal (oxyntic) cells, some of which are also present in the cardia and fundus. At the angularis incisura, the lesser curvature turns rather abruptly to the right, marking the anatomic beginning of the antrum, which comprises the distal 25% to 30% of the stomach.
The liver, colon, spleen, pancreas, and occasionally the kidney, abut the stomach (Fig. 26-2). The left lateral segment of the liver usually obscures part of the anterior stomach. Inferiorly, the stomach is attached to the transverse colon by the gastrocolic omentum. The lesser curvature is tethered to the liver by the hepatogastric ligament, also referred to as the lesser gastrocolic omentum. The lesser curvature is tethered to the greater curvature by the gastroepiploic arcade along the greater curvature. The right gastroepiploic artery arises from the splenic artery, and, left gastroepiploic artery arises from the gastroepiploic arcade along the greater curvature. The right gastroepiploic arcade along the greater gastric curvature, and the right and left gastroepiploic arteries form an arcade along the greater gastric curvature. The left gastric artery usually arises from the hepatic artery near the pylorus and hepatoduodenal ligament and runs proximally along the distal stomach. In the fundus along the proximal greater curvature, the short gastric arteries and veins arise from the splenic circulation. There also may be additional vascular branches to the proximal stomach from the phrenic and splenic circulation.

The veins draining the stomach generally parallel the arteries. The left gastric (coronary vein) and right gastric veins usually drain into the portal vein, though occasionally the coronary vein drains into the splenic vein. The right gastroepiploic vein drains into the superior mesenteric vein near the inferior border of the pancreatic neck, and the left gastroepiploic vein drains into the splenic vein.

The richness of the gastric blood supply and its many anastomotic connections have important clinical implications. At least two of the four named gastric arteries may be occluded without inducing gastric ischemia. This is done routinely when the stomach is mobilized and pedicled on the right gastroepiploic or ligated without inducing gastric ischemia. This is done routinely when the stomach is mobilized and pedicled on the right gastroepiploic or ligated without inducing gastric ischemia. This is done routinely when the stomach is mobilized and pedicled on the right gastroepiploic or ligated without inducing gastric ischemia. This is done routinely when the stomach is mobilized and pedicled on the right gastroepiploic or ligated without inducing gastric ischemia. This is done routinely when the stomach is mobilized and pedicled on the right gastroepiploic or ligated without inducing gastric ischemia. This is done routinely when the stomach is mobilized and pedicled on the right gastroepiploic or ligated without inducing gastric ischemia. This is done routinely when the stomach is mobilized and pedicled on the right gastroepiploic or ligated without inducing gastric ischemia. This is done routinely when the stomach is mobilized and pedicled on the right gastroepiploic or ligated without inducing gastric ischemia. This is done routinely when the stomach is mobilized and pedicled on the right gastroepiploic or ligated without inducing gastric ischemia. This is done routinely when the stomach is mobilized and pedicled on the right gastroepiploic or ligated without inducing gastric ischemia. This is done routinely when the stomach is mobilized and pedicled on the right gastroepiploic or ligated without inducing gastric ischemia. This is done routinely when the stomach is mobilized and pedicled on the right gastroepiploic or ligated without inducing gastric ischemia. This is done routinely when the stomach is mobilized and pedicled on the right gastroepiploic or ligated without inducing gastric ischemia. This is done routinely when the stomach is mobilized and pedicled on the right gastroepiploic or ligated without inducing gastric ischemia. This is done routinely when the stomach is mobilized and pedicled on the right gastroepiploic or ligated without inducing gastric ischemia.

**Key Points**

1. Whenever testing suggests *Helicobacter pylori* infection, treatment should be initiated and eradication confirmed.
2. Lifelong acid suppression should be considered in any patient admitted to a hospital because of peptic ulcer disease. Acid suppressive medication may be equivalent to surgical vagotomy in preventing recurrent peptic ulcer or ulcer complications.
3. If possible, gastric resection for peptic ulcer is avoided in the asthenic or high-risk patient.
4. Though less common in the United States, gastric cancer is a major cause of cancer-related morbidity and mortality worldwide.
5. Diagnostic laparoscopy with peritoneal lavage should be considered in the evaluation of clinical stage 2 and 3 patients with gastric cancer.
6. Multimodality therapy for gastric cancer, including resection in combination with perioperative chemotherapy or adjuvant chemoradiotherapy is associated with a survival advantage compared to surgery alone.
7. Most patients with primary gastric lymphoma can be treated without gastric resection.
8. Localized gastrointestinal stromal tumors of the stomach are treated with full thickness excision. Adjuvant (or neo-adjuvant) imatinib is indicated for higher-risk lesions.
9. Gastric neuroendocrine tumors may arise in the presence (types 1 and 2) or absence (type 3) of hypergastrinemia. Type 3 gastric neuroendocrine tumors should usually be treated with subtotal gastrectomy and regional lymphadenectomy.
10. Roux-en-Y gastrojejunostomy with a large (>50%) proximal gastric remnant should be avoided because marginal ulceration and/or gastric stasis (Roux syndrome) may become problematic.

**Arterial and Venous Blood Supply**

The stomach is the most richly vascularized portion of the alimentary tube with ample blood flow and a dense intramural vascular anastomotic network. The large majority of the gastric blood supply is from the celiac axis via four named arteries (Fig. 26-3). The left and right gastric arteries form an anastomotic arcade along the lesser gastric curvature, and the right and left gastroepiploic arteries form an arcade along the greater gastric curvature. The left gastric artery is consistently the largest artery to the stomach and usually arises directly from the celiac trunk and divides into an ascending and descending branch along the lesser gastric curvature. Approximately 20% of the time, the left gastric artery supplies an aberrant vessel that travels in the gastrohepatic ligament (lesser omentum) to the left side of the liver. Rarely, this is the only arterial blood supply to this part of the liver (replaced left hepatic artery), and inadvertent ligation may lead to clinically significant hepatic ischemia. The more common smaller accessory left hepatic artery may be ligated without significant consequences.

The second largest artery to the stomach is the right gastroepiploic artery, which consistently arises from the gastroduodenal artery behind the first portion of the duodenum. The left gastroepiploic artery arises from the splenic artery, and, together with the right gastroepiploic artery, forms the rich gastroepiploic arcade along the greater curvature. The right gastric artery usually arises from the hepatic artery near the pylorus and hepatoduodenal ligament and runs proximally along the distal stomach. In the fundus along the proximal greater curvature, the short gastric arteries and veins arise from the splenic circulation. There also may be additional vascular branches to the proximal stomach from the phrenic and splenic circulation.

Multimodality therapy for gastric cancer, including resection in combination with perioperative chemotherapy or adjuvant chemoradiotherapy is associated with a survival advantage compared to surgery alone.

Most patients with primary gastric lymphoma can be treated without gastric resection.

Localized gastrointestinal stromal tumors of the stomach are treated with full thickness excision. Adjuvant (or neo-adjuvant) imatinib is indicated for higher-risk lesions.

Gastric neuroendocrine tumors may arise in the presence (types 1 and 2) or absence (type 3) of hypergastrinemia. Type 3 gastric neuroendocrine tumors should usually be treated with subtotal gastrectomy and regional lymphadenectomy.

Roux-en-Y gastrojejunostomy with a large (>50%) proximal gastric remnant should be avoided because marginal ulceration and/or gastric stasis (Roux syndrome) may become problematic.

**Lymphatic Drainage**

Generally speaking, the gastric lymphatics parallel the blood vessels (Fig. 26-4). The cardia and medial half of the corpus commonly drain to nodes along the left gastric and celiac axis. The lesser curvature side of the antrum usually drains to the right gastric and pyloric nodes, while the greater curvature half of the distal stomach drains to the nodes along the right gastroepiploic chain. The proximal greater curvature side of the stomach usually drains into nodes along the left gastroepiploic or splenic hilum. The nodes along both the greater and lesser...
curvature commonly drain into the celiac nodal basin. There is a rich anastomotic network of lymphatics that drain the stomach, often in a somewhat unpredictable fashion. Thus, a tumor arising in the distal stomach may give rise to positive lymph nodes that are uninvolved. It also helps explain the not infrequent finding of positive lymph nodes, which may be many centimeters away from the primary tumor, with closer nodes that are uninvolved.

Not surprisingly, extensive and meticulous lymphadenectomy is considered by many surgeons to be an important part of an operation for gastric cancer. Surgeons and pathologists have numbered the primary and secondary lymph node groups to which the stomach drains (see Fig. 26-4).11,12

### Historic milestones in gastric surgery

<table>
<thead>
<tr>
<th>DATE</th>
<th>EVENT</th>
<th>DATE</th>
<th>EVENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>350 b.c.–201 A.D.</td>
<td>Existence of gastric ulceration was acknowledged by Diocles of Carystos (350 b.c.), Celsus, and Galen (131–201 A.D.).</td>
<td>1886</td>
<td>Heineke performs pyloroplasty.</td>
</tr>
<tr>
<td>1363</td>
<td>Guy de Chauliac describes closure of gastric wound.</td>
<td>1888</td>
<td>Mikulicz performs similar operation.</td>
</tr>
<tr>
<td>1586</td>
<td>Marcellus Donatus of Mantua describes gastric ulcer at autopsy.</td>
<td>1892</td>
<td>Jaboulay describes bypassing the intact pylorus with gastroduodenostomy.</td>
</tr>
<tr>
<td>1600–1700</td>
<td>Reports of surgeons cutting stomach to remove foreign bodies.</td>
<td>1902</td>
<td>Finney from Baltimore describes pyloroplasty technique.</td>
</tr>
<tr>
<td>1688</td>
<td>Muralto describes duodenal ulcer at autopsy.</td>
<td>1891–1913</td>
<td>Different techniques of gastrectomy are described by Witzel (1891), Stamm (1894), and Janeway (1913).</td>
</tr>
<tr>
<td>1737</td>
<td>Morgagni describes both gastric and duodenal ulcer at autopsy.</td>
<td>1920–1950</td>
<td>Subtotal gastrectomy grows popular as an operation for peptic ulcer. Von Haberer and Finsterer proponents.</td>
</tr>
<tr>
<td>1833</td>
<td>William Beaumont reports data recorded during his care of Alexis St. Martin who developed a gastric fistula from a left upper quadrant musket wound.</td>
<td>1943</td>
<td>Dragstedt and Owen describe transthoracic truncal vagotomy to treat peptic ulcer disease. By the early 1950s, it is well recognized that some patients developed gastric stasis after this procedure, and transabdominal truncal vagotomy and drainage (pyloroplasty or gastrojejunostomy) become a standard ulcer operation.</td>
</tr>
<tr>
<td>1869</td>
<td>Maury reportedly performs feeding gastrostomy to palliate esophageal stricture following consultation with Samuel D. Gross.</td>
<td>1952</td>
<td>Farmer and Smithwick describe good results with truncal vagotomy and hemigastrectomy for peptic ulcer.</td>
</tr>
<tr>
<td>1875</td>
<td>Sidney Jones in London publishes the first successful gastrostomy for feeding.</td>
<td>1953</td>
<td>Edwards and Herrington (Nashville) describe truncal vagotomy and antrectomy for peptic ulcer.</td>
</tr>
<tr>
<td>1879</td>
<td>Paen performed distal gastrectomy and gastroduodenostomy. The patient died 5 d later.</td>
<td>1955</td>
<td>Zollinger and Ellison describe the eponymous syndrome.</td>
</tr>
<tr>
<td>1880</td>
<td>Rydygier resected a distal gastric cancer, and the patient died 12 h later.</td>
<td>1957</td>
<td>Griffith and Harkins (Seattle) describe parietal cell vagotomy (highly selective vagotomy) for the elective treatment of peptic ulcer disease.</td>
</tr>
<tr>
<td>1884</td>
<td>Rydygier reports an unsuccessful gastrojejunostomy for benign gastric outlet obstruction.</td>
<td>1995–current</td>
<td>Dramatic increase in bariatric operations.</td>
</tr>
<tr>
<td>1885</td>
<td>Billroth performs a successful distal gastrectomy and gastrojejunostomy (Billroth II) for gastric cancer.</td>
<td>2000–current</td>
<td>Development of natural orifice translumenal endoscopic surgery, such as transgastric appendectomy and peroral pyloromyotomy. Development of robotic gastrectomy.</td>
</tr>
</tbody>
</table>

### Innervation

The vagus nerves provide the extrinsic parasympathetic innerva-

tion to the stomach, and acetylcholine is the most important neu-

rotransmitter. From the vagal nucleus in the floor of the fourth cerebral ventricle, the vagus traverses the neck in the carotid sheath and enters the mediastinum, where it gives off the recurrent laryngeal nerve and divides into several branches around the esophagus. These branches come together again above the esophageal hiatus and form the left (anterior) and right
(posterior) vagal trunks (mnemonic LARP). Near the GE junction the anterior vagus sends a branch (or branches) to the liver in the gastrohepatic ligament, and continues along the lesser curvature as the anterior nerve of Latarjet (Fig. 26-5). Similarly, the posterior vagus sends branches to the celiac plexus and continues along the posterior lesser curvature. The nerves of Latarjet send segmental branches to the body of the stomach before they terminate near the angularis incisura as the “crow’s foot,” sending branches to the antropyloric region. There may be additional branches to the distal stomach and pylorus that travel near the right gastric and/or gastroepiploic arteries. In 50% of patients, there are more than two vagal nerves at the esophageal hiatus. The branch that the posterior vagus sends to the posterior fundus is termed the criminal nerve of Grassi. This branch typically arises above the esophageal hiatus and is easily missed during truncal or highly selective vagotomy (HSV). Vagal fibers originating in the brain synapse with neurons in Auerbach’s myenteric plexus and Meissner’s submucosal plexus. In the stomach, the vagus nerves affect secretion (including acid), motor function, and mucosal bloodflow and cytoprotection. They also play a role in appetite control and perhaps even mucosal immunity and inflammation.\textsuperscript{14,15} Most of the axons contained in the vagal trunks are afferent (i.e., carrying stimuli from the viscera to the brain).

The extrinsic sympathetic nerve supply to the stomach originates at spinal levels T5 through T10 and travels in the splanchnic nerves to the celiac ganglion. Postganglionic sympathetic nerves then travel from the celiac ganglion to the stomach along the blood vessels.

Neurons in the myenteric and submucosal plexuses constitute the intrinsic nervous system of the stomach. There may be more intrinsic gastric neurons than extrinsic neurons, but their function is poorly understood.

The characterization of the vagus as the cholinergic system and the sympathetic system as the adrenergic system of innervation is a misleading oversimplification. Although acetylcholine is an important neurotransmitter mediating vagal function, and epinephrine is important in the sympathetic nerves, both systems (as well as the intrinsic neurons) have various and diverse neurotransmitters, including cholinergic, adrenergic, and peptidergic (e.g., substance P and somatostatin).

**Histology**

There are four distinct layers of the gastric wall: mucosa, submucosa, muscularis propria, and serosa (Fig. 26-6). The inner layer of the stomach is the mucosa, which is lined with columnar epithelial cells of various types. Beneath the basement membrane of the epithelial cells is the lamina propria, which contains connective tissue, blood vessels, nerve fibers, and inflammatory cells. Beneath the lamina propria is a thin muscle layer called the muscularis mucosa, the deep boundary of the mucosal layer of the gut. The epithelium, lamina propria, and muscularis mucosa constitute the mucosa (Fig. 26-7).\textsuperscript{16} The epithelium of the gastric mucosa is columnar glandular. Scanning electron micrographs show a smooth mucosal carpet punctuated by the openings of the gastric glands or units. The gastric glands are lined with different types of epithelial cells, depending upon their location in the stomach (Fig. 26-8 and Table 26-2).\textsuperscript{17,18} There are also endocrine cells present in the gastric glands. Progenitor or stem cells in the isthmus and base of the glands differentiate and...
replenish sloughed cells on a regular basis. Genetic studies show that there are several different subpopulations of stem cells in the gastric glands and that during conditions of stress even chief cells exhibit the plasticity required to regenerate other types of gastric epithelial cells.19 Throughout the stomach, the luminal carpet consists primarily of mucus-secreting surface epithelial cells (SECs) that extend down into the gland pits for variable distances. These cells also secrete bicarbonate and play an important role in protecting the stomach from injury due to acid, pepsin, and/or ingested irritants. In fact, all epithelial cells of the stomach (except the endocrine cells) contain carbonic anhydrase and are capable of producing bicarbonate.

In the cardia, the gastric glands are branched and secrete primarily mucus and bicarbonate, and little acid. In the fundus and body, the glands are more tubular, and the pits are deep. Parietal and chief cells are common in these glands (Fig. 26-9). Histamine-secreting enterochromaffin-like (ECL) cells and somatostatin-secreting D cells are also found. Parietal cells...
Figure 26-5. Vagal innervation of the stomach. br. = branch; n. = nerve; rt. = right. (Reproduced with permission from Menguy R: Surgery of Peptic Ulcer. Philadelphia, PA: Elsevier/Saunders; 1976.)

Figure 26-6. Layers of the gastric wall. (Reproduced with permission from Fawcett DW: Bloom and Fawcett’s Textbook of Histology, 11th ed. Philadelphia, PA: Elsevier/Saunders; 1986.)

Figure 26-7. Gastric mucosa. (Used with permission from Emma Furth, MD.)
secrete acid and intrinsic factor into the gastric lumen, and bicarbonate into the intercellular space. They have a characteristic ultrastructural appearance with secretory canaliculi (deep invaginations of the surface membrane) and cytoplasmic tubulovesicles containing the acid-producing apparatus H⁺/K⁺-ATPase (proton pump) (see Fig. 26-9). There are numerous mitochondria; in fact, the parietal cell is the most mitochondria-rich cell in the body. When the parietal cell is stimulated, the cytoplasmic tubulovesicles fuse with the membrane of the secretory canaliculus; when acid production ceases, the process is reversed. Arguably, parietal cells produce the only truly essential substance made by the stomach (i.e., intrinsic factor). Parietal cells tend to occupy the midportion of the gastric glands found in the corpus of the stomach.

Chief cells (also called zymogenic cells) secrete pepsinogen I, which is maximally activated at a pH of 2.5. They tend to be clustered toward the base of the gastric glands and have a low columnar shape. Ultrastructurally, chief cells have the characteristics of protein-synthesizing cells: basal granular endoplasmic reticulum, supranuclear Golgi apparatus, and apical zymogen granules (Fig. 26-10). When stimulated, the chief cells produce two immunologically distinct proenzyme forms of pepsinogen: predominantly pepsinogen I and some pepsinogen II, most of which is produced by SECs. These proenzymes are activated in an acidic luminal environment.

In the antrum, the gastric glands are again more branched and shallow, parietal cells are rare, and gastrin-secreting G cells and somatostatin-secreting D cells are present. A variety of hormone-secreting cells are present in various proportions throughout the gastric mucosa (Fig. 26-11). Histologic analysis suggests that in the normal stomach, 13% of the epithelial cells are oxyntic (parietal) cells, 44% are chief (zymogenic) cells, 40% are mucous cells, and 3% are endocrine cells. In general, the antrum produces gastrin but not acid, and the proximal stomach produces acid but not gastrin. The border between the corpus and antrum migrates proximally with age (especially on the lesser curvature side of the stomach).

Deep to the muscularis mucosa is the submucosa, which is rich in branching blood vessels, lymphatics, collagen, various inflammatory cells, and nerve fibers and ganglion cells of Meissner’s autonomic submucosal plexus. The collagen-rich submucosa gives strength to GI anastomoses. The mucosa and submucosa are folded into the grossly visible gastric rugae, which tend to flatten out as the stomach becomes distended.

Below the submucosa is the thick muscularis propria (also referred to as the muscularis externa), which consists of an incomplete inner oblique layer, a complete middle circular layer (continuous with the esophageal circular muscle and the circular muscle of the pylorus), and a complete outer longitudinal layer (continuous with the longitudinal layer of the esophagus and duodenum). Within the muscularis propria is the rich network of autonomic ganglia and nerves that make up Auerbach’s myenteric plexus. Specialized pacemaker cells, the interstitial cells of Cajal (ICC), also are present.

The outer layer of the stomach is the serosa, also known as the visceral peritoneum. This layer provides significant tensile strength to gastric anastomoses. When tumors originating in the mucosa penetrate and breach the serosa, microscopic or gross peritoneal metastases are common, presumably from shedding of tumor cells that would not have occurred if the serosa had not been penetrated. In this way, the serosa may be thought of as an outer envelope of the stomach.

**PHYSIOLOGY**

The stomach stores food and facilitates digestion through a variety of secretory and motor functions. Important secretory functions include the production of acid, pepsin, intrinsic factor, mucus, and a variety of GI hormones. Important motor functions include food storage (receptive relaxation and accommodation), grinding and mixing, controlled emptying of ingested food, and periodic interprandial “housekeeping.”

**Acid Secretion**

Hydrochloric acid in the stomach hastens both the physical and (with pepsin) the biochemical breakdown of ingested food. In an acidic environment, pepsin and acid facilitate proteolysis. Gastric acid also inhibits the proliferation of ingested pathogens, which protects against both infectious gastroenteritis and intestinal bacterial overgrowth and helps to maintain a healthy gastrointestinal microbiome. Long-term acid suppression with proton pump inhibitors (PPIs) has been associated with an
increased risk of community-acquired *Clostridium difficile* colitis and other gastroenteritis, presumably because of the absence of this protective germicidal barrier.22,23

**Parietal Cell.** The parietal cell is stimulated to secrete acid (Fig. 26-12) when one or more of three membrane receptor types is stimulated by acetylcholine (from vagally stimulated enteric neurons), gastrin (from G cells), or histamine (from ECL cells).7,24,25 The enzyme \( \text{H}^+/\text{K}^-\)-ATPase is the parietal cell proton pump. It is stored within the intracellular tubulovesicles and is the final common pathway for gastric acid secretion. When the parietal cell is stimulated, there is a cytoskeletal rearrangement and fusion of the tubulovesicles with the apical membrane of the secretory canaliculus. The heterodimer assembly of the enzyme subunits into the microvilli of the secretory canaliculus results in acid secretion, with extracellular potassium being exchanged for cytosolic hydrogen. Although electroneutral, this is an energy-requiring process because the hydrogen is secreted against a gradient of at least 1 million-fold, which explains why the parietal cell is packed with energy producing mitochondria. During acid production, potassium and chloride are also secreted into the apical secretory canaliculus through separate channels, providing potassium to exchange for \( \text{H}^+ \) via the \( \text{H}^+/\text{K}^-\)-ATPase, and chloride to accompany the secreted hydrogen. At the basolateral membrane, the combined activity of various

---

**Table 26-2**

<table>
<thead>
<tr>
<th>CELL TYPE</th>
<th>DISTINCTIVE ULTRASTRUCTURAL FEATURES</th>
<th>MAJOR FUNCTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface-foveolar mucous cells</td>
<td>Apical stippled granules up to 1 μm in diameter</td>
<td>Production of neutral glycoprotein and bicarbonate to form a gel on the gastric luminal surface; neutralization of hydrochloric acid</td>
</tr>
<tr>
<td>Mucous neck cell</td>
<td>Heterogeneous granules 1–2 μm in diameter dispersed throughout the cytoplasm</td>
<td>Progenitor cell for all other gastric epithelial cells; glycoprotein production; production of pepsinogens I and II</td>
</tr>
<tr>
<td>Oxyntic (parietal) cell</td>
<td>Surface membrane invaginations (canaliculi); tubulovesicle structures; numerous mitochondria</td>
<td>Production of hydrochloric acid; production of intrinsic factor; production of bicarbonate</td>
</tr>
<tr>
<td>Chief cell</td>
<td>Moderately dense apical granules up to 2 μm in diameter; prominent supranuclear Golgi apparatus; extensive basolateral granular endoplasmic reticulum</td>
<td>Production of pepsinogens I and II, and of lipase</td>
</tr>
<tr>
<td>Cardiopyloric mucous cell</td>
<td>Mixture of granules like those in mucous neck and chief cells; extensive basolateral granular endoplasmic reticulum</td>
<td>Production of glycoprotein; production of pepsinogen II</td>
</tr>
<tr>
<td>Endocrine cells</td>
<td>See Figure 26-11</td>
<td></td>
</tr>
</tbody>
</table>

*Bicarbonate is probably produced by other gastric epithelial cells in addition to surface-foveolar mucous cells.*


---

**Figure 26-9.** Ultrastructural features of the parietal (oxyntic) cell. SC = secretory canaliculus; M = mitochondria; TV = tubulovesicle. *(Reproduced with permission from Ming S-C, Goldman H: Pathology of the Gastrointestinal Tract, 2nd ed. Baltimore, MD: Williams & Wilkins; 1998.)*

**Figure 26-10.** Ultrastructural features of the chief (zymogenic) cell. GA = Golgi apparatus; GER = granular endoplasmic reticulum; ZG = zymogen granule. *(Reproduced with permission from Ming S-C, Goldman H: Pathology of the Gastrointestinal Tract, 2nd ed. Baltimore, MD: Williams & Wilkins; 1998.)*
cotransporters and ion exchangers accomplishes intracellular pH regulation and electrolyte homeostasis.24

The normal human stomach contains approximately 1 billion parietal cells, and total gastric acid production is proportional to parietal cell mass. Almost all of the parietal cells are in the proximal 2/3 stomach, though there are some parietal cells found in gastric antral glands. The potent acid-suppressing PPI drugs irreversibly interfere with the function of the H+/K+-ATPase molecule. These agents must be incorporated into the activated enzyme to be effective and thus work best when taken before or during a meal (when the parietal cell is stimulated). When PPI therapy is stopped, acid secretory capability gradually returns (within days) as new H+/K+-ATPase is synthesized.

Gastrin, acetylcholine, and histamine stimulate the parietal cell to secrete hydrochloric acid (see Fig. 26-12). Gastrin binds to type B cholecystokinin (CCK2) receptors on ECL cells and stimulates ECL cell histamine release, which binds to H2 receptors on the parietal cell. This stimulates adenylatecyclase (via a G-protein–linked mechanism) and increases cAMP which activates protein kinases, leading to increased levels of phosphoproteins and activation of the proton pump. Gastrin also binds to CCK2 receptors on the parietal cell, but this is less important for acid secretion than the gastrin effect on ECL cells. Acetylcholine from intrinsic neurons binds to M3 muscarinic receptors on the parietal cell, which (like gastrin binding to CCK2 receptors) stimulates phospholipase C via a G-protein–linked mechanism leading to increased production of inositol trisphosphate from membrane bound phospholipids. Inositol trisphosphate stimulates the release of calcium from intracellular stores, which leads to activation of protein kinases and activation of H+/K+-ATPase. Somatostatin released from mucosal D cells in the antral and oxyntic mucosa in response to luminal acid binds to SSTR2 receptors on parietal cells and inhibits acid release directly. Somatostatin also inhibits acid secretion in a paracrine fashion, binding to nearby ECL cells in the oxyntic mucosa and decreasing histamine release, and binding to nearby antral G cells to inhibit gastrin release.26

Physiologic Acid Secretion.27 Food ingestion is the physiologic stimulus for acid secretion (Fig. 26-13). The acid secretory response that occurs after a meal is traditionally described

---

**Figure 26-11.** Endocrine cells of the stomach—proportion by site. D = d cell (somatostatin); EC = enterochromaffin cell; ECL = enterochromaffin-like cell (histamine); G = g cell (gastrin). (Reproduced with permission from Feldman M, Friedman LS, Sleisenger MH, et al: Sleisenger and Fordtran’s Gastrointestinal and Liver Disease, 7th ed. Philadelphia, PA: Elsevier/Saunders; 2002.)

**Figure 26-12.** Control of acid secretion in the parietal cell. ATP = adenosine triphosphate; cAMP = cyclic adenosine monophosphate; CCK = cholecystokinin; H2 = histamine 2; IP3 = inositol trisphosphate; PIP2 = phosphatidylinositol 4,5-bisphosphate; PLC = phospholipase C. (Reproduced with permission from Zuidema GD, Yeo CJ: Shackelford’s Surgery of the Alimentary Tract, 5th ed. Vol. II. Philadelphia, PA: Elsevier/Saunders; 2002.)
in three phases: cephalic, gastric, and intestinal.\textsuperscript{28,29} The cephalic or vagal phase begins with the thought, sight, smell, and/or taste of food. These stimuli activate several cortical and hypothalamic sites (e.g., tractus solitarius, dorsal motor nucleus, and dorsal vagal complex), and signals are transmitted to the stomach by the vagal nerves which stimulate enteric submucosal neurons. Acetylcholine is released, leading to stimulation acid secretion from parietal cells. Vagal stimulation also leads to gastrin release from antral G cells via CGRP, and sensitizes ECL cells to gastrin.\textsuperscript{30,31} Although the acid secreted per unit of time in the cephalic phase is greater than in the other two phases, the cephalic phase is shorter. Thus, the cephalic phase accounts for no more than 30\% of total acid secretion in response to a meal. Sham feeding (chewing and spitting) stimulates gastric acid secretion only via the cephalic phase, and it results in acid secretion that is about half of that seen in response to IV pentagastrin or histamine.

When food reaches the stomach, the gastric phase of acid secretion begins. This phase lasts until the stomach is empty and accounts for about 60\% of the total acid secretion in response to a meal. The gastric phase of acid secretion has several components. Amino acids and small peptides directly stimulate antral G cells to secrete gastrin, which is carried in the bloodstream to the ECL and parietal cells, stimulating acid secretion in an endocrine fashion. In addition, proximal gastric distention stimulates acid secretion via a vagovagal reflex arc, which is mitigated by truncal or highly selective vagotomy (HSV). Antral distention also stimulates antral gastrin secretion. Finally, ongoing cephalic vagal input stimulates gastrin release, which in turn stimulates histamine release from ECL cells and acid secretion.

The intestinal phase of gastric secretion is poorly understood. It is thought to be mediated by a hormone released from the proximal small bowel mucosa in response to luminal chyme. This phase starts when gastric emptying of ingested food begins, and it continues as long as nutrients remain in the proximal small intestine. It accounts for about 10\% of meal-induced acid secretion.

Interprandial basal acid secretion is 2 to 5 mEq hydrochloric acid per hour, about 10\% of maximal acid output (MAO), and it is greater at night. Basal acid secretion probably contributes to the relatively low bacterial counts found in the stomach. Basal acid secretion is reduced 75\% to 90\% by vagotomy or continuous H\(_2\)-receptor blockade.

The pivotal role that ECL cells play in the regulation of gastric acid secretion is emphasized in Fig. 26-13. The acid stimulatory effect of gastrin is largely mediated by histamine released from mucosal ECL cells. H\(_2\)-receptor knockout mice do not secrete acid in response to gastrin.\textsuperscript{24} This explains why the H\(_2\)-receptor antagonists (H\(_2\)RAs) are effective inhibitors of acid secretion, even though histamine is only one of three parietal cell stimulants. The mucosal D cell, which releases somatostatin, is also an important regulator of acid secretion. Somatostatin inhibits histamine release from ECL cells and gastrin release from antral G cells. The function of D cells can be inhibited by \textit{Helicobacter pylori} infection, resulting in an exaggerated acid secretory response (see “\textit{Helicobacter pylori} Infection”).

Proton pump inhibitors are potent suppressors of gastric acid secretion. This results in hypergastrinemia and consequent ECL stimulation. In patients on long-term PPI (median 5.5 years), the degree of hypergastrinemia does not appear to correlate with the length of treatment.\textsuperscript{32} Chronic PPI use has been associated with ECL hyperplasia and type 1 gastric neuroendocrine tumor, but so far there has been no evidence linking these agents to malignant gastric epithelial or neuroendocrine tumors. Gastrin levels return to normal within a few days of PPI cessation, but during this time, some patients may experience gastric hyperacidity and dyspeptic symptoms, which may lead to difficulty in getting patients off the medication.\textsuperscript{33,34} This is less likely to occur with short-term PPI use and may be ameliorated by PPI dose tapering and/or initiation of H\(_2\) blockers prior to PPI cessation.

**Pepsinogen Secretion**

The most potent physiologic stimulus for pepsinogen secretion from chief cells is food ingestion; acetylcholine is the most important mediator. Somatostatin inhibits pepsinogen secretion. Pepsinogen I is produced by chief cells in acid producing glands, whereas pepsinogen II is produced by chief cells and by SECs in both acid producing and gastrin producing (i.e., antral) glands. Pepsinogen is cleaved to the active pepsin enzyme in an acidic environment and is maximally active at pH 2.5, and inactive at pH >5, although pepsinogen II may be activated over a wider pH range than pepsinogen I. Pepsin catalyzes the hydrolysis of proteins and is denatured at alkaline pH. Serum levels of pepsinogen I and II are increased in helicobacter gastritis, so elevated pepsinogen I and II levels and positive helicobacter serology are presumptive evidence of active helicobacter infection. Longstanding helicobacter infection may lead to atrophic gastritis, suggested by decreased pepsinogen I/II ratio (from chief cell loss) and hypergastrinemia (from parietal cell loss and hypochlorhydria).\textsuperscript{35}

**Intrinsic Factor**

Activated parietal cells secrete intrinsic factor in addition to hydrochloric acid. Presumably the stimulants are similar, but
acid secretion and intrinsic factor secretion may not be linked. Intrinsic factor binds to luminal vitamin B₁₂, and the complex is absorbed in the terminal ileum via mucosal receptors. Vitamin B₁₂ deficiency can be life threatening, and patients with total gastrectomy or pernicious anemia (i.e., patients with no parietal cells) require B₁₂ supplementation by a nonenteric route. Some patients develop vitamin B₁₂ deficiency following gastric bypass, presumably because there is insufficient intrinsic factor present in the small proximal gastric pouch and oral B₁₂ intake may be decreased. Under normal conditions, a significant excess of intrinsic factor is secreted, and acid-suppressive medication does not appear to inhibit intrinsic factor production and release.

**Gastric Mucosal Barrier**

The stomach’s durable resistance to autodigestion by caustic hydrochloric acid and active pepsin is intriguing. Some of the important elements of gastric barrier function and cytoprotection are listed in Table 26-3. When these defenses break down, ulceration occurs. A variety of factors are important in maintaining an intact gastric mucosal layer. The mucus and bicarbonate secreted by SECs form an unstirred mucous gel with a favorable pH gradient. Cell membranes and tight junctions prevent hydrogen ions from gaining access to the interstitial space. Hydrogen ions that do break through are buffered by the alkaline tide created by basolateral bicarbonate secretion from stimulated parietal cells. Any sloughed or denuded SECs are rapidly replaced by migration of adjacent cells, a process known as restitution. Mucosal blood flow plays a crucial role in maintaining a healthy mucosa, providing nutrients and oxygen for the cellular functions involved in cytoprotection. During acid secretion, there is a tremendous gradient favoring the movement of hydrogen ions from the lumen to the interstitium. This “back-diffused” hydrogen is buffered and rapidly removed by the rich blood supply. When “barrier breakers” such as bile or aspirin lead to increased back-diffusion of hydrogen ions from the lumen into the lamina propria and submucosa, there is a protective increase in mucosal blood flow. If this protective response is blocked, gross ulceration can occur. Important mediators of these protective mechanisms include prostaglandins, nitric oxide, intrinsic nerves, and peptides (e.g., calcitonin gene-related peptide, gastrin-releasing peptide [GRP], gastrin, and heat shock proteins). Sucralfate acts locally to enhance mucosal defenses. Protective reflexes involveafferent sensory neurons, and they can be blocked by the application of topical anesthetics to the gastric mucosa, or the experimental destruction of the afferent sensory nerves. In addition to these local defenses, there are important protective factors in saliva, duodenal secretions, and pancreatic or biliary secretions.

**Gastric Hormones**

The stomach is quite an elegant endocrine organ. It is the source of important peptides which work in an autocrine (EGF and surface epithelial cells, TGF and parietal cells), paracrine (somatostatin), endocrine (gastrin), and/or neurocrine (ghrelin) fashion.

**Gastrin.** Gastrin is produced by antral G cells and is the major hormonal stimulant of acid secretion during the gastric phase predominantly via an endocrine effect on histamine generating ECL cells and to a lesser extent via a direct effect on parietal cells. A variety of molecular forms exist: big gastrin (34 amino acids; G₁₇), little gastrin (17 amino acids; G₁₇), and mini-gastrin (14 amino acids; G₁₇). The large majority of gastrin released by the human antrum is G₁₇. The biologically active pentapeptide sequence at the C-terminal end of gastrin is identical to that of CCK. Luminal peptides and amino acids are the most potent stimulants of gastrin release, and luminal acid is the most potent inhibitor of gastrin secretion. The latter effect is predominantly mediated in a paracrine fashion by somatostatin released from antral D cells. Gastrin-stimulated acid secretion is significantly blocked by H₂ antagonists, suggesting that the principal mediator of gastrin-stimulated acid production is histamine from mucus-producing ECL cells and not direct stimulation of parietal cells by gastrin (see Fig. 26-13). In fact, chronic hypergastrinemia is associated with hyperplasia of gastric ECL cells and, rarely, gastric type I gastric neuroendocrine tumors (type I gastric carcinoid). Gastrin is toxic to gastric parietal cells and to other GI mucosal cells including gastric stem cells. It also is a regulator of gastric cellular proliferation, migration, invasion, apoptosis and angiogenesis. Mucosal biopsies of the gastric body from patients with gastrinoma show a thick mucosa with excess parietal cells, while similar biopsies in patients years after antrectomy (i.e., low gastrin state) show thin mucosa and decreased parietal cells. In animal studies, gastrin administration has been shown to stimulate the growth of established colon cancers and to cause pancreatic acinar cell hyperplasia. Important causes of hypergastrinemia include pernicious anemia, acid-suppressive medication, gastrinoma, retained antrum following distal gastrectomy and Billroth II surgery, and vagotomy.

**Ghrelin.** Ghrelin, first described in 1999, is a small peptide that is produced mainly in the stomach. It is produced by specialized P/D1 endocrine cells in gastric oxyntic glands. Ninety percent of the body’s ghrelin stores are in the stomach and duodenum. Ghrelin is a potent secretagogue of pituitary growth hormone and a weak secretagogue for ACTH and prolactin. It appears to be a major orexigenic regulator of appetite. Ghrelin crosses the blood brain barrier and stimulates appetite via hypothalamic receptors. It also stimulates appetite peripherally by stimulating vagal afferent fibers in the gastric wall.

---

**Table 26-3**

**Important components and mediators of mucosal defenses in the stomach**

<table>
<thead>
<tr>
<th>Components</th>
<th>Mediators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mucous barrier</td>
<td>Prostaglandins</td>
</tr>
<tr>
<td>Bicarbonate secretion</td>
<td>Nitric oxide</td>
</tr>
<tr>
<td>Epithelial barrier</td>
<td>Epidermal growth factor</td>
</tr>
<tr>
<td>Hydrophobic phospholipids</td>
<td>Calcitonin gene-related peptide</td>
</tr>
<tr>
<td>Tight junctions</td>
<td>Hepatocyte growth factor</td>
</tr>
<tr>
<td>Restitution</td>
<td>Histamine</td>
</tr>
<tr>
<td>Microcirculation (reactive hyperemia)</td>
<td>Gastrin-releasing peptide</td>
</tr>
<tr>
<td>Afferent sensory neurons</td>
<td></td>
</tr>
</tbody>
</table>
When ghrelin is elevated, appetite is stimulated, and when it is suppressed, appetite is decreased. Typically, ghrelin levels are elevated before a meal and decreased postprandially. Levels are high during starvation and decreased during hyperglycemia. Obesity and insulin resistance is associated with low ghrelin levels, but resection of the primary source of this hormone (i.e., the stomach) may partly account for the anorexia and weight loss seen in some patients following gastric resection including sleeve gastrectomy (Fig. 26-14).46-48 The effect of RYGBP on ghrelin physiology is controversial. This very effective weight loss procedure has been shown by some investigators to be associated with suppression of plasma ghrelin levels (and appetite) in humans (Fig. 26-15A).49,50

Other groups have failed to show a significant decrease in ghrelin levels following gastric bypass but have found such decreases following sleeve gastrectomy, another effective weight loss operation (Fig. 26-15B).50 Possibly, subtle differences in operative technique, patient selection, or experimental (including assay) conditions account for the disparate results of studies on the effect of bariatric surgery on ghrelin levels in obese patients. Interestingly the two common metabolites of ghrelin have different physiologic effects: acyl-ghrelin increases gastric emptying and appetite while deacyl ghrelin decreases gastric emptying and induces satiety.51 Obviously appetite control is complex with redundant and overlapping orexigenic and anorexigenic pathways and signals.50,52

**Somatostatin.** Somatostatin is produced by D cells located throughout the gastric mucosa. The predominant form in humans is somatostatin 14, though somatostatin 28 is present as well. The major stimulus for somatostatin release is antral acidification; acetylcholine from vagal nerve fibers inhibits its release. Somatostatin inhibits acid secretion from parietal cells and gastrin release from G cells. It also decreases histamine release from ECL cells. The proximity of the D cells to these target cells suggests that the primary effect of somatostatin is mediated in a paracrine fashion, but an endocrine (i.e., bloodstream) effect also is possible.

**Gastrin-Releasing Peptide.** GRP is the mammalian equivalent of bombesin, a hormone discovered more than two decades ago in an extract of skin from a frog. In the antrum, GRP stimulates both gastrin and somatostatin release by binding to receptors on the G and D cells. There are nerve terminals ending near the mucosa in the gastric body and antrum, which are rich in GRP immunoreactivity. When GRP is given peripherally, it stimulates acid secretion, but when it is given centrally into the cerebral ventricles of animals, it inhibits acid secretion, apparently via a pathway involving the sympathetic nervous system.

**Leptin.** Leptin is a protein primarily synthesized in adipocytes. It is also made by chief cells in the stomach, the main source of leptin in the GI tract.39,53 Leptin works at least in part via vagally mediated pathways to decrease food intake in animals. Not surprisingly, leptin, a satiety signal hormone, and ghrelin, a hunger signal hormone, are both synthesized in the stomach, an organ increasingly recognized as central to the mechanisms of appetite control.50,52

**Autocrine Proteins.** Gastric surface epithelial cells secrete a variety of proteins that are important regulators of SEC health, including trefoil factor family proteins and heat shock proteins.38 Parietal cells may also be influenced by molecules they secrete including transforming growth factor-α.

**Gastric Motility and Emptying**

Gastric motor function has several purposes.43,54-56 Interprandial motor activity clears the stomach of undigested debris, sloughed cells, and mucus. When feeding begins, the stomach relaxes to accommodate the meal. Regulated motor activity then breaks down the food into small particles and controls the output into the duodenum. The stomach accomplishes these functions by coordinated smooth muscle relaxation and contraction of the various gastric segments (proximal, distal, and pyloric). Smooth muscle myoelectric potentials are translated into muscular activity, which is modulated by extrinsic and intrinsic innervation and hormones. The mechanisms by which gastric distention is translated into a neurohormonal satiety signal have only been partially elucidated.39,52

**Intrinsic Gastric Innervation.** The extrinsic parasympathetic and sympathetic gastric innervation was discussed previously in “Innervation.” The intrinsic innervation consists of ganglia and nerves that constitute the enteric nervous system (Fig. 26-16).57 There are a variety of neurotransmitters that effect gastric smooth muscle; these are generally grouped as excitatory (augment muscular activity) and inhibitory (decrease muscular activity). Important excitatory neurotransmitters include acetylcholine, the tachykinins, substance P, and neurokinin A. Important inhibitory neurotransmitters include nitric oxide (NO) and vasoactive intestinal peptide (VIP). Serotonin has been shown to modulate both contraction and relaxation. A variety of other molecules affect motility, including GRP, histamine, neuropeptide Y, norepinephrine, and endogenous opioids.

Specialized cells in the muscularis propria also are important modulators of GI motility. These cells, called interstitial cells of Cajal, are distinguishable histologically from neurons and myocytes and appear to amplify both cholinergic excitatory
and nitric inhibitory input to the smooth muscle of the stomach and intestine. They are thought to be the cell of origin for gastrointestinal stromal tumors (GISTs), which are the most common mesenchymal neoplasms in the GI tract.

**Segmental Gastric Motility.** In general, the proximal stomach serves a short-term food storage function and helps regulate basal intragastric tone, and the distal stomach mixes and grinds the food. The pylorus helps the latter process when

---

closed, facilitating retropulsion of the solid food bolus back into the body of the stomach for additional breakdown. The pylorus opens intermittently to allow metered emptying of liquids and small solid particles into the duodenum.

Most of the motor activity of the proximal stomach consists of slow tonic contractions and relaxations, lasting up to 5 minutes. This activity is the main determinant of basal intragastric pressure, an important determinant of liquid emptying. Rapid phasic contractions may be superimposed on the slower tonic motor activity. When food is ingested, intragastric pressure falls as the proximal stomach relaxes. This proximal relaxation is mediated by two important vagovagal reflexes: receptive relaxation and gastric accommodation. Receptive relaxation refers to the reduction in gastric tone associated with the act of swallowing. This occurs before the food reaches the stomach and can be reproduced by mechanical stimulation of the pharynx or esophagus. Gastric accommodation refers to the proximal gastric relaxation associated with distention of the stomach. Accommodation is mediated through stretch receptors in the gastric wall and does not require esophageal or pharyngeal stimulation. Initially, as the meal enters the stomach, there is a drop in intragastric pressure mediated by nitric oxide. As the meal progresses, the intragastric pressure rises, parallel with the onset of satiety. Interestingly, satiety does not seem to be associated with any specific level of intragastric pressure. Obese patients have a delayed onset of satiety, so an obvious hypothesis to be tested is that a pharmacologic-induced increase in gastric tone leads to increased satiety and decreased food intake in this patient group. Because receptive relaxation and accommodation are mediated by afferent and efferent vagal fibers, they are significantly altered by truncal and highly selective vagotomy. Both these operations result in decreased gastric compliance, shifting the volume/pressure curve to the left. Presumably for any given amount of food ingested, the intragastric pressure is higher, and perhaps in some patients the onset of satiety is sooner. This may be one explanation for weight loss associated with vagotomy, and it also helps explain accelerated liquid gastric emptying postvagotomy, which likely contributes to dumping symptoms in some patients.

NO and VIP are the principal mediators of proximal gastric relaxation. But a variety of other agents increase proximal gastric relaxation and compliance, including dopamine, gastrin, CCK, secretin, GRP, and glucagon. Proximal gastric tone also is decreased by duodenal distention, colonic distention, and ileal perfusion with glucose (ileal brake).

The distal stomach breaks up solid food and is the main determinant of gastric emptying of solids. Slow waves of myoelectric depolarization sweep down the distal stomach at a rate of about three per minute. These waves originate from the proximal gastric pacemaker, high on the greater curvature. The pacing cells appear to be the interstitial cells of Cajal, which have been shown to have a similar function in the small intestine and colon. Most of these myoelectric waves are below the threshold for smooth muscle contraction in the quiescent state and thus are associated with negligible changes in pressure. Neural and/or hormonal input, which increases the plateau phase of the action potential, can trigger muscle contraction, resulting in a peristaltic wave associated with the electrical slow wave and of the same frequency (three per minute) (Fig. 26-17). There are measurable abnormalities in gastric slow wave activity in disorders of gastric motility such as gastroparesis, but reliable electrogastrography to aid in the diagnosis and management of these problems is not yet a clinical reality. It is likely that implantable gastric pacemakers benefit some patients with gastroparesis by favorably impacting this myoelectric coupling, normalizing gastric slow wave patterns.

During fasting, distal gastric motor activity is controlled by the migrating motor complex (MMC), the “gastrointestinal housekeeper” (Fig. 26-18). The purported function of the MMC is to sweep along any undigested food, debris, sloughed cells, and mucus after the fed phase of digestion is complete. The MMC lasts approximately 100 minutes (longer at night, shorter during daytime) and is divided into four phases. Phase I (about half the length of the entire cycle) is a period of relative motor inactivity. High-amplitude muscular contractions do not occur in phase I of the MMC. Phase II (about 25% of the entire MMC cycle) consists of some irregular, high-amplitude, generally nonpropulsive contractions. Phase III, a period of intense,
regular (about three per minute), propulsive contractions, only lasts about 5 to 10 minutes. Most phase III complexes of the GI MMC begin in the stomach, and the frequency approximates that of the myoelectric gastric slow wave. Phase IV is a transition period.

Neurohormonal control of the MMC is poorly understood, but it appears that different phases are regulated by different mechanisms. For example, vagotomy abolishes phase II of the gastric MMC but has little influence on phase III that persists even in the autotransplanted stomach, totally devoid of extrinsic neural input. This suggests that phase III is regulated by intrinsic nerves and/or hormones. Indeed, the initiation of phase III of the MMC in the distal stomach corresponds temporally to elevation in serum levels of motilin, a hormone produced in the duodenal mucosa. Resection of the duodenum abolishes distal gastric phase III in dogs, and resection of the duodenum in humans (e.g., with pancreaticoduodenectomy, the Whipple procedure) commonly results in early postoperative delayed gastric emptying. There are clearly motilin receptors on gastric smooth muscle and nerves. Other modulators of gastric MMC activity include NO, endogenous opioids, intrinsic cholinergic and adrenergic nerves, and duodenal pH. The onset of MMC phase III signals the return of hunger in humans, but oddly ghrelin, a major orexigenic hormone, appears to have little to do with phase III.63

Feeding abolishes the MMC and leads to the fed motor pattern. The fed motor pattern of gastric activity starts within 10 minutes of food ingestion and persists until all the food has left the stomach. The neurohormonal initiator of this change is unknown, but CCK and the vagus appear to play some role since sham feeding transiently induces antral motor activity resembling the fed motor pattern which is blocked by the CCK receptor antagonist loxiglumide. Gastric motility during the fed pattern resembles phase II of the MMC, with irregular but continuous phasic contractions of the distal stomach. During the fed state, about half of the myoelectric slow waves are associated with strong higher frequency distal gastric contractions. Some are prograde and some are retrograde, serving to mix and grind the solid components of the meal. The magnitude of gastric contractions and the duration of the pattern are influenced by the consistency and composition of the meal.

The pylorus functions as an effective regulator of gastric emptying and an effective barrier to duodenogastric reflux. Bypass, transection, or resection of the pylorus may lead to uncontrolled gastric emptying of food and the dumping syndrome (see “Postgastrectomy Problems”). Pyloric dysfunction

Figure 26-17. The relationship between intracellular electrical activity and muscle cell contraction. Note that contractile activity is always associated with electrical activity, but the converse is not so. During mechanical quiescence, there are regular depolarizations that do not reach threshold. In the stimulated state, the threshold for contraction is reached, and motor activity is demonstrable. (Reproduced with permission from Kim CH, Malagelada JR: Electrical activity of the stomach: clinical implications, Mayo Clin Proc. 1986 Mar;61(3):205-210.)

or disruption may also result in uncontrolled entry of duodenal contents into the stomach. Perfusion of the duodenum with lipids, glucose, amino acids, hypertonic saline, or hydrochloric acid results in closure of the pylorus and decreased transpyloric flow. Ileal perfusion with fat has the same effect. A variety of neurohumoral pathways are involved with these physiologic responses, and there is evidence that different pathways may be involved for different stimuli.

The pylorus is readily apparent grossly as a thick ring of muscle and connective tissue. The density of nerve tissue in the pyloric smooth muscle is several folds higher than in the antrum, with increased numbers of neurons staining positive for substance P, neuropeptide Y, VIP, and galanin. Interstitial cells of Cajal are closely associated with pyloric myocytes, and the myoelectric slow wave of the pylorus has the same frequency as that seen in the distal stomach. The motor activity of the pylorus is both tonic and phasic. During phase III of the MMC, the pylorus is open as gastric contents are swept into the duodenum. During the fed phase, the pylorus is closed most of the time. It relaxes intermittently, usually in synchronization with lower-amplitude, minor antral contractions. The higher-amplitude, more major antral contractions are usually met with a closed pylorus, facilitating retropulsion and further grinding of food.

Modulation of pyloric motor activity is complex. There is evidence for both inhibitory and excitatory vagal pathways. Some contractile vagal effects are mediated by opioid pathways because they are blocked by naloxone. Electrical stimulation of the duodenum causes the pylorus to contract, whereas electrical stimulation of the antrum causes pyloric relaxation. Nitric oxide is an important mediator of pyloric relaxation. Other molecules that may play a physiologic role in controlling pyloric smooth muscle include serotonin, VIP, prostaglandin E1, and galanin (pyloric relaxation); and histamine, CCK, and secretin (pyloric contraction).

**Gastric Emptying.** The control of gastric emptying is complex. In general, liquid emptying is faster than solid emptying. Osmolarity, acidity, caloric content, nutrient composition, and particle size are important modulators of gastric emptying. Stimulation of duodenal osmoreceptors, glucoreceptors, and pH receptors clearly inhibits gastric emptying by a variety of neurohumoral mechanisms. CCK has been consistently shown to inhibit gastric emptying at physiologic doses (Fig. 26-19). Recently, it has been noted that the anorexigenic hormone leptin, secreted largely by fat but also by gastric mucosa, inhibits gastric emptying, perhaps through the same pathway as CCK (which also has properties of a satiety hormone). The orexigenic hormone ghrelin has the opposite effect.

**Liquid Emptying.** The gastric emptying of water or isotonic saline follows first-order kinetics, with a half emptying time around 12 minutes. Thus, if one drinks 200 mL of water, about 100 mL enters the duodenum by 12 minutes, whereas if one drinks 400 mL of water, about 200 mL enters the duodenum by 12 minutes. This emptying pattern of liquids is modified considerably as the caloric density, osmolarity, and nutrient composition of the liquid changes (Fig. 26-20). Up to an osmolarity of about 1 M, liquid emptying occurs at a rate of about 200 kcal per hour. Duodenal osmoreceptors and hormones (e.g., secretin and VIP) are important modulators of liquid gastric emptying. Generally, liquid emptying is delayed in the supine position.

Traditionally, liquid emptying has been attributed to the activity of the proximal stomach, but it is probably more complicated than previously thought. Clearly, receptive relaxation and gastric accommodation play a role in gastric emptying of liquids. Patients with a denervated (e.g., vagotomized), resected, or plicated (e.g., fundoplication) proximal stomach have decreased gastric compliance and may show accelerated gastric emptying of liquids.

Some observations suggest an active role for the distal stomach in liquid emptying. For instance, even if the proximal intragastric pressure is lower than duodenal pressure, normal gastric emptying of liquids can occur. Also, diabetic patients...
Signs and Symptoms
The most common symptoms of gastric disease are pain, weight loss, early satiety, and anorexia. Nausea, vomiting, bloating, and anemia also are frequent complaints. Several of these symptoms (pain, bloating, nausea, and early satiety) are often described by physicians as dyspepsia, synonymous with the common nonmedical term indigestion. Common causes of dyspepsia include gastroesophageal reflux disease (GERD), helicobacter gastritis, and other disorders of the stomach, gallbladder, and pancreas. Although none of the aforementioned symptoms alone is specific for gastric disease, when elicited in the context of a careful history and physical examination, they point to a differential diagnosis, which can be refined with certain tests. Early endoscopy should be considered in patients presenting with recent onset of alarm symptoms (weight loss, anemia, dysphagia, vomiting) particularly those over 55 years of age (Table 26-5).

Diagnostic Tests
Esophagogastroduodenoscopy. Esophagogastroduodenoscopy (EGD) is a safe and accurate outpatient procedure performed under conscious sedation. Smaller flexible scopes with excellent optics and a working channel are easily passed transnally in the unsedated patient. Following an 8-hour fast, the flexible scope is advanced under direct vision into the esophagus, stomach, and duodenum. The fundus and GE junction are inspected by retroflexing the scope. To rule out cancer with a high degree of accuracy, all patients with gastric ulcer diagnosed on upper GI series or found at EGD should have multiple biopsy specimens of the base and rim of the lesion. Brush cytology also should be considered. Gastritis should be biopsied both for histologic examination and assessment (see discussion on gastritis in “Helicobacter Pylori Infection”) and for a tissue urease test and histologic evaluation to rule out the presence of *H pylori*. If Helicobacter infection is detected, it should be treated because of the etiologic association with peptic ulcers, mucosa-associated lymphoid tissue (MALT), and gastric cancer; in addition, eradication may ameliorate symptoms. The most serious complications of EGD are perforation (which is rare, but can occur anywhere from the cervical esophagus to the duodenum), aspiration, and respiratory depression from excessive sedation. Although EGD is a more sensitive test than double-contrast upper GI series, these modalities should be considered complementary rather than mutually exclusive.

Barium Upper GI Study. Plain abdominal X-rays may be helpful in the diagnosis of gastric perforation (pneumoperitoneum) or delayed gastric emptying (large air-fluid level). Double-contrast upper GI series may be better than EGD at elucidating gastric diverticula, fistula, tortuosity, stricture location, and size or morphology of hiatal hernia. Although there are radiologic characteristics of ulcers that suggest the presence or absence of malignancy, gastric ulcers always require adequate biopsy.

Computed Tomographic Scanning and Magnetic Resonance Imaging. Usually, significant gastric disease can be diagnosed without these sophisticated imaging studies. However, one or the other should be part of the routine staging work-up for patients with a malignant gastric tumor. Magnetic resonance imaging (MRI) may prove clinically useful as a quantitative test for gastric emptying, and it may even hold some promise for the
analysis of myoelectric derangements in patients with gastroparesis. Virtual gastroscopy using multi detector CT scan or MRI is not yet widely used, but these techniques may prove useful for screening and staging of gastric disease. CTA or MRA is useful in evaluating the blood supply to the stomach after endovascular treatment of aortic and/or visceral arterial disease or in patients with previous upper abdominal operation in whom gastric conduit construction is contemplated, e.g., with esophagectomy.

Arteriography can be helpful in the occasional poor-risk patient with exsanguinating gastric hemorrhage, in the patient with occult gastric bleeding, or when CTA or MRA is inconclusive in delineating vascular anatomy.

Endoscopic Ultrasound. Endoscopic ultrasound (EUS) is useful in the evaluation and management of gastric mass lesions. Local staging of gastric adenocarcinoma with EUS is quite accurate, and this modality can be used to plan therapy. At many centers, patients with transmural and/or node positive adenocarcinoma of the stomach are considered for preoperative (neoadjuvant) chemoradiation therapy. EUS is the best way to clinically stage these patients locoregionally. Suspicious nodes can be sampled with EUS-guided endoscopic needle biopsy. Malignant tumors that are confined to the mucosa on EUS may be amenable to endoscopic mucosal resection (EMR). EUS also can be used to assess tumor response to chemotherapy. Submucosal masses are commonly discovered during routine EGD. Large submucosal masses should be resected unless benign pathology is a certainty, but observation may be appropriate for some small submucosal masses (e.g., lipoma or leiomyoma). There are endoscopic characteristics of benign and malignant mesenchymal tumors, and thus, EUS can provide reassurance, but no guarantee, that small lesions under observation are probably benign. Thus, EUS-guided needle biopsy should be considered. Submucosal varices also can be assessed by EUS.

Gastric Secretory Analysis. Analysis of gastric acid output requires gastric intubation, and it is performed infrequently nowadays. This test may be useful in the evaluation of patients with hypergastrinemia, including the Zollinger-Ellison syndrome (ZES), patients with refractory ulcer or GERD, and patients with recurrent ulcer after operation. Historically, gastric analysis was performed most commonly to test for the adequacy of vagotomy in postoperative patients with recurrent or persistent ulcer. Now this can be done by assessing peripheral...
pancreatic polypeptide levels in response to sham feeding. A 50% increase in pancreatic polypeptide within 30 minutes of sham feeding suggests intact vagal function.

Normal basal acid output (BAO) is greater than 5 mEq/h. MAO is the average of the two final stimulated 15-minute periods and is usually 10 to 15 mEq/h. Peak acid output is defined as the highest of the four stimulated periods. Patients with a gastrinoma commonly have a high BAO, often above 30 mEq/h, but consistently above 15 mEq/h unless there has been previous vagotomy or gastric resection. In patients with gastrinoma, the ratio of BAO to MAO exceeds 0.6. Normal acid output in the patient prescribed acid-suppressive medication usually means that the patient is noncompliant. To assess acid-secretory capacity in the absence of medication effect, H₂ blockers and PPIs should be withheld for several days before gastric analysis.

**Scintigraphy.** The standard scintigraphic evaluation of gastric emptying involves the ingestion of a test meal with one or two isotopes and scanning the patient under a gamma camera. A curve for gastric emptying is plotted, and the half-time is calculated. Normal standards exist at each facility. Duodenogastric reflux can be quantitated by the IV administration of hepatobiliary iminodiacetic acid (HIDA scan), which is concentrated and excreted by the liver into the duodenum. Software allows a semiquantitative assessment of how much of the isotope refluxes into the stomach. Positron emission tomography (PET) scan or CT/PET scan is useful in staging certain patients with gastric malignancy.

**Tests for Helicobacter pylori.** A variety of tests can help the clinician to determine whether the patient has active *H. pylori* infection. The predictive value (positive and negative) of any of these tests when used as a screening tool depends on the prevalence of *H. pylori* infection in the screened population. A positive test is quite accurate in predicting *H. pylori* infection, but a negative test can be unreliable. Thus, in the appropriate clinical setting, treatment for *H. pylori* should be initiated on the basis of a positive test, but not necessarily withheld if the test is negative. *Helicobacter* infection should be treated when the diagnosis is made and eradication is confirmed.

A positive serologic test is presumptive evidence of active infection if the patient has never been treated for *H. pylori*. Histologic examination of gastric mucosal biopsy using special stains is the gold standard test for helicobacter infection. Other sensitive tests include commercially available rapid urease tests, which assay for the presence of urease in mucosal biopsy specimens (strong presumptive evidence of infection). Urease is an omnipresent enzyme in *H. pylori* strains that colonize the gastric mucosa. The carbon labeled urea breath test has become the standard test to confirm eradication of *H. pylori* following appropriate treatment. In this test, the patient ingests urea labeled with nonradioactive ¹³C or ¹⁴C. The labeled urea is acted upon by the urease present in the *H. pylori* and converted into appropriate infection and may also be used to confirm cure after treatment. Helicobacter culture may be useful to assess antimicrobial resistance in persistently recalcitrant cases.

**Antroduodenal Motility Testing and Electrogastrography.** Antroduodenal motility testing and electrogastrography (EGG) are performed in specialized centers and may be useful in the evaluation of the occasional patient with dyspeptic symptoms. EGG consists of the transcutaneous recording of gastric myoelectric activity. Antroduodenal motility testing is done with a tube placed transnasally or transorally into the distal duodenum. There are pressure-recording sensors extending from the stomach to the distal duodenum. The combination of these two tests together with scintigraphy provides a thorough assessment of gastric motility.

**HELIcobacter pylori INFECTION**

Over 50% of people worldwide are infected with *Helicobacter pylori*. Infection with *H. pylori* is a chronic disease and does not resolve spontaneously without specific treatment. Worldwide, *H. pylori*-induced gastritis accounts for 80% to 90% of all gastritis. Chronic gastritis associated with *H. pylori* is the most important risk factor for peptic ulcer and gastric adenocarcinoma. Successful *H. pylori* treatment largely eliminates recurrent peptic ulcer in infected patients, and eradication of *H. pylori* worldwide would eliminate most cases of gastric infection.
adenocarcinoma, a major cause of cancer death worldwide. \(^7\)
Helicobacter pylori infection is also associated with MALT lymphoma, dyspepsia, hyperplastic gastric polyps, and even immune thrombocytopenic purpura.

Human beings are the only reservoir for \(H\) pylori. Infection is presumed to occur by oral ingestion of the bacterium, which dramatically alters the gastric microbiome. In helicobacter-infected individuals, 90% of gastric bacteria are helicobacter, whereas in helicobacter-negative patients 90% of gastric bacteria are a combination of firmicutes, actinobacteria, bacteroidetes, proteobacteria, and fusobacteria. The prevalence of \(H\) pylori infection varies among populations and is strongly correlated with socioeconomic conditions. In developing countries, \(H\) pylori infection usually occurs in childhood, and over 80% of adults are infected. Reinfection after curative treatment is common. Infection rates are lower in industrialized countries, and the prevalence of infection in the United States has been declining since the second half of the 19th century as hygiene and sanitation have improved. Nonetheless, \(H\) pylori infection is predicted to remain endemic in the United States for the next century. Family members of infected individuals and healthcare workers are at increased risk of infection.

With specialized flagella and a rich supply of urease, \(H\) pylori is uniquely equipped for survival in the hostile environment of the stomach.\(^{[79-81]}\) Helicobacter strains that lack either flagella or urease are nonpathogenic. The pathogenesis of helicobacter infection involves survival in the acidic gastric lumen, flagellated movement from the lumen across the mucus layer to the surface epithelial cell, adhesion to the surface epithelial cell, and toxin production. Up to 15% of the protein in a Helicobacter organism is composed of cytoplasmic urease that converts periplasmic urea into \(CO_2\) and ammonia. This buffers the surrounding acid, allowing the bacteria to survive the inimical luminal environment until it can burrow deeply into the surface mucus, propelled by its flagella (Fig. 26-23). \(H\) pylori typically does not invade the surface epithelial cell layer. Rather, it triggers a host immune response by attaching to gastric epithelial cells. Important Helicobacter adhesins mediating surface cell injury include neutrophil activating protein A, heat shock protein 60, and sialic acid–binding adhesin. Helicobacter-produced toxins include vacuolating cytotoxin A and cag A (cytotoxin-associated gene A). The initial inflammatory response to Helicobacter infection is characterized by recruitment of neutrophils, followed sequentially by T and B lymphocytes, plasma cells, and macrophages (Fig. 26-24). The resultant chronic gastric inflammation in affected individuals is characterized by enhanced mucosal expression of multiple cytokines and the presence of reactive oxygen and nitrogen species, and long-term infection is associated with mucosal cell DNA damage and chromosomal instability and increased apoptosis (Fig. 26-25).\(^{[80,81]}\) The net effect is a weakening of mucosal defenses. The mechanism by which the helicobacter organism avoids recognition and destruction by the mucosal immune system is a topic of interest and active research.\(^{[82]}\)

Acute \(H\) pylori infection causes a nonerosive pangastritis that is invariably followed by the development of chronic gastritis. Chronic antral gastritis with sparing of the proximal stomach occurs in about 10% of infected patients, and this predisposes to peptic ulcer disease (PUD). The other 90% of Helicobacter-infected patients develop chronic inflammation of the proximal stomach (corpus dominant gastritis), which can lead to gastric cancer in about 1% to 3% of this group.

\(H\) pylori infection is the major cause of peptic ulceration. Patients with \(H\) pylori infection and antral gastritis are three and one-half times more likely to develop PUD than patients without \(H\) pylori infection. Up to 90% of patients with duodenal ulcers, and at least 70% of patients with gastric ulcers, have \(H\) pylori infection. It is clear from multiple randomized prospective studies that curing \(H\) pylori infection dramatically alters the natural history of PUD, decreasing the recurrent ulcer rate from more than 75% in patients treated with a course of acid-suppressive therapy alone (in whom \(H\) pylori is not eradicated) to less than 20% in patients treated with a course of antibacterial therapy (Fig. 26-26).\(^{[83]}\)

In patients with duodenal ulcer caused by helicobacter, the associated antral gastritis leads to relative hypergastrinemia by depleting antral somatostatin, the primary inhibitor of antral gastrin release. \(H\) pylori infection is associated with decreased
levels of somatostatin, decreased somatostatin messenger RNA production, and fewer somatostatin-producing D cells. The mechanism of decreased antral somatostatin synthesis and release may be related to (a) antral alkalinization due to helicobacter urease (acid in the antrum releases somatostatin); (b) toxic cytokine effect on antral D cells; and/or (c) Helicobacter production of N-α-methylhistamine, an H₂ receptor agonist, which binds H₂ receptors on the antral D cell and decreases somatostatin release. Since the gastritis does not involve the oxyntic mucosa, hypergastrinemia leads to hyperacidity and parietal cell hyperplasia. The acid hypersecretion and the antral gastritis are thought to lead to antral epithelial metaplasia in the postpyloric duodenum. This duodenal metaplasia allows H pylori to colonize the duodenal mucosa, and this is where the duodenal ulcer occurs. In fact, in patients with gastric metaplasia of the duodenum, the risk of developing a duodenal ulcer increases 50-fold. When H pylori colonizes the duodenum, there is a significant decrease in acid-stimulated duodenal bicarbonate release. When H pylori infection is successfully treated, acid secretory physiology tends to normalize. Relapse of duodenal ulcer after eradication of H pylori may signal reinfection of the gastric mucosa by the organism.

Many patients with antral dominant helicobacter gastritis never develop duodenal ulcer, and some patients with peptic ulcer do not have Helicobacter. This obviously suggests that there are other important pathogenetic factors involved in peptic ulcer. And even in the presence of active H pylori infection, strong acid suppression usually heals peptic ulcer, an observation consistent with the old dictum “no acid, no ulcer.” But successful helicobacter treatment eliminates ulcer recurrence and the need for long-term PPI. And long-term PPI in patients with active Helicobacter infection may lead to corpus predominant gastritis, which leads to atrophic gastritis and increases the risk of gastric cancer. Thus, Helicobacter infection should be treated and eradication confirmed.

Testing for H pylori infection should be performed in patients with peptic ulcer, gastritis, significant dyspepsia, MALT lymphoma, and early gastric cancer. Noninvasive methods for diagnosis of H pylori infection include the urea breath test, serology, and detection of stool antigen. The urea breath test has a sensitivity and specificity of greater than 90% and is useful for initial diagnosis of infection and for follow-up after eradication therapy since it is positive only in the presence of active infection. The stool antigen test is another noninvasive test to detect active H pylori infection, but it is recommended that only locally validated tests be used. Because H pylori induces a strong immunologic response, serological testing is useful but may not be as accurate as the urea breath test or the stool antigen test, and a positive serology persists after eradication of H pylori infection, so serology is not useful to confirm successful treatment of Helicobacter infection. H pylori infection can also be diagnosed by histologic evaluation of gastric biopsies and/or the rapid urease test on fresh biopsies. Culture of H pylori is not routine and is usually reserved for recurrent infection and for antibiotic sensitivity testing when second-line therapy has failed. All tests for H pylori have a false negative rate. Empiric Helicobacter treatment can be considered despite negative tests if clinical likelihood of infection is high, e.g., a compliant nonsmoking, non–NSAID-consuming patient facing operation for nonhealing peptic ulcer or a patient with unexplained gastritis.

Patients with a positive test should be treated and eradication confirmed. Spontaneous cure without treatment is very rare. It is important to note that none of the therapeutic regimens reported to date cure H pylori infection in 100% of patients. To be effective, antimicrobial drugs must be combined with gastric acid secretion inhibitors or bismuth salts. The Maastricht V/Florence Consensus Report provides current recommendations for diagnosis and treatment of H pylori infection in various clinical scenarios, including recommendations for areas with high metronidazole and clarithromycin resistance. Ideally, a treatment regimen is chosen with 90% effectiveness. Treatment failure requires an alternative course of therapy. Failure to eradicate infection after two trials should prompt Helicobacter culture and sensitivity testing and referral to a specialist. With assiduous treatment, Helicobacter eradication can be achieved in nearly every patient. Patients with atrophic gastritis require endoscopic surveillance (see discussion of gastritis later in this
chapter) because the same sequence of inflammation to metaplasia to dysplasia to carcinoma, that is well known to occur in the esophagus from reflux-induced inflammation (and in the colon from inflammatory bowel disease), is now increasingly well recognized to occur in the stomach with Helicobacter-induced gastritis. Helicobacter also clearly has an etiologic role in the development of gastric lymphoma.

**PEPTIC ULCER DISEASE**

Peptic ulcers are focal defects in the gastric or duodenal mucosa that extend into the submucosa or deeper. They may be acute or chronic and, ultimately, are caused by an imbalance between mucosal defenses and acid/peptic injury (Fig. 26-27).\(^{89,90}\) Peptic ulcer remains a common outpatient diagnosis, but the number of
PUD is one of the most common GI disorders in the United States with a prevalence of about 2%, and a lifetime cumulative prevalence of about 10%, peaking around age 70 years. The costs of PUD, including lost work time and productivity, are estimated to be above $8 billion per year in the United States. In 1998, approximately 1.5% of all Medicare hospital costs were spent treating PUD, and the crude mortality rate for peptic ulcer was 1.7 per 100,000 individuals. Using the National Inpatient Sample, it can be estimated that the mortality rate in patients hospitalized in 2006 with duodenal ulcer was 3.7% compared to 2.1% for gastric ulcer, and the age adjusted hospitalization rate was 56.5 per 100,000, down 21% from the previous decade. Recent studies have shown an increase in the rates of hospitalization and mortality in elderly patients for the peptic ulcer complications of bleeding and perforation. This may be due in part to the increasingly common use of NSAIDs and aspirin in this elderly cohort, many of whom also have *H. pylori* infection.

### Pathophysiology and Etiology

A variety of factors may contribute to the development of PUD. Although it is now recognized that the large majority of duodenal and gastric ulcers are caused by *H. pylori* infection (see previous discussion on *H. pylori*) and/or NSAID use ([Fig. 26-28](#)), the final common pathway to ulcer formation is acid-peptic injury of the gastroduodenal mucosal barrier. Acid suppression heals both duodenal and gastric ulcers and prevents recurrence if continued. In general, *H. pylori* predisposes to ulceration, both by acid hypersecretion and by compromise of mucosal defense mechanisms. NSAID use causes ulcers predominantly by compromise of mucosal defenses. Duodenal ulcer was traditionally viewed as a disease of increased acid-peptic action on the duodenal mucosa, whereas gastric ulcer was viewed as a disease of weakened mucosal defenses. An increased understanding of peptic ulcer pathophysiology has blurred this overly simplistic distinction. Clearly, weakened mucosal defenses play a role in both duodenal and gastric ulcers, and acid hypersecretion may result in a duodenal or gastric ulcer in the setting of normal mucosal defenses.

Elimination of *H. pylori* infection or NSAID use is important for optimal ulcer healing, and perhaps is even more important in preventing ulcer recurrence and/or complications. A variety of other diseases are known to cause peptic ulcer, including ZES (gastrinoma), antral G-cell hyperfunction and/or hyperplasia, systemic mastocytosis, trauma, burns, and major physiologic stress. Other causative agents include drugs (all NSAIDs, aspirin, and cocaine), smoking, and psychologic stress. In the United States, probably more than 90% of serious peptic ulcer complications can be attributed to *H. pylori* infection, NSAID use, and/or cigarette smoking.

### Acid Secretion and Peptic Ulcer

A variety of abnormalities related to mucosal acid exposure have been described in patients with duodenal ulcer ([Fig. 26-29](#)). Although duodenal ulcer patients as a group have a higher mean BAO and mean MAO compared to normal controls, many duodenal ulcer patients have basal and peak acid outputs in the normal range, and there is no correlation between acid secretion and the severity of the ulcer disease. As a group, duodenal ulcer patients produce more acid than normal controls in response to any known acid secretory stimulus. Although they usually have normal fasting serum gastrin levels, DU patients often produce more gastric acid at

---

**Figure 26-26.** *Helicobacter* treatment dramatically decreases the recurrence rate of duodenal and gastric ulcer. (Reproduced with permission from Peek RM, Blaser MJL: Pathophysiology of *Helicobacter pylori*-induced gastritis and peptic ulcer disease, Am J Med. 1997 Feb;102(2):200-207.)

**Figure 26-27.** Balance of aggressive and defensive factors in the gastric mucosa. (Reproduced with permission from Suerbaum S, Michetti P: *Helicobacter pylori* infection, N Engl J Med. 2002 Oct 10;347(15):1175-1186.)
any given dose of gastrin than controls. Considering that many duodenal ulcer patients do produce excessive gastric acid, it has been argued that a "normal" fasting gastrin level in these patients is inappropriately high, and that there is an impaired feedback mechanism, especially in light of the apparently increased sensitivity of the parietal cell mass to gastrin. Many of these long-standing observations now seem reasonable in light of recently gained understanding of the perturbations in acid and gastrin secretion associated with \textit{H pylori} infection. Some patients with duodenal ulcer also have increased rates of gastric emptying that deliver an increased acid load per unit of time to the duodenum. Finally, the buffering capacity of the duodenum in many patients with duodenal ulcer is compromised due to decreased duodenal bicarbonate secretion and duodenal gastric metaplasia.

In patients with gastric ulcer, acid secretion is variable. Currently, five types of gastric ulcer are described, although the original Johnson classification contained three types (Fig. 26-30).\textsuperscript{96} The most common, Johnson type I gastric ulcer, is typically located near the angularis incisura on the lesser curvature, close to the border between antral and corpus mucosa. Patients with type I gastric ulcer usually have normal or decreased acid secretion. Type II gastric ulcer is associated with active or quiescent duodenal ulcer disease, and type III gastric ulcer is prepyloric ulcer disease. Both type II and type III gastric ulcers are associated with normal or increased gastric acid secretion and surgically are treated similar to duodenal ulcer. Type IV gastric ulcers occur near the GE junction, and acid secretion is normal or below normal. Type V gastric ulcers are medication induced and may occur anywhere in the stomach. Patients with gastric ulcers may have weak mucosal defenses that permit an abnormal amount of injurious acid back-diffusion into the mucosa. Duodenogastric reflux may play a role in weakening the gastric mucosal defenses, and a variety of components in duodenal juice, including bile, lysolecithin, and pancreatic juice, have been shown to cause injury and inflammation in the gastric mucosa. NSAIDs and aspirin have similar effects. Although chronic gastric ulcer usually is associated with surrounding gastritis, it is unproven that the latter leads to the former.

\textbf{Nonsteroidal Anti-Inflammatory Drugs in Peptic Ulcer Disease.} Chronic use of NSAIDs (including aspirin) increases the risk of peptic ulcer disease about fivefold and upper GI bleeding at least twofold.\textsuperscript{97-100} Complications of PUD (specifically hemorrhage and perforation) are much more common in patients taking NSAIDs. More than half of patients who present with peptic ulcer hemorrhage or perforation report the recent use of NSAIDs, including aspirin. Many of these patients remain asymptomatic until they develop these life-threatening complications.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{pathophys_b_and_d.png}
\caption{Pathophysiologic abnormalities in DU vary in frequency.}  
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{hpylori.png}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{gastric_table.png}
\caption{Conditions associated with peptic ulcer}  
\end{figure}
The overall risk of significant serious adverse GI events in patients taking NSAIDs is more than three times that of controls (Table 26-6). This risk increases to five times in patients more than age 60 years old. In elderly patients taking NSAIDs, the likelihood that they will require an operation related to a GI complication is 10 times that of the control group, and the risk that they will die from a GI cause is about four and one-half times higher. This problem is put into perspective when one realizes that approximately 20 million patients in the United States take NSAIDs on a regular basis; perhaps even more regularly take aspirin. Persons who take NSAIDs also have a higher hospitalization rate for serious GI events than those who do not.

Factors that clearly put patients at increased risk for NSAID-induced GI complications include age >60, prior GI event, high NSAID dose, concurrent steroid intake, and concurrent anticoagulant intake. Proton pump inhibitors have been shown to significantly decrease upper GI bleeding risk in patients on chronic warfarin, low dose aspirin, and/or antplatelet agents.101-103 Any patient taking NSAIDs or aspirin who has one or more of these risk factors should receive concomitant acid suppressive medication,104 preferably PPI (Table 26-7). High-dose H2 blockers have been shown to be somewhat less effective than PPIs in preventing GI complications in these high-risk patients on antiplatelet therapy, but clearly, they are better than no acid suppression.105

**Smoking, Stress, and Other Factors.** Epidemiologic studies suggest that smokers are about twice as likely to develop PUD as nonsmokers. Smoking increases gastric acid secretion and duodenogastric reflux. Smoking decreases both gastroduodenal prostaglandin production and pancreaticoduodenal bicarbonate production. These observations may be related, and any or all could explain the observed association between smoking and PUD.

Although difficult to measure, both physiologic and psychologic stress undoubtedly play a role in the development of peptic ulcer in some patients.106 In 1842, Curling described duodenal ulcer and/or duodenitis in burn patients. Decades later, Cushing described the appearance of acute peptic ulceration in patients with head trauma (Cushing’s ulcer). Even the ancients recognized the undeniable links between PUD and stress. Patients still present with ulcer complications (bleeding, perforation, and obstruction) that are seemingly exacerbated by stressful life events. The use of crack cocaine has been linked to juxtagastric peptic ulcers with a propensity to perforate. Alcohol is commonly mentioned as a risk factor for PUD, but confirmatory data are lacking.

**Clinical Manifestations**

More than 90% of patients with PUD complain of abdominal pain. The pain is typically nonradiating, burning in quality, and located in the epigastrium. The mechanism of the pain is unclear. Patients with duodenal ulcer often experience pain 2 to 3 hours after a meal and at night. Two-thirds of patients with duodenal ulcers will complain of pain that awakens them from sleep. The pain of gastric ulcer more commonly occurs with eating and is less likely to awaken the patient at night. A history of PUD, use of NSAIDs, over-the-counter antacids, or antisecretory drugs is suggestive of the diagnosis. Other signs and symptoms include nausea, bloating, weight loss, stool positive for occult blood, and anemia. Duodenal ulcer is about twice as common in men compared to women, but the incidence of gastric ulcer is similar in men and women. On average, gastric ulcer patients are older than duodenal ulcer patients, and the incidence is increasing in the elderly, perhaps because of increasing NSAID and aspirin.

**Diagnosis**

In the young patient with dyspepsia and without alarm symptoms, it may be appropriate to initiate empirical PPI therapy for PUD without upper endoscopy or upper GI series. NSAIDs and aspirin should be stopped if the patient is taking these drugs, and *Helicobacter* should be ruled out with testing and treated if present. It is prudent to discuss with the patient the small possibility of an alternative diagnosis, including malignancy, even if symptoms improve with the initiation of empiric therapy. Patients with persistent dyspepsia, and those who cannot stop NSAIDs or aspirin for health reasons should have an upper endoscopy, and all patients, regardless of age, should have this study if any alarm symptoms (see Table 26-5) are present. A double-contrast upper GI X-ray study may be useful. Once an ulcer has been confirmed endoscopically or radiologically, obvious possible causes (*Helicobacter*, NSAIDs, gastrinoma, cancer) should always be considered. All gastric ulcers should be adequately biopsied, and any sites of gastritis should be biopsied to rule out *H pylori*, and for histologic evaluation. Additional testing for *H pylori* may be indicated. It is reasonable to test all peptic

---

**Figure 26-30.** Modified Johnson classification for gastric ulcer. I. Lesser curve, incisura. II. Body of stomach, incisura + duodenal ulcer (active or healed). III. Prepyloric. IV. High on lesser curve, near gastroesophageal junction. V. Medication-induced (NSAID/ acetylsalicylic acid), anywhere in stomach. (Reproduced with permission from Cameron JL: Current Surgical Therapy, 9th ed. Philadelphia, PA: Elsevier/Mosby; 2008.)
ulcer patients and those with nonulcer dyspepsia for *H pylori* (Table 26-8). A baseline serum gastrin level to rule out gastrinoma should be considered if the peptic ulcer is unusual (distal duodenal or jejunal) or if the patient is *Helicobacter* and NSAID negative.

### Complications

The three most common complications of PUD, in decreasing order of frequency, are bleeding, perforation, and obstruction. Most peptic ulcer–related deaths in the United States are due to bleeding. Inhospital mortality and length of stay can be predicted by the AIMS65 score, with a score of 0 predicting negligible mortality and a score of 5 predicting a 30% inhospital mortality. Bleeding peptic ulcers are by far the most common cause of upper GI bleeding in patients admitted to a hospital (Fig. 26-31). Patients with a bleeding peptic ulcer typically present with melena and/or hematemesis. Nasogastric aspiration is usually confirmatory of upper GI bleeding. Abdominal pain is quite uncommon. Shock may be present, necessitating aggressive resuscitation and blood transfusion.

---

**Table 26-6**

Hospitalization rates for GI events with and without NSAID use in selected large populations

<table>
<thead>
<tr>
<th>STUDY</th>
<th>THERAPIES USED</th>
<th>CLINICAL UPPER GI EVENTS</th>
<th>COMPLICATED UPPER GI EVENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSAID CONTROL</td>
<td>NSAID CONTROL</td>
<td>NSAID CONTROL</td>
<td>NSAID CONTROL</td>
</tr>
<tr>
<td>MUCOSA</td>
<td>NSAIDs (n = 4439)</td>
<td>Misoprostol 200 μg four times a day + NSAID (n = 4404)</td>
<td>3.1%</td>
</tr>
<tr>
<td>CLASS</td>
<td>Ibuprofen 800 mg three times a day, diclofenac 75 mg twice a day (n = 3987)</td>
<td>Celecoxib 400 mg twice a day (n = 3995)</td>
<td>3.5%</td>
</tr>
<tr>
<td>VIGOR</td>
<td>Naproxen 500 mg twice a day (n = 4047)</td>
<td>Rofecoxib 50 mg four times a day (n = 4029)</td>
<td>4.5%</td>
</tr>
</tbody>
</table>

*1MUCOSA and VIGOR trials included only rheumatoid arthritis patients; CLASS trial included osteoarthritis (73%) and rheumatoid arthritis (27%).

*2Incidence for MUCOSA trial represents doubling of results provided at 6 months (although median follow-up was <6 months). Incidences for VIGOR and CLASS trials represent rates per 100 patient-years, although VIGOR median follow-up was 9 months, and CLASS data include only the first 6 months of the study.

*3Includes perforations, obstructions, bleeding, and uncomplicated ulcers discovered on clinically indicated work-up.

*4Includes perforation, obstruction, bleeding (documented due to ulcer or erosions in MUCOSA and CLASS; major bleeding in VIGOR).

*21% of patients in CLASS study were taking low-dose aspirin.

Note: All differences between controls and study drugs were significant except clinical upper GI events in overall CLASS study (*P* = .09).


---

**Table 26-7**

Patients taking NSAIDs or aspirin need concomitant acid suppressing medication if any of the following risk factors are present

- Age over 60 years
- History of acid/peptic disease
- Concurrent steroid intake
- Concurrent anticoagulant intake
- High-dose or chronic NSAID use
- High-dose or chronic aspirin use >325 mg/day


---

**Table 26-8**

Indications for diagnosis and treatment of *Helicobacter pylori*

**Established**

- Active peptic ulcer disease (gastric or duodenal ulcer)
- Confirmed history of peptic ulcer disease (not previously treated for *H pylori*)
- Gastric mucosa-associated lymphoid tissue lymphoma (low grade)
- After endoscopic resection of early gastric cancer
- Uninvestigated dyspepsia (depending on *H pylori* prevalence)

**Controversial**

- Nonulcer dyspepsia
- Gastroesophageal reflux disease
- Persons using NSAIDs
- Unexplained iron deficiency anemia
- Populations at higher risk for gastric cancer
Early endoscopy is important to diagnose the cause of the bleeding and to assess the need for hemostatic therapy.

Three-fourths of the patients who come to the hospital with bleeding peptic ulcer will stop bleeding if given acid suppression and nothing by mouth. However, one fourth will continue to bleed or will rebleed after an initial quiescent period, and virtually all the mortalities (and all the operations for bleeding) occur in this group. This group can be fairly well delineated based on clinical factors related to the magnitude of the hemorrhage, comorbidities, age, and endoscopic findings. Shock, hematemesis, transfusion requirement exceeding four units in 24 hours, and certain endoscopic stigmata (active bleeding or visible vessel) define this high-risk group. Risk stratification tools have proven useful in predicting rebleeding and death, and in identifying a low risk cohort. As can be seen in Table 26-9, the maximal Blatchford score is 23, and the maximal Rockall score is 11. The former does not use endoscopic criteria and may be better in identifying the low-risk cohort. Studies have shown that a Blatchford score of 1 or less, or a Rockall score of 2 or less, identifies patients who are very unlikely to be suffering from life-threatening upper GI bleeding. The shorter modified Blatchford score may be just as useful (BUN, Hgb, pulse, BP; maximal score 16).111

High-risk patients benefit from endoscopic therapy to stop the bleeding, while low-risk patients with low-risk lesions can be promptly discharged and treated as outpatients. The most common endoscopic hemostatic modalities used are injection with epinephrine and electrocautery. In a case with exposed vessel, mechanical hemostasis using clips is useful to control the bleeding.112 Biopsy should be performed to evaluate for H pylori infection. Persistent bleeding or rebleeding after endoscopic therapy is an indication for repeat endoscopic treatment. Surgery should be considered after two endoscopic failures. Elderly patients and patients with multiple comorbidities do not tolerate repeated episodes of hemodynamically significant hemorrhage, and they may benefit from early elective operation after initially successful endoscopic treatment, especially if they have a high-risk ulcer.

Planned surgery under controlled circumstances often yields better outcomes than emergent surgery. Deep bleeding ulcers on the posterior duodenal bulb or lesser gastric curvature are high-risk lesions because they often erode large arteries less amenable to nonoperative treatment, and early operation should be considered.

Perforated peptic ulcer usually presents as an acute abdomen. The patient can often give the exact time of onset of the excruciating abdominal pain. Initially, a chemical peritonitis develops from the gastric and/or duodenal secretions, but within hours a bacterial peritonitis supervenes. The patient is in obvious distress, and the abdominal examination shows peritoneal signs. Usually, marked involuntary guarding and rebound tenderness is evoked by a gentle examination. Upright chest X-ray shows free air in about 80% of patients (Fig. 26-32). Once the diagnosis has been made, the patient is given analgesia and antibiotics, resuscitated with isotonic fluid, and taken to the operating room. Fluid sequestration into the third space of the inflamed peritoneum can be impressive, so preoperative fluid resuscitation is mandatory. Sometimes, the perforation has sealed spontaneously by the time of presentation, and surgery can be avoided if the patient is doing well. Nonoperative management is appropriate only if there is objective evidence that the leak has sealed (i.e., radiologic contrast study), and in the absence of clinical peritonitis.

Gastric outlet obstruction occurs in no more than 5% of patients with PUD. It is usually due to duodenal or prepyloric ulcer disease, and it may be acute (from inflammatory swelling and peristaltic dysfunction) or chronic (from cicatrix). Patients typically present with nonbilious vomiting and may have profound hypokalemic hypochloremic metabolic alkalosis and dehydration. Pain or discomfort is common. Weight loss may be prominent, depending on the duration of symptoms. A succussion splash may be audible with stethoscope placed in the epigastrium. Initial treatment is nasogastric suction, IV hydration and electrolyte repletion, and acid suppression. The diagnosis is confirmed by endoscopy. Most patients admitted to the hospital nowadays with obstructing ulcer disease require intervention, either balloon dilation or operation. Cancer must be ruled out because most patients who present with the symptoms of gastric outlet obstruction will have a pancreatic, gastric, or duodenal malignancy.

Medical Treatment of Peptic Ulcer Disease

PPIs are the mainstay of medical therapy for PUD, but high-dose H$_2$RAs and sucralfate are also quite effective. Patients hospitalized for ulcer complications should receive high-dose intravenous PPI and, when discharged, should be considered for lifelong PPIs unless the definitive cause is eliminated or a definitive operation performed. Peptic ulcer patients should stop smoking and avoid alcohol and NSAIDs (including aspirin). Patients who require NSAIDs or aspirin to treat other medical conditions should always take concomitant PPIs or high dose H$_2$ receptor blockers. Testing for H pylori infection is performed, and if it is found, it should be treated with one of several acceptable regimens (Table 26-10).113 If initial H pylori testing is negative and ulcer symptoms persist, an empirical trial of anti-H pylori therapy is reasonable since false-negative H pylori tests are not uncommon. Generally, acid suppression can be stopped after 3 months if the ulcerogenic stimulus (e.g., H pylori, NSAIDs, or aspirin) has been removed. However, long-term maintenance PPI therapy should be considered in all patients admitted to hospital with ulcer complications, all high-risk patients on NSAIDs or aspirin (the elderly or debilitated), and all patients requiring anticoagulation or antiplatelet agents or those with a history of recurrent ulcer or bleeding. Consideration should also be given to maintenance PPI therapy in refractory smokers with a history of peptic ulcer. Sucralfate acts locally on mucosal defects and is well tolerated, and occasionally it is useful as a supplement to acid suppression.
### Table 26-9

**Risk-stratification tools for upper gastrointestinal hemorrhage**

**A. BLATCHFORD SCORE**

<table>
<thead>
<tr>
<th>AT PRESENTATION</th>
<th>POINTS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Systolic blood pressure</strong></td>
<td></td>
</tr>
<tr>
<td>100–109 mmHg</td>
<td>1</td>
</tr>
<tr>
<td>90–99 mmHg</td>
<td>2</td>
</tr>
<tr>
<td>&lt;90 mmHg</td>
<td>3</td>
</tr>
<tr>
<td><strong>Blood urea nitrogen</strong></td>
<td></td>
</tr>
<tr>
<td>6.5–7.9 mmol/L</td>
<td>2</td>
</tr>
<tr>
<td>8.0–9.9 mmol/L</td>
<td>3</td>
</tr>
<tr>
<td>10.0–24.9 mmol/L</td>
<td>4</td>
</tr>
<tr>
<td>≥25 mmol/L</td>
<td>6</td>
</tr>
<tr>
<td><strong>Hemoglobin for men</strong></td>
<td></td>
</tr>
<tr>
<td>12.0–12.9 g/dL</td>
<td>1</td>
</tr>
<tr>
<td>10.0–11.9 g/dL</td>
<td>3</td>
</tr>
<tr>
<td>&lt;10.0 g/dL</td>
<td>6</td>
</tr>
<tr>
<td><strong>Hemoglobin for women</strong></td>
<td></td>
</tr>
<tr>
<td>10.0–11.9 g/dL</td>
<td>1</td>
</tr>
<tr>
<td>&lt;10.0 g/dL</td>
<td>6</td>
</tr>
<tr>
<td><strong>Other variables at presentation</strong></td>
<td></td>
</tr>
<tr>
<td>Pulse ≥100 beats/min</td>
<td>1</td>
</tr>
<tr>
<td>Melena</td>
<td>1</td>
</tr>
<tr>
<td>Syncope</td>
<td>2</td>
</tr>
<tr>
<td>Hepatic disease</td>
<td>2</td>
</tr>
<tr>
<td>Cardiac failure</td>
<td>2</td>
</tr>
</tbody>
</table>

**B. ROCKALL SCORE**

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>POINTS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
</tr>
<tr>
<td>&lt;60 y</td>
<td>0</td>
</tr>
<tr>
<td>60–79 y</td>
<td>1</td>
</tr>
<tr>
<td>≥80 y</td>
<td>2</td>
</tr>
<tr>
<td><strong>Shock</strong></td>
<td></td>
</tr>
<tr>
<td>Heart rate &gt;100 beats/min</td>
<td>1</td>
</tr>
<tr>
<td>Systolic blood pressure &lt;100 mmHg</td>
<td>2</td>
</tr>
<tr>
<td><strong>Coexisting illness</strong></td>
<td></td>
</tr>
<tr>
<td>Ischemic heart disease, congestive heart failure, other major illness</td>
<td>2</td>
</tr>
<tr>
<td>Renal failure, hepatic failure, metastatic cancer</td>
<td>3</td>
</tr>
<tr>
<td><strong>Endoscopic diagnosis</strong></td>
<td></td>
</tr>
<tr>
<td>No lesions observed, Mallory-Weiss syndrome</td>
<td>0</td>
</tr>
<tr>
<td>Peptic ulcer, erosive disease, esophagitis</td>
<td>1</td>
</tr>
<tr>
<td>Cancer of the upper GI tract</td>
<td>2</td>
</tr>
<tr>
<td><strong>Endoscopic stigmata of recent hemorrhage</strong></td>
<td></td>
</tr>
<tr>
<td>Clean base ulcer, flat pigmented spot</td>
<td>0</td>
</tr>
<tr>
<td>Blood in upper GI tract, active bleeding, visible vessel, clot</td>
<td>2</td>
</tr>
</tbody>
</table>

*Panel A shows the values used in the Blatchford risk-stratification score, which ranges from 0 to 23, with higher scores indicating higher risk. Panel B shows the Rockall score, with point values assigned for each of three clinical variables (age and the presence of shock and coexisting illnesses) and two endoscopic variables (diagnosis and stigmata of recent hemorrhage). The complete Rockall score ranges from 0 to 11, with higher scores indicating higher risk. Patients with a clinical Rockall score (Age + Shock + Coexisting illness) of 0 or a complete Rockall score of 2 or less are considered to be at low risk for rebleeding or death.*

Surgical Treatment of Peptic Ulcer Disease
The indications for surgery in PUD are (in order of decreasing frequency) perforation, obstruction, bleeding, and intractability or nonhealing. Gastric cancer must always be considered in patients with gastric ulcer or gastric outlet obstruction. Today, most patients undergoing emergent operation have simple patch of a perforated ulcer or oversewing of a bleeding ulcer. Simultaneous performance of vagotomy either truncal or highly selective is increasingly uncommon, probably due to surgeon unfamiliarity with the procedure and reliance on postoperative PPIs to decrease acid secretion. But even in the current era, vagotomy may improve outcomes in emergency ulcer surgery. Before denying the stable low-risk patient a highly selective vagotomy or truncal vagotomy and drainage as an adjunct to simple patch or oversew, the surgeon should consider that many patients having emergency operation for peptic ulcer will not take long-term PPI, do not have Helicobacter, or will continue to smoke or take NSAIDs.

Unfortunately, the data from many excellent randomized clinical trials evaluating elective operation for peptic ulcer over the last several decades may be irrelevant to most patients presenting for ulcer surgery today. The large majority of these excellent studies were done in the pre-PPI, pre-Helicobacter, pre-NSAID era, and focused on elective operation for intractable disease, an unusual indication for operation nowadays. Thus, today’s surgeon should take great care in applying this literature to inform surgical decision making.

Traditionally, the vast majority of peptic ulcers were treated by a variant of one of the three basic operations: parietal cell vagotomy, also called highly selective vagotomy (HSV) or proximal gastric vagotomy, vagotomy and drainage (V+D), and vagotomy and distal gastrectomy. Recurrence rates are lowest but morbidity highest with the latter procedure, while the opposite is true for HSV (Table 26-11).

HSV severs the vagal nerve supply to the proximal two-thirds of the stomach, where essentially all the parietal cells are located, and preserves the vagal innervation to the antrum and pylorus and the remaining abdominal viscera (Fig. 26-33). Thus, the operation decreases total gastric acid secretion by about 75%, and GI side effects are rare. Elective HSV has largely been supplanted by long-term PPI treatment, but the operation, which has a learning curve, may still be useful in the patient (elective or emergent) who is noncompliant with, intolerant of, or cannot afford medical treatment. Historically, HSV has not performed particularly well for type II (gastric and duodenal) and type III (prepyloric) gastric ulcer, perhaps because of hypergastrinemia caused by gastric outlet obstruction and persistent antral stasis.

Table 26-10

<table>
<thead>
<tr>
<th>Helicobacter pylori therapies (10–14 days)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clarithromycin triple therapy</strong></td>
</tr>
<tr>
<td>standard or double dose PPI twice a day</td>
</tr>
<tr>
<td>clarithromycin 500 mg twice a day</td>
</tr>
<tr>
<td>amoxicillin 1 g twice a day; or metronidazole 500 mg three times a day</td>
</tr>
<tr>
<td><strong>Metronidazole triple therapy</strong></td>
</tr>
<tr>
<td>standard or double dose PPI twice a day</td>
</tr>
<tr>
<td>metronidazole 500 mg twice a day</td>
</tr>
<tr>
<td>amoxicillin 1 g twice a day</td>
</tr>
<tr>
<td><strong>Levofloxacin triple therapy</strong></td>
</tr>
<tr>
<td>standard dose PPI twice a day</td>
</tr>
<tr>
<td>amoxicillin 1 g twice a day</td>
</tr>
<tr>
<td>levofloxacin 500 mg daily</td>
</tr>
<tr>
<td><strong>Sequential therapy</strong></td>
</tr>
<tr>
<td>standard or double dose PPI (10–14 days)</td>
</tr>
<tr>
<td>amoxicillin 1 g twice a day; then clarithromycin 500 mg twice a day and metronidazole 500 mg twice a day (5–7 days)</td>
</tr>
<tr>
<td><strong>Bismuth quadruple therapy</strong> (commonly used when above regimens fail to eradicate H pylori)</td>
</tr>
<tr>
<td>standard dose PPI twice a day</td>
</tr>
<tr>
<td>bismuth subsalicylate 300 mg four times a day</td>
</tr>
<tr>
<td>tetracycline 500 mg four times a day</td>
</tr>
<tr>
<td>metronidazole 250 mg four times a day</td>
</tr>
</tbody>
</table>

PPI = proton pump inhibitor.

Table 26-11

<table>
<thead>
<tr>
<th>Clinical results of surgery for duodenal ulcer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PARIETAL CELL VAGOTOMY</strong></td>
</tr>
<tr>
<td>Operative mortality rate (%)</td>
</tr>
<tr>
<td>Ulcer recurrence rate (%)</td>
</tr>
<tr>
<td>Dumping (%)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Diarrhea (%)</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

The Taylor procedure, a straightforward laparoscopic operation, consists of a posterior truncal vagotomy and anterior seromyotomy (but anterior HSV is probably equivalent), and it is an attractive and simple alternative to HSV with similar results.

Truncal vagotomy and pyloroplasty, and truncal vagotomy and gastrojejunostomy are the paradigmatic vagotomy and drainage procedures. HSV may be substituted for truncal vagotomy. The advantage of V + D is that it can be performed safely and quickly by the experienced surgeon. The main disadvantages are the side effect profile (10% of patients have significant dumping and/or diarrhea). During truncal vagotomy (Fig. 26-34), care must be taken not to perforate the esophagus, a potentially lethal complication. Intraoperative frozen section confirmation of at least two vagal trunks is prudent; additional vagal trunks are common. Unlike HSV, V + D is widely accepted as a successful definitive operation for complicated PUD. It has been described as a useful part of the operative treatment for bleeding duodenal and gastric ulcer, perforated duodenal and gastric ulcer, and obstructing duodenal and gastric (types II and III) ulcer. When applied to gastric ulcer, the ulcer should be excised or biopsied.

Truncal vagotomy denervates the antropyloric mechanism, and therefore, some sort of procedure is necessary to ablate or bypass the pylorus. Gastrojejunostomy is a good choice in patients with gastric outlet obstruction or a severely diseased proximal duodenum. The anastomosis is done between the proximal jejunum and the most dependent portion of the greater gastric curvature, in either an antecolic or retrocolic fashion (Fig. 26-35). Marginal ulceration is a potential complication.
Mechanical complications are also possible such as afferent or efferent loop obstruction, internal hernia, and intussusception. Pyloroplasty is useful in patients who require a pyloroduodenotomy to deal with the ulcer complication (e.g., posterior bleeding duodenal ulcer), in those with limited or focal scarring in the pyloric region, or when gastrojejunostomy is technically difficult. The most commonly performed pyloroplasty is the Heineke-Mikulicz type (Fig. 26-36). Other occasionally useful techniques include the Finney (Fig. 26-37) and the Jaboulay pyloroplasties (Fig. 26-38). These more extensive pyloroplasty techniques may make subsequent distal gastric resection more difficult and/or hazardous.

Although vagotomy and antrectomy (V + A) is associated with a very low ulcer recurrence rate and is applicable to many patients with complicated PUD (e.g., bleeding duodenal and gastric ulcer, obstructing peptic ulcer, nonhealing gastric ulcer, and recurrent ulcer), V + A has a higher operative mortality risk (compared with HSV or V + D), and is irreversible. Following antrectomy, GI continuity may be reestablished with a Billroth I gastroduodenostomy (Fig. 26-39) or a Billroth II loop gastrojejunostomy (Fig. 26-40). Since antrectomy routinely leaves a 60% to 70% gastric remnant, routine reconstruction as a Roux-en-Y gastrojejunostomy should be avoided (Fig. 26-41). Although the Roux-en-Y operation is an excellent procedure for keeping duodenal contents out of the stomach and esophagus, in the presence of a large gastric remnant, this reconstruction will predispose to marginal ulceration and/or gastric stasis.

V + A should be avoided in hemodynamically unstable patients, and in patients with extensive inflammation and/or scarring of the proximal duodenum, because secure
anastomosis (Billroth I) or duodenal closure (Billroth II) may be difficult.

Distal gastrectomy without vagotomy (usually about a 50% gastrectomy to include the ulcer) has traditionally been the procedure of choice for type I gastric ulcer. The addition of vagotomy should be considered for type II and III gastric ulcers (because the pathophysiology is more analogous to duodenal ulcer), or if the patient is believed to be at increased risk for recurrent ulcer, or perhaps even if Billroth II reconstruction is contemplated (to decrease the chance of marginal ulcer). Subtotal gastrectomy (75% distal gastrectomy) without vagotomy is rarely used to treat PUD today, although it was the most popular ulcer operation at the middle of the last century.

Pylorus preserving gastrectomy (PPG) was first reported as a surgical option for gastric ulcer that could minimize both dumping and duodenogastric reflux. Though not widely adopted for this indication, in some centers PPG is considered a good minimally invasive surgical option for early gastric cancer.

**Choice of Operation for Peptic Ulcer.** The choice of operation for the individual patient with PUD depends on a variety of factors, including the type of ulcer (duodenal, gastric, recurrent, or marginal), the indication for operation, and the condition of the patient. Other important considerations are intra-abdominal factors (duodenal scarring/inflammation, adhesions, or difficult exposure), the ulcer diathesis status of the patient, the surgeon’s experience and personal preference, whether *H pylori* infection is present, the need for NSAID therapy, previous treatment, and the likelihood of future compliance with treatment. Table 26-12 shows the surgical options for managing various aspects of PUD.
In general, resective procedures have a lower ulcer recurrence rate, but a higher morbidity and mortality rate (see Table 26-11) compared to nonresective ulcer operations. Because ulcer recurrence often is related to *H. pylori* and/or NSAIDs, it is usually managed adequately without reoperation. Thus, gastric resection to minimize recurrence in duodenal ulcer disease is usually not justified; resection for gastric ulcer remains the standard because of the risk of cancer. Clearly, the modern trend in peptic ulcer operation could be described as “less is more.”

**Bleeding Peptic Ulcer**

Bleeding is the most common cause of ulcer-related death, but only rarely do patients with bleeding gastric or duodenal ulcer require operation today. The success of endoscopic treatment and medical therapy for bleeding PUD has resulted in the selection of a small subgroup of high-risk patients for today’s surgeon. It is likely that patients currently coming to operation for bleeding PUD are at higher risk for a poor outcome than ever before. The surgical options for treating bleeding PUD include suture ligation of the bleeder; suture ligation and definitive nonresective ulcer operation (HSV or V + D); and gastric resection (usually, including vagotomy and ulcer excision). Gastric ulcer requires biopsy if not resected.

The management of bleeding peptic ulcer is summarized in the algorithm provided in Fig. 26-42. All patients admitted to the hospital with bleeding peptic ulcer should be adequately

---

**Figure 26-38.** A through D. Jaboulay pyloroplasty. ant. = anterior; Duod. = duodenum; Stom. = stomach. (Reproduced with permission from Zuidema GD: Shackelford’s Surgery of the Alimentary Tract, 4th ed. Philadelphia, PA: Elsevier/Saunders; 1996.)
resuscitated and started on IV PPI. Most patients will stop bleeding with these measures alone, but about 25% will continue to bleed or will rebleed in hospital. It is important to identify this high-risk group early with clinical and endoscopic parameters because, essentially, all the deaths from bleeding ulcer occur in this group. Surgical consultation is mandatory, and endoscopic hemostatic therapy (cautery, epinephrine injection, clipping) is indicated and usually successful in these high-risk patients.

Indications for operation include massive hemorrhage unresponsive to initial endoscopic control, recurrent hemorrhage...
requiring multiple transfusions after two attempts at endoscopic control, ongoing hemorrhage and transfusion with limited availability of blood for transfusion or lack of availability of a therapeutic endoscopist, early rehospitalization for bleeding ulcer, and concurrent indications for surgery such as perforation or obstruction. Patients with massive bleeding from high-risk lesions (e.g., posterior duodenal ulcer with erosion of gastroduodenal artery, or lesser curvature gastric ulcer with erosion of left gastric artery or branch) should be considered for operation as should those presenting in shock, those requiring more than four units of blood in 24 hours or eight units of blood in 48 hours, and those with ulcers >2 cm in diameter. The mortality rate for surgery for bleeding peptic ulcer is around 20%. Angiography and embolization may be useful in some patients.

**Operation for Bleeding Peptic Ulcer (Fig. 26-43)**
The two operations most commonly used for bleeding duodenal ulcer are oversewing of the ulcer with or without vagotomy and drainage, or V + A. Oversewing alone results in a higher rebleeding rate but a lower operative mortality rate than definitive operation. When the mortality for reoperation for rebleeding is considered, the overall mortality is probably comparable for the two approaches. Patients who are in shock or medically unstable should not have gastric resection.

An initial pyloromyotomy incision allows access to the bleeding posterior duodenal ulcer, and an expeditious Kocher maneuver allows the surgeon to control the hemorrhage with the left hand if necessary. Heavy suture material on a stout needle is used to place figure-of-eight sutures or a U-stitch to secure the bleeding vessel at the base of the posterior duodenal ulcer. Multiple sutures are usually necessary. Once the surgeon is unequivocally convinced that hemostasis is secure, a pyloroplasty can be performed. If the patient is stable, vagotomy may be considered if the surgeon is experienced and the vagotomy is straightforward. If the patient is not a high operative risk and V + A is selected, smaller duodenal ulcers are resected with the specimen; larger bleeding duodenal ulcers must often be left behind in the duodenal stump. In this situation, suture hemostasis must be attained and a secure duodenal closure accomplished. The anterior wall of the open duodenum can be sutured to either the proximal or distal lip of the posterior ulcer once the bleeding vessel has been sutured. The duodenal closure can be buttressed with omentum and the duodenum should be decompressed, either with a lateral duodenostomy or retrograde tube via the proximal jejunum or well secured nasogastric tube secured with tip well into afferent limb. Right upper quadrant closed suction peritoneal drainage is important. Use of a feeding jejunostomy is also considered. A Billroth II anastomosis reestablishes gastrointestinal continuity.

The initial management of bleeding gastric ulcers and the indications for operation are similar to those for bleeding duodenal ulcer. These lesions tend to occur in older and/or medically complicated patients, and this fact may increase the operative risk. However, experience shows that planned surgery in a resuscitated patient results in a better operative survival rate than emergent operation in a patient who has rebled and is in shock. Distal gastric resection to include the bleeding ulcer is the procedure of choice for bleeding gastric ulcer. Second best is V + D with oversewing and biopsy of the ulcer to rule out cancer. Oversewing of the bleeder and biopsy followed by long-term acid suppression is a reasonable alternative in high-risk or unstable patients.
Perforated Peptic Ulcer (Fig. 26-44)
Perforation is the second most common complication of peptic ulcer, but nowadays it is a much more common indication for operation than bleeding. As with bleeding ulcer, NSAID and/or aspirin use have been inextricably linked with perforated PUD, especially in the elderly population. Surgery is almost always indicated for ulcer perforation, although occasionally nonsurgical treatment can be used in the stable patient without peritonitis in whom radiologic studies document a sealed perforation. Patients with acute perforation and GI blood loss (either chronic or acute) should be suspected of having a second ulcer or a GI cancer.

The options for surgical treatment of perforated duodenal ulcer are simple patch closure, patch closure and HSV, or patch closure and V + D. Simple patch closure, currently the most commonly performed operation for perforated peptic ulcer, is the procedure of choice in patients with hemodynamic instability and/or exudative peritonitis signifying a perforation >24 hours old. In stable patients without longstanding perforation, particularly those with chronic symptoms or failure of medical treatment, the addition of HSV should be considered. Vagotomy and drainage is also an acceptable definitive operation for perforated duodenal ulcer, but occasionally side effects are disabling, and if gastrojejunostomy has been performed,

Figure 26-42. Algorithm for the treatment of bleeding peptic ulcer. ASA = acetylsalicylic acid; EGD = esophagogastroduodenoscopy; O.R. = operating room; PPI = proton pump inhibitor; PRBC = unit of packed red blood cells; PT = prothrombin time; PTT = partial thromboplastin time; Rx = treatment.
Operation for bleeding peptic ulcer

Hemodynamically unstable? Or High operative risk?

No

Yes

BMI < 21? or difficult duodenum?

No

Duodenal ulcer

Oversew + TV/D or TV/A

Rebleed

Duodenal ulcer

Oversew + TV/D

Distal gastrectomy**

Duodenal ulcer*

Gastric ulcer*

Type 1, 2, 3

Oversew

1) Biopsy and oversew
2) Wedge resection

Duodenal ulcer*

Gastric ulcer*

Type 4

Rebleed

Csendes proc
Pauchet proc
Kelling-Madlener proc
(see text)

*Add lifelong PPI
**Add TV for type 2 + 3

Figure 26-43. Algorithm for operation for bleeding peptic ulcer. BMI = body mass index; Bx = biopsy; PPI = proton pump inhibitor; proc = procedure; TV = truncal vagotomy; TV/A = truncal vagotomy and antrectomy; TV/D = truncal vagotomy and drainage.

Operation for Perforated Peptic Ulcer

Hemodynamically unstable? Or High operative risk? Or Perforation >24 hours

Chronic ulcer Hx? Or Perforation on RX? Or NSAIDs/ASA necessary?

No

Yes**

Duodenal ulcer*

Gastric ulcer*

Patch + HSV, or Patch + TV/D

Patch

Bx + Patch or Wedge resection

Duodenal ulcer*

Gastric ulcer*

Wedge resection and TV/D or HSV
Distal gastrectomy including perforation***

* In all patients, test and treat for H pylori, and if vagotomy not performed (most patients today) consider lifelong PPI.
** Avoid truncal vagotomy and avoid gastrectomy if BMI<21
*** Consider adding vagotomy for type II and type III gastric ulcer

Figure 26-44. Algorithm for operation for perforated peptic ulcer. ASA = acetylsalicylic acid; BMI = body mass index; Bx = biopsy; HSV = highly selective vagotomy; Hx = history; PPI = proton pump inhibitor; Rx = treatment; TV/D = truncal vagotomy and drainage.
marginal ulcer can be life-threatening. In the United States and Western Europe, there is clearly a trend away from definitive operation for perforated duodenal ulcer.92,121

In the stable patient without multiple operative risk factors, perforated gastric ulcers are best treated by distal gastric resection. Vagotomy is usually added for type II and III gastric ulcers. Patch closure with biopsy; or local excision and closure; or biopsy, closure, truncal vagotomy, and drainage are alternative operations in the unstable or high-risk patient, or in the patient with a perforation in an inopportune location. All perforated gastric ulcers, even those in the prepyloric position, should be biopsied if they are not removed at surgery.

Obstructing Peptic Ulcer
Acute ulcers associated with obstruction due to edema and/or motor dysfunction may respond to intensive antisecretory therapy and nasogastric suction. But most patients with significant obstruction from chronic ulceration will require some sort of more substantial intervention. Endoscopic balloon dilation can often transiently improve obstructive symptoms, but many of these patients ultimately fail and come to operation.123

The standard operation for obstructing PUD is vagotomy and antrectomy. Alternatively, vagotomy and gastrojejunostomy should be considered if a difficult duodenal stump is anticipated with resection. HSV and gastrojejunostomy may be comparable to V + A for obstructing ulcer disease,124 and this procedure is appealing because it can be done laparoscopically and does not complicate future resection, if needed. However, potentially curable gastric or duodenal cancers can be missed with this approach.

Intractable or Nonhealing Peptic Ulcer
Intractability should be an unusual indication for peptic ulcer operation nowadays. The patient referred for surgical evaluation because of intractable PUD should raise red flags for the surgeon: Maybe the patient has a missed cancer; maybe the patient is noncompliant (not taking prescribed PPI, still taking NSAIDs, still smoking); maybe the patient has Helicobacter despite the presence of a negative test or previous treatment. Because acid secretion can be totally blocked and H pylori eradicated with modern medication, the question remains: “Why does the patient have a persistent ulcer diathesis?” The surgeon should review the differential diagnosis of nonhealing ulcer before any consideration of operative treatment (Table 26-13).

Surgical treatment should be considered in patients with nonhealing or intractable PUD who have multiple recurrences, large ulcers (>2 cm), complications (obstruction, perforation, or hemorrhage), or suspected malignancy. Definitive operation, particularly gastric resection, should be considered most cautiously in the thin or marginally nourished individual.

It is important that the surgeon not fall into the trap of performing a large, irreversible operation on these patients, based on the unproven theory that if all other methods have failed to heal the ulcer, a large operation is required. Although there are good data in the surgical literature suggesting that the majority of patients do well after the larger elective ulcer operations, most of these data are several decades old and may not be particularly relevant to the modern patient.114 Candidates for ulcer operation today are different than those of 30 to 50 years ago. One might argue that current medical care has healed the typical peptic ulcers, and that patients presenting with true intractability or nonhealing will be more difficult to treat and are likely to have chronic problems after a major ulcer operation.

If surgery is necessary, a lesser operation may be preferable. It is prudent to avoid truncal vagotomy and/or distal gastrectomy as the initial elective operation for intractable peptic ulcer in the thin or asthenic patient. Alternatives for intractable duodenal ulcer include HSV with or without gastrojejunostomy (reversible drainage operation). In patients with nonhealing gastric ulcer, wedge resection with HSV should be considered in thin or frail patients. Otherwise, distal gastrectomy (to include the ulcer) is recommended. It is unnecessary to add a vagotomy in patients with type I or type IV (juxta-esophageal) gastric ulcers because they are usually associated with acid hyposecretion. Type IV gastric ulcers may be difficult to resect as part of a distal gastrectomy, and a variety of surgical techniques have been described to treat these more proximal lesions (Fig. 26-45).

Zollinger-Ellison Syndrome127-129
ZES is caused by the hypersecretion of gastrin, typically by a duodenal or pancreatic neuroendocrine tumor (i.e., gastrinoma). Most cases (80%) are sporadic, but 20% are inherited. The inherited or familial form of gastrinoma is associated with multiple endocrine neoplasia type I (MEN I), which is characterized by parathyroid, pituitary, and pancreatic (or duodenal) tumors. Gastrinoma is the most common islet cell tumor in patients with MEN I. Patients with MEN I usually have multiple gastrinomas, and surgical cure is usually not achievable; sporadic gastrinomas are more often solitary and are more often amenable to surgical cure. About 50% to 60% of gastrinomas are malignant, with lymph node, liver, or other distant metastases at operation. Five-year survival in patients presenting with metastatic disease is approximately 40%. More than 90% of patients with sporadic, completely resected gastrinoma will be cured.

The most common symptoms of ZES are epigastric pain, GERD, and diarrhea. More than 90% of patients with gastrinoma have peptic ulcers. Most ulcers are in a typical location (proximal duodenum), but atypical ulcer location (distal duodenum, jejunum, or multiple ulcers) should prompt an evaluation for gastrinoma. Gastrinoma also should be considered in the differential diagnosis of recurrent or refractory peptic ulcer, secretory diarrhea, gastric rugal hypertrophy, esophagitis with stricture, bleeding or perforated ulcer, familial ulcer, peptic ulcer with hypercalcemia, and gastric neuroendocrine tumor (carcinoid). The majority of patients with ZES have been symptomatic for several years before definitive diagnosis and,

### Table 26-13
Differential diagnosis of intractability or nonhealing peptic ulcer disease

<table>
<thead>
<tr>
<th>Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gastric</td>
</tr>
<tr>
<td>Pancreatic</td>
</tr>
<tr>
<td>Duodenal</td>
</tr>
<tr>
<td>Persistent <em>Helicobacter pylori</em> infection</td>
</tr>
<tr>
<td>Tests may be false-negative</td>
</tr>
<tr>
<td>Consider empiric treatment</td>
</tr>
<tr>
<td>Noncompliant patient</td>
</tr>
<tr>
<td>Failure to take prescribed medication</td>
</tr>
<tr>
<td>Surreptitious use of NSAIDs</td>
</tr>
<tr>
<td>Motility disorder</td>
</tr>
<tr>
<td>Zollinger-Ellison syndrome</td>
</tr>
</tbody>
</table>
in general, patients with ZES and MEN1 are diagnosed in their 20s and 30s, while those with sporadic ZES more typically are diagnosed in their 40s and 50s.

ZES is an important element in the differential diagnosis of hypergastrinemia (Fig. 26-46). All patients with gastrinoma have an elevated gastrin level, and hypergastrinemia in the presence of elevated BAO strongly suggests gastrinoma. Patients with gastrinoma usually have a BAO >15 mEq/h or >5 mEq/h if they have had a previous procedure for peptic ulcer. Acid secretory medications should be held for several days before gastrin measurement, because acid suppression may falsely elevate gastrin levels. Causes of hypergastrinemia can be divided into those associated with hyperacidity and those associated with hypacidity (see Fig. 26-46). The diagnosis of ZES is confirmed by the secretin stimulation test. An IV bolus of secretin (2 U/kg) is given, and gastrin levels are checked before and after injection. An increase in serum gastrin of 200 pg/mL or greater suggests the presence of gastrinoma. Patients with gastrinoma should have serum calcium and parathyroid hormone levels determined to rule out MEN1 and, if present, parathyroidectomy should be considered before resection of gastrinoma.

About 80% of primary tumors are found in the gastrinoma triangle (Fig. 26-47), and many tumors are small (<1 cm), making preoperative localization difficult. Transabdominal ultrasound is quite specific, but not very sensitive. CT will detect most lesions >2 cm in size, and MRI is comparable. EUS is more sensitive than noninvasive imaging tests, but it still misses many smaller lesions or lesions in inaccessible locations (e.g., the pancreatic tail). Somatostatin receptor scintigraphy (the octreotide scan) or Gallium-68 dotatate PET/CT are sensitive and specific when the pretest probability of gastrinoma is high and may identify sites of regional or distant metastatic disease (Fig. 26-48). Angiographic localization studies are infrequently performed for gastrinoma. Both diagnostic angiography and transhepatic selective venous sampling of the portal system have been supplanted by selective arterial secretin infusion, which helps to localize the tumor as inside or outside the gastrinoma triangle. This study too, is rarely performed given increasing availability of endoscopic ultrasonography and accurate nuclear medicine imaging.

All patients with sporadic (nonfamilial) gastrinoma should be considered for surgical exploration. The lesions can be located in over 90% of patients, and a majority are cured by extirpation of the gastrinoma.130,131 A thorough intraoperative exploration of the gastrinoma triangle and pancreas is essential, but other sites (i.e., liver, stomach, small bowel, mesentery, and pelvis) should be evaluated as part of a thorough intra-abdominal evaluation to find the primary tumor, which is most often solitary and often in the duodenal wall. The duodenum and pancreatic head should be extensively mobilized and intraoperative ultrasound should be used. Intraoperative EGD with translumination may be considered. If the tumor cannot be located, longitudinal duodenotomy with inspection and palpation of the duodenal wall is performed. Lymph nodes from the portal, peripancreatic, and celiac drainage basins should be removed. Ablation or resection of hepatic metastases when identified should be considered.

The management of gastrinoma in patients with MEN I is controversial because patients are infrequently cured by operation. Acid hypersecretion in patients with gastrinoma can always be managed with high-dose PPIs. Highly selective vagotomy may make management easier in some patients and should be considered in those with surgically untreatable or unresectable gastrinoma. Gastrectomy for ZES is not indicated.

**STRESS GASTRITIS AND STRESS ULCER**

Stress gastritis is a peculiar entity that has all but disappeared from the clinical (if not endoscopic) lexicon, largely due to better critical care and acid suppression or cytoprotective agents (e.g., sucralfate) in the intensive care unit (ICU). Stress gastritis and stress ulcer are probably due to inadequate gastric mucosal
blood flow during periods of intense physiologic stress. Adequate mucosal blood flow is important to maintain the mucosal barrier and to buffer any back-diffused hydrogen ions. When blood flow is inadequate, these processes fail and mucosal breakdown occurs. Modern intensive care, with emphasis on adequate tissue perfusion and oxygenation, has undoubtedly decreased the severity of gastric mucosal injury seen in the ICU today. Although it is still common to see small mucosal erosions when performing upper endoscopy in the ICU, it is rare for these lesions to coalesce into the larger bleeding erosions that plagued the ICU patient 30 to 50 years ago. The rationale for routine acid suppression in the ICU, supported by excellent data from clinical trials and the laboratory, is that less mucosal injury will be caused in the potentially weakened gastric mucosa if there is less luminal acid.\textsuperscript{132} There are some studies suggesting that routine acid suppression leads to overgrowth of gastric bacteria, which increases the incidence and/or severity of aspiration pneumonia in the ICU.\textsuperscript{133,134} Nevertheless, acid suppression, particularly in the severely ill patient, remains an important part of clinical pathways in most ICUs.\textsuperscript{135} In the extraordinarily rare patient requiring operation today for hemorrhagic stress gastritis, the surgical options include V + D with oversewing of the major bleeding lesions, or near total gastrectomy. Angiographic embolization and endoscopic hemostatic treatment should be considered as well.

**ATROPHIC GASTRITIS**

Atrophic gastritis is characterized by atrophy or disappearance of gastric glands and loss of parietal and chief cells. The most common cause is chronic *H. pylori* infection, particularly in the corporal distribution (as opposed to the antral distribution which is more typically associated with peptic ulcer disease). Autoimmune destruction of cells (pernicious anemia) and chemical irritation (e.g., bile reflux) can also result in atrophic gastritis. Some patients with atrophic gastritis develop intestinal metaplasia in the gastric mucosa that may progress to dysplasia and then to gastric cancer. Numerous cofactors have been implicated, including diet, altered gastric microbiome, genetics, and hypergastrinemia. Patients with atrophic gastritis are at risk for

---

**Figure 26-46.** Algorithm for diagnosis and management of hypergastrinemia. BAO = basal acid output; B\textsubscript{1} = Billroth 1; B\textsubscript{2} = Billroth 2; Bx = biopsy; ECL = enterochromaffin-like; EGD = esophagogastroduodenoscopy; GJ = gastrojejunostomy; H\textsubscript{2}RA = histamine 2 receptor antagonist; insuff = insufficiency; MEN\textsubscript{1} = multiple endocrine neoplasia type I; PPI = proton pump inhibitor; R/O = rule out; SB = small bowel; S/P = status post; TV = truncal vagotomy; TV and A = truncal vagotomy and antrectomy.
gastric cancer and should undergo periodic endoscopic surveillance. Metaplastic atrophic gastritis and dysplastic atrophic gastritis in particular, are markers of increased risk for gastric cancer. Patients with high grade dysplasia may benefit from gastrectomy. Cancer risk is related to the extent of the atrophic gastritis and intestinal metaplasia, and grading systems have been developed to stratify cancer risk based on endoscopic findings. Two such systems are the operative link on gastritis assessment (OLGA) and the operative link on gastric intestinal metaplasia (OLGIM) assessment. These systems define the severity (“stage”) of atrophic gastritis based on the histologic grading of at least five gastric biopsies (lesser and greater curve antrum; lesser and greater curve corpus; angularis incisura). Since pathologists are more likely to agree on the histological diagnosis of intestinal metaplasia than they are on atrophic gastritis, the latter tool (OLGIM) may be more useful in stratifying gastric cancer risk. Patients stratified as stage 3 or 4 gastritis and those with pernicious anemia may benefit from surveillance endoscopy every 3 years. Serum markers are also useful in helping to identify patients with atrophic gastritis who usually have increased serum gastrin and iron deficiency due to parietal cell loss and hypochlorhydria or achlorhydria; decreased pepsinogen I levels due to chief cell loss; and $B_12$ deficiency due to parietal cell loss and concomitant loss of intrinsic factor.

**MALIGNANT NEOPLASMS OF THE STOMACH**

The most common primary malignant gastric neoplasm is adenocarcinoma (95%); lymphoma and GIST account for most of the remaining cases (Table 26-14). Other rare primary malignancies include neuroendocrine tumor, angiosarcoma, carcinosarcoma, and squamous cell carcinoma. Occasionally the stomach is a site of hematogenous metastasis from other sites (e.g., melanoma or breast cancer). Malignant tumors from adjacent organs may also invade the stomach by direct extension (e.g., colon or pancreas) or by peritoneal dissemination (e.g., ovary or appendiceal).

**Adenocarcinoma**

**Epidemiology.** Gastric cancer is the fourth most common cancer type and the second leading cause of cancer death worldwide. Over the past century, there has been a dramatic decrease in the incidence of gastric cancer in most Western industrialized countries (Fig. 26-49). This decrease has been largely in the so-called intestinal form rather than in the diffuse form of gastric cancer. In Asia and Eastern Europe, gastric cancer remains a leading cause of cancer death. In 2017 in the United States, approximately 28,000 new cases of stomach cancer were diagnosed (17,750 in men and 10,250 in women), and 10,960 deaths will be attributed to this disease (6720 in men and 4240 in women). The estimated 5-year survival rate is 27%, up from about 15% in 1975.

In general, gastric cancer is a disease of the elderly, and it is twice as common in blacks as in whites. In younger patients, tumors are more often of the diffuse variety and tend to be large, aggressive, and poorly differentiated, sometimes involving the entire stomach (limitis plastic). Gastric cancer has a higher incidence in groups of lower socioeconomic status.

<table>
<thead>
<tr>
<th>Types of gastric tumors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malignant tumors</td>
</tr>
<tr>
<td>Carcinoma</td>
</tr>
<tr>
<td>Lymphoma</td>
</tr>
<tr>
<td>Gastrointestinal stromal tumor (GIST)</td>
</tr>
<tr>
<td>Neuroendocrine tumor</td>
</tr>
<tr>
<td>Benign tumors</td>
</tr>
<tr>
<td>Hyperplastic polyp</td>
</tr>
<tr>
<td>Adenomatous polyp</td>
</tr>
<tr>
<td>Leiomyoma</td>
</tr>
<tr>
<td>Lipoma</td>
</tr>
<tr>
<td>Schwannoma</td>
</tr>
<tr>
<td>Heterotopic pancreas</td>
</tr>
</tbody>
</table>
Etiology. Gastric cancer is more common in patients with pernicious anemia, blood group A, or a family history of gastric cancer. When patients migrate from a high-incidence region to a low-incidence region, the risk of gastric cancer decreases in the subsequent generations born in the new region. This strongly suggests an environmental influence on the development of gastric cancer. Environmental factors appear to be more important in the pathogenesis of the intestinal form of gastric cancer compared to the diffuse form. The commonly accepted risk factors for gastric cancer are listed in Table 26-15.

Diet and Drugs A diet high in pickled, salted, or smoked food is found in many regions of high gastric cancer risk. Dietary nitrates have been implicated as a possible cause of gastric cancer. Gastric bacteria (more abundant in the achlorhydric stomach of patients with atrophic gastritis, a risk factor for gastric cancer) convert nitrate into nitrite, a known carcinogen. A diet high in fresh fruits and vegetables and rich in vitamin C and E has been shown to decrease the risk of gastric cancer. The reduced consumption of nitrate-rich preserved foods seen with the widespread availability of refrigeration has been suggested as a cause of the dramatic decrease in gastric cancer seen in North America and Western Europe over the last century.

Figure 26-49. Gastric cancer incidence and death rates per 100,000 population in men and women, in different regions and countries. (Reproduced with permission from The Global Cancer Observatory - All Rights Reserved, September, 2018.)

<table>
<thead>
<tr>
<th>Factors increasing or decreasing the risk of gastric cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Increase risk</strong></td>
</tr>
<tr>
<td>Family history</td>
</tr>
<tr>
<td>Diet (high in nitrates, salt, fat)</td>
</tr>
<tr>
<td>Familial polyposis</td>
</tr>
<tr>
<td>Gastric adenomas</td>
</tr>
<tr>
<td>Hereditary nonpolyposis colorectal cancer</td>
</tr>
<tr>
<td><em>Helicobacter pylori</em> infection</td>
</tr>
<tr>
<td>Atrophic gastritis, intestinal metaplasia, dysplasia</td>
</tr>
<tr>
<td>Previous gastrectomy or gastrojejunostomy (&gt;10 y ago)</td>
</tr>
<tr>
<td>Tobacco use</td>
</tr>
<tr>
<td>Ménétrier’s disease</td>
</tr>
<tr>
<td><strong>Decrease risk</strong></td>
</tr>
<tr>
<td>Aspirin</td>
</tr>
<tr>
<td>Diet (high fresh fruit and vegetable intake)</td>
</tr>
<tr>
<td>Vitamin C</td>
</tr>
</tbody>
</table>
Tobacco use probably increases the risk of stomach cancer, and alcohol use probably has no effect. Regular aspirin use may be protective.

*Helicobacter pylori* The risk of gastric cancer in patients with chronic *H pylori* infection is increased about threefold. Compared to uninfected patients, patients with a history of gastric ulcers are more likely to develop gastric cancer (incidence ratio 1.8, 95% confidence interval 1.6–2.0), and patients with a history of duodenal ulcers are at decreased risk for gastric cancer (incidence ratio 0.6, 95% confidence interval 0.4–0.7). This may be due to the fact that some patients develop antral-predominant disease (predisposing to duodenal ulcer and somehow protecting against gastric cancer), while other patients develop corpus-predominant gastritis, resulting in hypochlorhydria and somehow predisposing to gastric ulcer and gastric cancer (Fig. 26-50). The theoretical sequence for development of gastric adenocarcinoma is diagrammed in Fig. 26-51. Recently, it has been demonstrated that bone marrow-derived stem cells play a key role in the pathogenesis of gastric adenocarcinoma in patients with chronic *H pylori* infection. However, it must be recognized that gastric adenocarcinoma is a multifactorial disease. Not all patients with gastric cancer have *H pylori*, and there are some geographic areas with a high prevalence of chronic *H pylori* infection and a low prevalence of gastric cancer (the “African enigma”). Finally, *H pylori*-infected patients seem to be at decreased risk for the development of adenocarcinoma of the distal esophagus and cardia region. Perhaps corporeal gastritis decreases acid secretion, creating a less damaging refluxate and thus reducing the risk for Barrett’s esophagus, the precursor lesion for these tumors.

**Epstein-Barr Virus** About 10% of gastric adenocarcinomas carry the EBV virus. Recently it has been suggested that EBV infection is a late step in gastric carcinogenesis, since EBV transcripts are present in cancer cells but not in the metaplastic cells of precursor epithelium.

**Genetic Factors** A variety of genetic abnormalities have been described in gastric cancer (Table 26-16). Most gastric cancers are aneuploid. The most common genetic abnormalities in sporadic gastric cancer affect the *p53* and *COX-2* genes. Over two-thirds of gastric cancers have deletion or suppression of the important tumor-suppressor gene *p53*. Additionally, approximately the same proportion have overexpression of *COX-2*. In the colon, tumors with upregulation of this gene have suppressed apoptosis, more angiogenesis, and higher metastatic potential. Gastric tumors that overexpress *COX-2* are more aggressive. Recently, a germline mutation in the *CDH1* gene encoding E-cadherin was shown to be associated with hereditary diffuse gastric cancer. Prophylactic total gastrectomy should be considered in patients with these mutations.

**Premalignant Conditions of the Stomach** Figure 26-52 shows the prevalence of some premalignant conditions associated with the development of early gastric cancer in a series of 1900 cases from Tokyo. By far the most common precancerous lesion is atrophic gastritis. There is a growing appreciation of the important influence of the chronic inflammatory milieu on the genome of mucosal cells. Chronic inflammation leads to both genetic

---

**Figure 26-50.** *Helicobacter*, gastritis, and the pathogenesis of duodenal ulcer (DU) or gastric cancer.

**Figure 26-51.** Gastric carcinogenesis. (Reproduced with permission from Yamada T, Alpers DH, Laine L, et al: Textbook of Gastroenterology, 4th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2003.)
and epigenetic changes in mucosal cells, which in the stomach leads to the development of gastritis-associated cancer.\textsuperscript{143,144}

**Polyps** Benign gastric polyps are classified as neoplastic (adenoma and fundic gland polyps) or nonneoplastic (hyperplastic polyp, inflammatory polyp, hamartomatous polyp).\textsuperscript{145} In general, inflammatory and hamartomatous polyps have little or no malignant potential. Fundic gland polyps, commonly seen in patients on long-term PPI therapy, are not premalignant, but in patients with familial adenomatous polyposis (FAP), dysplasia in these lesions is not uncommon, and there are numerous reports of gastric cancer arising in a background of fundic gland polyps in this setting. Hyperplastic polyps usually occur in the setting of chronic inflammation. Large hyperplastic polyps (>2 cm) may harbor dysplasia or carcinoma in situ, and gastric cancer may develop remote from the hyperplastic polyp in an area of associated chronic inflammation. Gastric adenomas are premalignant. Patients with familial adenomatous polyposis (FAP) have a high prevalence of gastric adenomatous polyps (about 50%), and are 10 times more likely to develop adenocarcinoma of the stomach than the general population.\textsuperscript{146} Screening EGD is indicated in these families. Patients with hereditary nonpolyposis colorectal cancer may also be at risk for gastric cancer.\textsuperscript{147}

**Atrophic Gastritis** Chronic atrophic gastritis (Fig. 26-53) is by far the most common precursor for gastric cancer, particularly the intestinal subtype (see Fig. 26-52). The prevalence of atrophic gastritis is higher in older age groups, but it is also common in younger people in areas with a high incidence of gastric cancer. In many patients, \textit{H pylori} is critical in the pathogenesis of atrophic gastritis. Correa described three distinct patterns of chronic atrophic gastritis: autoimmune (involves the acid-secreting proximal stomach), hypersecretory (involving the distal stomach), and environmental (involving multiple random areas at the junction of the oxyntic and antral mucosa).\textsuperscript{139}

**Intestinal Metaplasia** Gastric carcinoma often occurs in an area of intestinal metaplasia, and the risk of gastric cancer is proportional to the extent of intestinal metaplasia of the gastric mucosa. These observations suggest that intestinal metaplasia is a precursor lesion to gastric cancer. There are different pathologic subtypes of intestinal metaplasia in the stomach, based upon the histologic and biochemical characteristics of the changed mucosal glands. In the complete type of intestinal metaplasia, the glands are lined with goblet cells and intestinal absorptive cells.
Slightly elevated tumors

Slightly depressed tumors

Tumors without elevation or depression relative to the surrounding mucosa are classified as Type 0-II.

Type 0-IIa

Slightly elevated tumors (superficial elevated)

Type 0-IIb

Tumors without elevation or depression (superficial flat)

Type 0-IIc

Slightly depressed tumors (superficial depressed)

Type 0-III (excavated)

Tumors with deep depression

*Tumors with less than 3-mm elevation are usually classified as Type 0-IIa, with more elevated tumors being classified as Type 0-I.
Gross Morphology and Histologic Subtypes  Gastric cancer has been subdivided into four morphologic subtypes: polypoid, fungating, ulcerative, and scirrhous. The first two are characterized by a largely intraluminal tumor. Polypoid tumors are not ulcerated; fungating tumors are predominantly intraluminal with ulceration. In the latter two gross subtypes, the bulk of the tumor mass is confined to the wall of the stomach. Ulcerative tumors are self-descriptive. Scirrhous tumors infiltrate the entire thickness of the stomach and cover a very large surface area, commonly involve the entire stomach and have a particularly poor prognosis. Although these latter lesions may be technically resectable with total gastrectomy, it is common for both the esophageal and duodenal margins of resection to show microscopic evidence of tumor infiltration; distant metastasis, overt or occult, is frequent and death from recurrent disease within 6 months is common. Palliative chemotherapy may prolong median survival.\textsuperscript{151}

The location of the primary tumor in the stomach is essential in planning an operation. Several decades ago, the large majority of gastric cancers were in the distal stomach. Recently, there has been a proximal migration of tumors, so currently, the distribution is closer to 40\% distal, 30\% middle, and 30\% proximal.

Histology  The most important prognostic indicators in gastric cancer are both histologic: lymph node involvement and depth of tumor invasion. Tumor grade (degree of differentiation: well, moderately, or poorly) is also important prognostically.

There are several histologic classifications of gastric cancer. The World Health Organization recognizes several histologic types (Table 26-18). The Japanese classification is similar but more detailed. The commonly used Lauren classification separates gastric cancers into intestinal type (53\%), diffuse type (33\%), and unclassified (14\%). The intestinal type is associated with chronic atrophic gastritis, severe intestinal metaplasia, and dysplasia, and tends to be less aggressive than the diffuse type. The diffuse type of gastric cancer is more likely to be poorly differentiated and is associated with younger patients and proximal tumors. The Ming classification also is useful and easy to remember, with only two types—expanding (67\%) and infiltrative (33\%).

Recently, the significance of human epidermal growth factor receptor-2 (HER2) was reported in patients with gastric cancer. In breast cancer, overexpression of HER2 has been reported in 15\% to 25\% of cases, and it is well recognized as an unfavorable prognostic factor. The development of molecular targeted agents such as trastuzumab has improved the survival of HER2-positive patients. Likewise, in gastric cancer, HER2 overexpression has been reported in 13\% to 30\% of patients. HER2 targeting with trastuzumab resulted in improved survival in patients with stage IV gastric cancer, and immunohistochemistry (IHC) staining for HER2 should be performed in recurrent or metastatic cases.\textsuperscript{152} Expression of other growth receptors in gastric cancer have been characterized as well, including HER1 (epidermal growth factor receptor) and HER3. The latter is associated with poor prognosis, but efforts to target these receptors for therapeutic benefit are still exploratory.\textsuperscript{153}

Pathologic Staging  Ultimately, prognosis is related to pathologic stage. The most widespread system for staging of gastric cancer is the tumor-node-metastasis (TNM) staging system based on depth of tumor invasion, extent of lymph node metastases, and presence of distant metastases. This system was developed by the American Joint Committee on Cancer and the International Union Against Cancer, and it has undergone several modifications since it was originally conceived (Table 26-19).

Clinical Manifestations.  Most patients who are diagnosed with gastric cancer in the United States have advanced stage III or IV disease at the time of diagnosis. The most common symptoms are weight loss and decreased food intake due to anorexia and early satiety. Abdominal pain (usually not severe and often ignored) is also common. Other symptoms include nausea, vomiting, and bloating. Acute GI bleeding is somewhat unusual (5\%), but chronic occult blood loss is common and manifests as iron
deficiency anemia and heme-positive stool. Dysphagia is common if the tumor involves the cardia of the stomach. Paraneoplastic syndromes such as Trousseau’s syndrome (thrombophlebitis), acanthosis nigricans (hyperpigmentation of the axilla and groin), or peripheral neuropathy are rarely present.

Physical examination typically is normal. Other than signs of weight loss, specific positive physical findings usually indicate incurability. A focused examination in a patient in whom gastric cancer is a likely part of the differential diagnosis should include an examination of the neck, chest, abdomen, and rectum. Cervical, supraclavicular (on the left referred to as Virchow’s node), and axillary lymph nodes may be enlarged, and can be sampled with fine-needle aspiration cytology. Malignant pleural effusions or ascites, or aspiration pneumonitis may be present. An abdominal mass may indicate a large (usually T4) primary tumor, liver metastases, or carcinomatosis (including Krukenberg’s tumor of the ovary). A palpable umbilical nodule (Sister Joseph’s nodule) is pathognomonic of advanced disease. Rectal exam may reveal heme-positive stool and hard nodularity extraluminally and anteriorly, indicating so-called drop metastases, or rectal shelf of Blumer in the pouch of Douglas.

**Diagnostic Evaluation.** Distinguishing between peptic ulcer and gastric cancer on clinical grounds alone can be difficult. Patients over the age of 55 years who have new-onset dyspepsia as well as all patients with dyspepsia and alarm symptoms (weight loss, recurrent vomiting, dysphagia, evidence of GI bleeding, or anemia) or with a family history of gastric cancer should undergo prompt upper endoscopy and biopsy if a mucosal lesion is noted. Essentially, all patients in whom gastric cancer is part of the differential diagnosis should have endoscopy and biopsy. If suspicion for cancer is high and the biopsy is negative, the patient should be reendoscoped and more aggressively biopsied. In some patients with gastric tumors, upper GI series can be helpful in planning treatment. Although a good double-contrast barium upper GI examination is sensitive for gastric tumors (up to 75% sensitive), in most centers, endoscopy has become the gold standard for the diagnosis of gastric malignancy. In addition, recent advances in endoscopy have contributed to the earlier diagnosis of gastric cancer. Magnifying endoscopy with narrow-band imaging (NBI) has undergone technological improvements and can observe the microvascular architecture of the mucosa and microsurface pattern of the lesion. Magnifying endoscopy with NBI has been reported to be accurate and reliable in the diagnosis of early gastric cancer.154

Preoperative staging of gastric cancer is best accomplished with abdominal/pelvic CT scanning with IV and oral contrast. MRI is probably comparable. The best way to stage the tumor locally is via EUS, which gives fairly accurate (80%) information about the depth of tumor penetration into the gastric wall, and can usually show enlarged (>5 mm) perigastric and celiac lymph nodes. However, there are limitations to tumor staging with EUS. It is highly operator dependent and may underestimate lymph node involvement because normal-sized nodes (<5 mm) can harbor metastases. EUS is most accurate in distinguishing early gastric cancer (T1) from more advanced tumors.

**Positron Emission Tomography Scanning** Whole-body PET scanning derives its power from the preferential accumulation of positron-emitting 18F-fluorodeoxyglucose in tumor compared to nontumor cells. It is most useful in the evaluation of distant metastasis in gastric cancer, but it can also be useful in locoregional staging. PET scan is accurate when combined with spiral CT (PET-CT)155 and should be considered before major surgery in patients with particularly high-risk or locally advanced tumors.

**Staging Laparoscopy and Peritoneal Cytology** Laparoscopy has emerged as a valuable adjunct to gastric cancer staging, particularly in patients with more substantial tumors. This modality allows for rapid identification of macroscopic peritoneal metastases. Peritoneal lavage identifies an additional subset of patients with microscopic dissemination. The prognostic
significance of the latter has been established by several investigators.156,157 Gastrectomy should be deferred in patients with positive peritoneal cytology without obvious peritoneal metastases. Patients with gastric cancer who undergo R0 resection (i.e., no gross residual disease) and are found to have positive peritoneal cytology (no gross carcinomatosis) have a much poorer prognosis than those with negative cytology (median survival 14.8 months vs. 98.5 months).158 Stand-alone laparoscopy may influence management in up to 36% of cases and is increasingly advocated to allow appropriate initial treatment selection. The yield is likely highest in patients with T3 or T4 tumors, proximal tumors, or evidence of regional nodal involvement158, such patients may benefit from neoadjuvant therapy, and laparoscopy should be offered prior to initiation of treatment. Systemic therapy is the cornerstone of therapy for patients with Stage IV disease and surgery is generally reserved for palliation of symptoms (e.g., an obstructing distal tumor) in patients with metastases identified during laparoscopy.

Treatment. Surgical resection is the only potentially curative treatment for gastric cancer,159 and most patients with clinically resectable locoregional disease should undergo gastrectomy. The goals of curative surgical treatment are resection of all tumor (i.e., R0 resection and adequate lymphadenectomy to afford accurate staging and provide locoregional control. Generally, the surgeon strives for a grossly negative margin of at least 5 cm, although an evidence base for this is lacking and recent retrospective analyses have suggested that more conservative resections may be adequate.160 Conversely, complete resection of diffuse tumors sometimes proves challenging, and wider gross margins guided by frozen section are sometimes appropriate. Prior to extending the resection on the basis of a positive frozen section margin, the surgeon should determine whether the microscopic tumor cells are within the wall or on the serosal. The latter may indicate incurable disseminated disease, rendering additional resection proximally or distally moot, particularly when it makes the anastomosis or stump closure more difficult or hazardous.

More than 15 resected lymph nodes are required for adequate staging, a relevant marker of quality of care.161 Thera-

apeutic nihilism should be avoided, and in the low-risk patient, an aggressive attempt to resect all tumor should be made. The primary tumor may be resected en bloc with adjacent involved organs (e.g., distal pancreas, transverse colon, or spleen) during the course of curative gastrectomy. Palliative gastrectomy may be indicated in the rare patient with incurable disease, but most patients presenting with stage IV gastric cancer can be managed without major operation.163

Extent of Gastrectomy The standard operation for gastric cancer is radical subtotal gastrectomy. Unless required for R0 resection, total gastrectomy confers no additional survival benefit and may have adverse nutritional or quality-of-life consequences and higher perioperative morbidity and mortality.164 Subtotal gastric resection typically entails ligation of the left and right gastric and gastroepiploic arteries at their origins, as well as the en bloc removal of the distal 2/3 of the stomach, including the pylorus and 2 cm of duodenum, the greater and lesser omentum, and all associated lymphatic tissue (Fig. 26-56). In the absence of involvement by direct extension, the spleen and pancreatic tail are not removed.

Reconstruction is usually by Billroth II gastrojejunostomy or Roux-en-Y gastrojejunostomy. The former is associated with shorter operative time and precludes roux limb stasis. The latter mitigates bile reflux and, therefore, may be associated with better quality of life long term. In East Asia, especially Japan, Billroth-I gastroduodenostomy is frequently performed after distal gastrectomy. Billroth-I gastroduodenostomy consists of one anastomosis (which is usually straightforward and keeps the duodenum in the food stream). In the United States, traditional surgical teaching eschews gastroduodenostomy following gastric cancer resection because of the possibility of anastomotic recurrence and obstruction. A risk of remnant carcinoma attributed to bile reflux has been invoked in support of Roux-en-Y reconstruction over Billroth II or I. Strong evidence linking reconstruction approach to long-term oncologic outcome is lacking.

Total gastrectomy with Roux-en-Y esophagojejunoanostomy may be required for R0 resection (Fig. 26-57) and is frequently the optimal operation for patients with proximal gastric adenocarcinoma. The construction of a jejunal pouch is associated with superior nutritional recovery in some but not all reports, and is a consideration.162 Proximal subtotal gastric resection,
a technically feasible alternative to total gastrectomy for some proximal gastric tumors, requires esophagogastrectomy to a denervated distal gastric remnant, and functional outcomes are generally poor.\textsuperscript{165} Pyloroplasty in this setting virtually guarantees bile esophagitis, and if the pylorus is left intact, gastric emptying may be problematic. An isoperistaltic jejunal interposition (Henley loop) between the esophagus and antrum may mitigate some of the adverse symptoms associated with this operation but adds additional complexity.

**Extent of Lymphadenectomy** The 3rd edition of Japanese classification of gastric carcinoma defines lymph node stations on the basis of anatomic landmarks. Lymph node stations 1 to 12 and 14V are classified as regional and metastasis to any other lymph nodal stations constitute distant disease (M1) (see Fig. 26-4). So-called D1 lymphadenectomy in distal gastrectomy requires the dissection of stations 1, 3, 4sb, 4d, 5, 6, and 7. Additional resection of stations 8a, 9, 11p, and 12a constitute D2 lymphadenectomy. D1 lymphadenectomy in total gastrectomy requires dissection of stations 1 through 7; D2 lymphadenectomy includes stations 8a to 12a as well. The operation most commonly performed in the United States for gastric cancer is a D1 resection and involves removal of the primary tumor with perigastric nodes. The standard operation for gastric cancer in Asia and specialized U.S. centers is D2 gastrectomy, which involves a more extensive lymphadenectomy (removal of the D1 and D2 nodes). In addition to the tissue removed in a D1 resection, D2 gastrectomy includes the superior peritoneum overlying the mesocolon and, selectively, the pancreas, as well as nodes along the common hepatic and splenic arteries, and the celiac axis. Splenectomy and distal pancreatectomy are not routinely performed because this has been shown to increase the morbidity and mortality of the operation.\textsuperscript{166,167}

The purported survival advantage of D2 gastrectomy in gastric cancer is illustrated in Table 26-20, which shows the 5-year survival rates for gastric cancer stratified by pathologic stage for the United States and Japan. Randomized prospective trials have not confirmed this survival advantage. Two studies showed increased operative mortality with D2 gastrectomy, but the most recent study did not (Table 26-21).\textsuperscript{166,168,169} With extended lymphadenectomy, much of the morbidity and mortality is attributable to performance of splenectomy and distal pancreatectomy, which are no longer routinely included as part of the D2 gastrectomy. Because D2 lymphadenectomy in total gastrectomy requires the dissection of station 10 (i.e., splenic hilar lymph nodes) splenectomy is still selectively performed, particularly for locally advanced fundic tumors.

Longer-term follow-up from the Dutch lymphadenectomy trial demonstrating a disease-specific survival advantage with D2 dissection\textsuperscript{170} as well as recognition that pancreas and spleen preserving dissection can be performed with low morbidity\textsuperscript{169} have provided momentum for increased utilization of D2 gastrectomy at high-volume centers in the United States and Europe. Some experts have argued that the D2 operation affords better staging and informs more rational decision-making regarding multimodality therapy. Without question, D2 provides a better yield of evaluable nodes. Given the frequent inadequacy of lymph node evaluation with gastrectomy in the United States and the association with poorer outcomes, greater attention to the conduct of the operation and pathologic specimen evaluation is clearly desirable. Whether better outcomes after more extensive dissection are an epiphenomenon of improved pathologic staging, or a function of therapeutic benefit, remains unclear.

**Chemotherapy and Radiation for Gastric Cancer** The actuarial 5-year survival rates for resected gastric adenocarcinoma stages I, II, and III in the United States are approximately 75%, 50%, and 25%, respectively. Because most surgical patients have stage II disease or greater, adjuvant therapy is indicated in the majority of patients who undergo initial resection. Adjuvant chemotherapy alone has not proven effective, at least in studies from Europe and the United States.\textsuperscript{171} Several studies from

<table>
<thead>
<tr>
<th>Table 26-20</th>
<th>Gastric cancer 5-year survival and operative mortality in the United States and Japan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage I</td>
<td>91%</td>
</tr>
<tr>
<td>Stage II</td>
<td>72%</td>
</tr>
<tr>
<td>Stage III</td>
<td>44%</td>
</tr>
<tr>
<td>Stage IV</td>
<td>9%</td>
</tr>
<tr>
<td>Operative mortality</td>
<td>1%</td>
</tr>
</tbody>
</table>

**Figure 26-57.** Reconstruction after total gastrectomy. Jejunal pouch (not shown here) should be considered. (Reproduced with permission from Zinner MJ: Atlas of Gastric Surgery. New York, NY: Elsevier/Churchill Livingstone; 1992.)
Japan and Korea have indicated a survival advantage with adjuvant chemotherapy after D2 gastrectomy. The discordance between outcomes in Asia and those from the United States and Europe have been attributed to differences in disease biology or treatment approaches. Proponents of the latter suggest that D2 lymphadenectomy provides sufficient locoregional control and that chemotherapy alone has efficacy after optimal surgery. A preferred adjuvant approach in the United States incorporates chemotherapy and radiation based on the results of the Intergroup trial. This prospective randomized study of adjuvant treatment with chemotherapy (5-fluorouracil and leucovorin) and radiation (4500 cGy) demonstrated a survival benefit in resected patients with stage II and III adenocarcinoma of the stomach. Only 10% of patients entered in the study actually had D2 gastrectomy, and most (54%) had less than an adequate D1 gastrectomy. Because adequacy of lymphadenectomy has been correlated with survival, particularly in patients with stage III gastric cancer, it has been suggested that the benefits of adjuvant chemoradiation shown in this study would be vitiated by a more extensive operation.

Neoadjuvant chemotherapy has emerged as a viable alternative to adjuvant chemoradiotherapy in Europe and the United States. Theoretical advantages of this approach include more consistent completion of multimodality therapy, downstaging, earlier treatment of micrometastatic disease, and the ability to gauge response at the in situ tumor. The MAGIC trial, a randomized controlled trial comparing perioperative epirubicin, cisplatin, and 5-fluorouracil to surgery alone demonstrated a survival advantage and supported this approach in patients with at least stage II disease. A subset of patients with very symptomatic tumors may not be eligible for this approach, and the perception that systemic therapy is an ineffective detour for patients who require locoregional control with surgery is sometimes hard to overcome. Regardless, neoadjuvant approaches are increasingly utilized, and even more rigorous regimens incorporating radiotherapy or targeted agents have been explored, sometimes with promising outcomes.

Recent clinical trials from Asia suggest the potential benefit of adjuvant chemotherapy after D2 lymphadenectomy in patients with advanced gastric cancer. These trials compared surgery alone and surgery plus adjuvant chemotherapy including oral fluoropyrimidines in resected advanced gastric cancer. A study from the Japan Clinical Oncology Group showed a 69% overall 5-year survival rate in patients with clinically curable T2b, T3, and T4 gastric cancer, treated with D2 gastrectomy alone. A subsequent trial from Korea demonstrated a survival advantage with adjuvant capecitabine and oxaliplatin after D2 gastrectomy compared to D2 gastrectomy alone. It is uncertain whether this approach can be translated to patients in the United States.

Although the prognosis of metastatic or recurrent gastric cancer is poor, systemic chemotherapy provides a significant survival benefit over the best supportive care. Agents that have shown activity against gastric cancer include 5-fluorouracil (5-FU), cisplatin, doxorubicin, methotrexate, taxanes, and camptothecin. Until recently, 5-FU–based chemotherapy, especially in combination with platinums, played a key role in the treatment for the unresectable gastric cancer as well as several types of cancer, such as colon and lung. In the 1990s, the introduction of novel anticancer agents such as camptothecin, taxanes, third-generation platinums, and new oral fluoropyrimidines, improved the prognosis of unresectable gastric cancer.

It is likely that targeted molecular agents will have an increasing role in treating gastric cancer. Recently, Trastuzumab, a humanized molecular antibody reactive against the extracellular domain of HER2, increased the effectiveness of cytotoxic chemotherapy in patients with HER2 over-expressing advanced gastric cancer. Other large trials are ongoing. Determination of HER2 gene amplification status may have prognostic significance.

**Endoscopic Resection**

The short- and long-term morbidity associated with gastrectomy and the relatively infrequent dissemination of superficial (i.e., T1) tumors to regional nodes have compelled exploration of endoscopic resection for selected lesions. Numerous East Asian centers have demonstrated that some patients with early gastric cancer are adequately treated with endoscopic mucosal resection (EMR). EMR is most appropriate for patients in whom the probability of lymph node metastasis is low. According to the Japanese treatment guidelines for gastric cancer, EMR is a standard treatment for well differentiated gastric cancer confined to the mucosa (T1a), measuring less than 2 cm and without signs of ulceration. Such lesions are associated with a negligible risk of lymph node metastasis.

### Table 26-21

<table>
<thead>
<tr>
<th>AUTHORS</th>
<th>NUMBER OF PATIENTS</th>
<th>TYPE OF SURGERY</th>
<th>POSTOPERATIVE COMPLICATIONS</th>
<th>POSTOPERATIVE MORTALITY</th>
<th>5 YEAR SURVIVAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bonenkamp et al.</td>
<td>711</td>
<td>D1</td>
<td>25</td>
<td>4</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D2</td>
<td>43</td>
<td>10</td>
<td>47</td>
</tr>
<tr>
<td>Cuschieri et al.</td>
<td>400</td>
<td>D1</td>
<td>28</td>
<td>6.5</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D2</td>
<td>46</td>
<td>13</td>
<td>33</td>
</tr>
<tr>
<td>Degiuli et al.</td>
<td>267</td>
<td>D1</td>
<td>12</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>D2</td>
<td>17.9</td>
<td>2.2</td>
<td></td>
</tr>
</tbody>
</table>

En bloc resection is required to evaluate margins for confirmation of complete resection. The development of endoscopic submucosal dissection (ESD) allows en bloc resection of larger tumors. This has increased the feasibility of endoscopic resection of larger lesions (<3 cm) at experienced centers. If pathologic evaluation of the resected specimen does not demonstrate ulceration, penetration of the muscularis mucosae, or lymphatic invasion, the risk of lymph node metastases is less than 1%. Even the occasional patient with higher risk stigmata may be managed endoscopically, particularly in the presence of comorbidities that preclude safe operation.

Screening for Gastric Cancer. In Japan, it clearly has been shown that patients participating in gastric cancer screening programs have a significantly decreased risk of dying from gastric cancer. Thus, screening is effective in a high-risk population. Screening the general population in the United States (a low-risk country) is probably not justified, but patients clearly at risk for gastric cancer probably should have periodic endoscopy and biopsy. This includes patients with familial adenomatous polyposis, hereditary nonpolyposis colorectal cancer, gastric adenomas, Ménétrier’s disease, intestinal metaplasia or dysplasia, and remote gastrectomy or gastrojejunostomy.

Gastric Lymphoma

Gastric lymphomas generally account for about 4% of gastric malignancies. Over half of patients with non-Hodgkin’s lymphoma have involvement of the GI tract. The stomach is the most common site of primary GI lymphoma, and over 95% are non-Hodgkin’s type. Most are B-cell type, thought to arise in mucosa associated lymphoid tissue (MALT), although most high-grade gastric lymphomas are without any characteristics of the low-grade MALT neoplasm. About half of gastric lymphomas are histologically low grade, and about half are high grade. Interestingly, the normal stomach is relatively devoid of lymphoid tissue. However, in the setting of chronic gastritis, the stomach acquires MALT, which can undergo malignant degeneration. Again, H pylori is thought to be the culprit. In populations with a high incidence of gastric lymphoma, there is a high incidence of H pylori infection; patients with gastric lymphoma also usually have H pylori infection.

Low-grade MALT lymphoma, essentially a monoclonal proliferation of B cells, presumably arises from a background of chronic gastritis associated with H pylori. These relatively innocuous tumors then undergo degeneration to high-grade lymphoma, which is the usual variety seen by the surgeon. Remarkably, when the H pylori is eradicated and the gastritis improves, the low-grade MALT lymphoma often disappears. Thus, low-grade MALT lymphoma is not a surgical lesion. Careful follow-up is necessary particularly in those lesions with a t(11:18) translocation, thought to be a risk factor for a more aggressive MALT lesion. If low-grade lymphoma persists after H pylori eradication, radiation should be considered for disease clinically confined to the stomach (stage I), while chemotherapy with or without radiation is used for more advanced lesions (Fig. 26-58).

Patients with high-grade gastric lymphoma require aggressive oncologic treatment for cure and present with many of the same symptoms as gastric cancer patients. However, systemic symptoms such as fever, weight loss, and night sweats occur in about 50% of patients with gastric lymphoma. The tumors may bleed and/or obstruct. Lymphadenopathy and/or organomegaly suggest systemic disease. Diagnosis is by endoscopy and biopsy. Much of the tumor may be submucosal, and an assiduous attempt at biopsy is necessary. Primary lymphoma is usually nodular with enlarged gastric folds. A diffusely infiltrative process akin to limits plastica is more suggestive of secondary gastric involvement by lymphoma. A diligent search for extragastric disease is necessary before the diagnosis of localized primary gastric lymphoma is made. This includes EUS; CT scanning of the chest, abdomen, and pelvis; and bone marrow biopsy. Most patients with high-grade gastric lymphoma are currently treated with chemotherapy and radiation, without surgical resection. Treatment-related perforation or bleeding is an unusual but recognized complication. For disease limited to the stomach and regional nodes, radical subtotal D2 gastrectomy may be performed, especially for bulky tumors with bleeding and/or obstruction. Palliative gastrectomy for tumor complications also has a role. Certainly, a multidisciplinary team should be involved in managing patients with primary gastric lymphoma.

Gastrointestinal Stromal Tumor

GISTs arise from interstitial cells of Cajal (ICC) and are distinct from leiomyoma and leiomyosarcoma, which arise from smooth muscle. Prognosis in patients with GIST tumors depends on tumor size, location, and mitotic count. Metastasis, when it occurs, is typically by the hematogenous route. Virtually all GISTs should be resected along with a margin of normal tissue. Most GISTs (and almost no smooth muscle tumors) express c-KIT (CD117) or the related PDGF receptor A, as well as CD34; almost all smooth muscle tumors (and almost no GISTs) express actin and desmin. These markers can often be detected on specimens obtained by fine-needle aspiration and are useful in differentiating between GIST and smooth muscle tumor histopathologically. Lesions that are definitively leiomyoma by histopathologic criteria can be observed if small and asymptomatic. Larger or symptomatic gastric leiomyomas are adequately treated by enucleation or wedge resection. Lesions that are definitively GIST or leiomyosarcoma are best treated by resection with negative margins. Most equivocal lesions should be resected provided that the patient has a reasonable operative risk.

Two-thirds of all GISTs occur in the stomach and have a more favorable prognosis than do GISTs occurring in other locations. Epithelial cell stromal GIST is the most common cell type arising in the stomach, and cellular spindle type is the next most common. The glomus tumor type is seen only in the stomach. Smaller lesions are usually found incidentally, although they occasionally may ulcerate and cause bleeding. Larger lesions may produce symptoms of weight loss, abdominal pain, fullness, early satiety, and bleeding. An abdominal mass may be palpable. Metastasis is by the hematogenous route, most often to liver.

Diagnosis is by endoscopy and biopsy, although the interpretation of the latter may be problematic. When performed, a transmural (i.e., endoscopic) approach to biopsy is preferred to a percutaneous one, to avoid the potential for fragmentation and peritoneal seeding. A nondiagnostic biopsy does not preclude resection of a suspicious appearing lesion. Metastatic workup entails CT of the abdomen, and pelvis (chest X-ray sufficient in lieu of CT of the chest for most patients). Most GISTs are solitary. Local resection with clear margins is adequate surgical treatment but is sometimes impractical for larger prepyloric or pyloric channel tumors, or those near the GE junction.
True invasion of adjacent structures by the primary tumor is occasionally seen with larger more aggressive lesions. If safe, en bloc resection of involved surrounding organs is appropriate to remove all tumor.

The risk of tumor recurrence or metastasis behavior has been stratified into four groups according to the tumor size and mitotic count. Very low risk is defined by size <2 cm and <5 mitoses/50 HPF (high-power field). Low risk is defined by size 2 to 5 cm and <5 mitoses/50 cm. Intermediate risk is defined by size <5 cm and 6 to 10 mitoses/50 HPF or size 5 to 10 cm and >5 mitoses/50 HPF. High risk is defined by size >5 cm and >5 mitoses/50 HPF, size >10 cm regardless of mitotic rate or >10 mitoses/50 HPF regardless of size. As mentioned, stomach lesions are associated with lower risk than are tumors in other locations. Classification based on tumor location, size, and mitotic rate have been proposed to evaluate the risk of recurrence and metastasis and role for adjuvant therapy. In an effort to further refine risk stratification, a number of nomograms have been introduced incorporating tumor features. These tools draw upon and are subject to the limitations of the institutional data from which they are derived, but they are sometimes helpful in counseling patients.

Mutations in the oncodriver c-kit and PDGFRA are present in a majority of GISTs. This has been exploited through the use of imatinib (Gleevec), a tyrosine kinase inhibitor. Several clinical trials in a metastatic disease setting demonstrated marked improvements in median survival from 9 months to greater than 5 years. These striking results not only established imatinib as the primary therapy for metastatic GIST, but they also compelled broader efforts to target solid tumors with small molecule inhibitors. Notably, up to 50% of treated patients develop resistance to imatinib by 2 years, and several
second-line agents have been utilized for patients with refractory disease, most notable sunitinib.

The efficacy of imatinib as adjuvant therapy for high risk GIST has been demonstrated in two randomized clinical trials, ACOSOG Z9001 and SSG XVIII.190,191 The former trial randomized patients to 1 year of adjuvant imatinib or placebo and showed an improvement in recurrence-free survival with imatinib. The latter trial demonstrated an overall survival advantage with 3 years compared to 1 year of therapy. Imatinib is now recommended in high risk groups as an adjuvant therapy, for three years or longer. Preoperative therapy with imatinib may be indicated in selected patients with larger lesions that may be more difficult to completely resect or require multivisceral resection.

Molecular profiling has been embraced with growing recognition that specific tumor subtypes are insensitive to imatinib. Patients with PDGFRA D842V mutations, for example, do not respond to imatinib.192 Management of metastatic GIST is principally medical, but surgery has a selected role. An algorithm for the treatment of patients with metastatic GIST is shown in Fig. 26-59.

**Gastric Neuroendocrine Tumors**

Compared to neuroendocrine tumors of the midgut and hindgut, neuroendocrine tumors of the stomach are rare. Gastric neuroendocrine tumors comprise about 1% of all neuroendocrine tumors and less than 2% of gastric neoplasms. They arise from gastric enterochromaffin-like (ECL) cells and may have malignant potential. The apparent incidence of gastric neuroendocrine tumors is increasing, perhaps related to increased detection or the increasing use of acid suppressive medication. The latter may cause hypergastrinemia, and gastrin has a recognized trophic effect on gastric ECL cells.

Nomenclature remains a point of confusion; carcinoid and well-differentiated neuroendocrine tumor (NET) are synonymous according to WHO classification.

Gastric neuroendocrine tumors are classified into one of three different types. Type I is the most common, accounting for about 75% of cases. Type I lesions occur in patients with chronic hypergastrinemia secondary to pernicious anemia or atrophic gastritis. These lesions occur more frequently in women, are often multiple and small, and have low malignant potential (<5% metastasize). The role of long-term acid suppression with resultant hypergastrinemia in the pathogenesis of type I gastric carcinoids is unclear. Type II gastric neuroendocrine tumors are associated with MEN1 and ZES. These lesions also tend to be small and multiple, but they have a somewhat higher malignant potential than type I lesions (10% metastasize). Type II lesions are more common in the setting of MEN1; they are quite uncommon in patients with sporadic ZES. The constellation of gastric acidity, hypergastrinemia, and gastric neuroendocrine tumors suggests gastrinoma until proven otherwise. Type III gastric neuroendocrine tumors are sporadic. They are most often solitary (usually >2 cm) and occur more commonly in men. They are not associated with hypergastrinemia. Most patients have regional nodal or distant metastases at the time of diagnosis, and some present with symptoms of carcinoid syndrome.

Gastric neuroendocrine tumors are usually diagnosed with endoscopy and biopsy. The type can be determined based upon clinical context, patient history, the presence or absence of atrophic gastric mucosa, gastric pH and gastrin level. Some tumors are submucosal and may be quite small. They are often confused with heterotopic pancreas or small leiomyomas. Biopsy may be difficult because of the submucosal location, and EUS can be helpful in defining the size and depth of the lesion. Plasma

---

**Figure 26-59.** Algorithm for the treatment of malignant gastrointestinal stromal tumor. (Reproduced with permission from Balachandran VP, DeMatteo RP: Gastrointestinal stromal tumors: who should get imatinib and for how long? Adv Surg. 2014;48:165-183.)
chromogranin A levels are frequently elevated. CT scan and octreotide or gallium dotatate scans are useful for staging.

Type I and II patients with numerous diminutive lesions can be followed with serial endoscopy. Small lesions confined to the mucosa (typically type I or type II lesions) less than 1 cm may be treated endoscopically with EMR if there are only a few lesions (<5). Occasionally a slightly larger lesion (1–2 cm) necessitates local surgical excision. Larger lesions and type III lesions should be removed by D1 or D2 gastrectomy. Antrectomy to mitigate gastric secretion in type I patients with refractory growing lesions was invoked as a viable treatment strategy in the past, but it is rarely indicated.

Survival is excellent for node-negative patients (>90% 5-year survival); node-positive patients have a 50% 5-year survival. Gastrinoma should be resected if located in patients with type II carcinoid. The 5-year survival for patients with type I gastric carcinoid is close to 100%; for patients with type III lesions, the 5-year survival is less than 50%. Somatostatin analogue therapy may delay progression of metastatic disease. Surgical debulking may have a role in selected patients with limited metastatic disease.

**BENIGN GASTRIC NEOPLASMS**

**Leiomyoma**

The typical leiomyoma is submucosal and firm. If ulcerated, it has an umbilicated appearance and may bleed. Histologically, these lesions appear to be of smooth muscle origin. Lesions <2 cm are usually asymptomatic and benign. Larger lesions may cause symptoms such as bleeding, obstruction, or pain. Asymptomatic lesions <2 cm may be carefully observed or enucleated if fine-needle aspiration and immune markers confirm smooth muscle tumor; larger lesions and symptomatic lesions should be removed by wedge resection (often possible laparoscopically). When lesions thought to be leiomyoma are observed rather than resected, the patient should be made aware of their presence and the small possibility for malignancy.

**Lipoma**

Lipomas are benign submucosal fatty tumors that are usually asymptomatic, found incidentally on upper GI series or EGD. Endoscopically, they have a characteristic appearance; there also is a characteristic appearance on EUS. Excision is unnecessary unless the patient is symptomatic.

**Gastroparesis**

Gastric motility disorders include delayed gastric emptying (gastroparesis), rapid gastric emptying, and motor and sensory abnormalities (e.g., functional dyspepsia). Surgically relevant secondary disorders of gastric motility (e.g., dumping, gastric stasis, and Roux syndrome) are discussed under Postgastrectomy Problems. Gastroparesis is the most surgically relevant primary disorder of gastric motility.195,196

Most patients with primary gastroparesis present with nausea, vomiting, bloating, early satiety, and/or abdominal pain. Eighty percent of these patients are women; some are diabetic. Postprandial vomiting significantly complicates the management of blood glucose in the latter group, predisposing to hypoglycemia following preprandial insulin. In patients with gastroparesis, it is important to rule out mechanical gastric outlet obstruction, and small-bowel obstruction. An upper GI series may suggest slow gastric emptying and relative atony, or it may be normal. EGD may show bezoars or retained food but is frequently normal. Gastric emptying scintigraphy shows delayed solid emptying, and often delayed liquid emptying. Gastroparesis can be a manifestation of a variety of problems (Table 26-22). Medical treatment includes promotility agents, antiemetics, and, perhaps, botulinum injection into the pylorus.

If the diabetic gastroparetic patient is not a candidate for pancreas transplant, both gastrostomy (for decompression) and jejunostomy tubes (for feeding and prevention of hypoglycemia) can be helpful in managing these patients. Other surgical options include implantation of a gastric pacemaker, pyloroplasty or peroral endoscopic pyloromyotomy (particularly in patients responsive to pyloric Botox injection), and gastric resection.197 Generally, gastric resection should be done only after other therapeutic options have been exhausted.

### Table 26-22

<table>
<thead>
<tr>
<th>Etiology of gastroparesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idiopathic</td>
</tr>
<tr>
<td>Endocrine or metabolic</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
</tr>
<tr>
<td>Thyroid disease</td>
</tr>
<tr>
<td>Renal insufficiency</td>
</tr>
<tr>
<td>After gastric surgery</td>
</tr>
<tr>
<td>After resection</td>
</tr>
<tr>
<td>After vagotomy</td>
</tr>
<tr>
<td>Central nervous system disorders</td>
</tr>
<tr>
<td>Brain stem lesions</td>
</tr>
<tr>
<td>Parkinson’s disease</td>
</tr>
<tr>
<td>Peripheral neuromuscular disorders</td>
</tr>
<tr>
<td>Myotonia dystrophica</td>
</tr>
<tr>
<td>Duchenne muscular dystrophy</td>
</tr>
<tr>
<td>Connective tissue disorders</td>
</tr>
<tr>
<td>Scleroderma</td>
</tr>
<tr>
<td>Polymyositis/dermatomyositis</td>
</tr>
<tr>
<td>Infiltrative disorders</td>
</tr>
<tr>
<td>Lymphoma</td>
</tr>
<tr>
<td>Amyloidosis</td>
</tr>
<tr>
<td>Diffuse gastrointestinal motility disorder</td>
</tr>
<tr>
<td>Chronic intestinal pseudo-obstruction</td>
</tr>
<tr>
<td>Medication-induced</td>
</tr>
<tr>
<td>Electrolyte imbalance</td>
</tr>
<tr>
<td>Potassium, calcium, magnesium</td>
</tr>
<tr>
<td>Miscellaneous conditions</td>
</tr>
<tr>
<td>Infections (especially viral)</td>
</tr>
<tr>
<td>Paraneoplastic syndrome</td>
</tr>
<tr>
<td>Ischemic conditions</td>
</tr>
<tr>
<td>Gastric ulcer</td>
</tr>
</tbody>
</table>


---

1152
The most common causes of acute upper GI bleeding in emergency department or hospitalized patients are peptic ulcer, gastritis, Mallory-Weiss syndrome, and esophagogastric varices. Less common causes include benign or malignant neoplasm, angiodyplasia, Dieulafoy’s lesion, portal gastropathy, Ménétrier’s disease, and watermelon stomach. Arterioenteric fistula should always be considered in the patient who has an aortic graft or who has undergone repair of a visceral artery aneurysm.

The most important issues in the early hospital management of patients with acute upper GI bleeding are resuscitation and risk stratification. Large-bore IV access and Foley catheterization is accomplished, and nasogastric intubation is considered. Risk stratification is essentially accomplished by answering the following questions:

a. What is the magnitude and acuity of the hemorrhage? Hypotension, tachycardia, oliguria, low hematocrit, pallor, altered mentation, and/or hematemesis suggest a large blood loss that has occurred over a short period of time. This is a high-risk situation.
b. Does the patient have significant chronic disease, particularly lung, liver, kidney, and/or heart disease, which compromises physiologic reserve? If yes, this is a high-risk situation.
c. Is the patient anticoagulated, or immunosuppressed? If yes, this is a high-risk situation.
d. On endoscopy, is the patient bleeding from varices, or is there active bleeding, or is there a visible vessel, or is there a deep ulcer overlying a large vessel (e.g., posterior duodenal ulcer overlying the gastroduodenal artery)? Could the patient be bleeding from an arterio-enteric fistula? If yes, this is a high-risk situation.

When judged to be low risk, most patients will stop bleeding with supportive treatment and IV PPI. Selected patients may be discharged from the emergency department and managed on an outpatient basis.

If the patient is deemed to be high risk based on one or more of the aforementioned questions, then the following should be done immediately:

1. Type and cross-match for transfusion of blood products.
2. Admit to ICU or monitored bed in specialized unit.
3. Consult surgeon.
4. Consult gastroenterologist.
5. Start intravenous PPI.
6. Perform upper endoscopy within 12 hours, after resuscitation and correction of coagulopathy. Endoscopic hemostasis should be considered in most high-risk patients with acute upper GI bleeding.

Although the surgeon should be involved early in the hospital course of all high-risk patients with acute upper GI bleeding, most of these patients will be adequately managed without operation. Mucosal lesions can usually be controlled with endoscopic hemotherapy and medical management. Occasionally, arteriography can be helpful. Operation for bleeding ulcer is discussed previously (see “Operation for Bleeding Peptic Ulcer” and Fig. 26-43).

Isolated Gastric Varices
Isolated gastric varices are those that occur in the absence of esophageal varices and are classified as type I (fundic) or type II (distal to fundus including proximal duodenum). The presence of isolated gastric varices is usually associated with portal hyper-tension or splenic vein thrombosis. Although there is a significant bleeding risk from isolated gastric varices on long-term follow-up, there is no indication for the routine application of prophylactic measures.

Patients with acute upper GI bleeding from isolated gastric varices should be considered high risk. Although data are limited, octreotide and/or vasopressin infusion may decrease bleeding, if tolerated. Balloon tamponade with a Sengstaken-Blakemore tube may provide temporary control of exsanguinating hemorrhage from type isolated gastric varices, but if this is used, endotracheal intubation for airway protection is prudent. Endoscopic treatment with sclerotherapy or varix ligation is less successful than in esophageal varices but should be considered. Interventional radiology should be consulted and balloon-occluded retrograde transvenous obliteration considered. A transjugular intrahepatic portosystemic shunt (TIPSS) may be useful if there is nonsegmental portal hypertension. If the patient has splenic vein thrombosis and left-sided (sinistral) or segmental portal hypertension, splenectomy is quite effective in controlling bleeding from isolated gastric varices. The operative mortality is 5%. Liver transplantation should always be considered in the cirrhotic patient.

Hypertrophic Gastropathy (Ménétrier’s Disease)
There are two clinical syndromes characterized by epithelial hyperplasia and giant gastric folds: ZES and Ménétrier’s disease. The latter is characteristically associated with protein-losing gastropathy and hypochlorhydria. There are large rugal folds in the proximal stomach, and the antrum is usually spared. Mucosal biopsy shows diffuse hyperplasia of the surface mucus-secreting cells and usually decreased parietal cells (Fig. 26-60). It has recently been suggested that Ménétrier’s disease is caused by local overexpression of transforming growth factor-α in the gastric mucosa, which stimulates the epidermal growth factor receptor, a receptor tyrosine kinase, on gastric SECs. This results in the selective expansion of surface mucus cells in the gastric body and fundus. A few patients with this unusual disease have been successfully

Figure 26-60. Mucosal biopsy in Ménétrier’s disease. (Reproduced with permission from Ming S-C, Goldman H: Pathology of the Gastrointestinal Tract, 2nd ed. Baltimore, MD: Williams & Wilkins; 1998.)
treated with the epidermal growth factor receptor blocking monoclonal antibody cetuximab.201

Most patients with Ménétrier’s disease are middle-aged men who present with epigastric pain, weight loss, diarrhea, and hypoproteinemia. There may be an increased risk of gastric cancer. Sometimes, the disease regresses spontaneously. Occasionally it is associated with *H pylori* infection, and the disease improves with helicobacter eradication. Total gastrectomy may be indicated for bleeding, severe hypoproteinemia, or cancer.

**Watermelon Stomach (Gastric Antral Vascular Ectasia)**

The parallel red stripes atop the mucosal folds of the distal stomach give this rare entity its sobriquet. Histologically, gastric antral vascular ectasia (GAVE) is characterized by dilated mucosal blood vessels that often contain thrombi, in the lamina propria. Mucosal fibromuscular hyperplasia and hyalinization often are present (Fig. 26-61). The histologic appearance can resemble portal hypertensive gastropathy, but the latter usually affects the proximal stomach, whereas watermelon stomach predominantly affects the distal stomach. β-Blockers and nitrates, useful in the treatment of portal hypertensive gastropathy, are ineffective in patients with gastric antral vascular ectasia. Patients with GAVE are usually elderly women with chronic GI blood loss requiring transfusion. Most have an associated autoimmune connective tissue disorder, and at least 25% have chronic liver disease. Nonsurgical treatment options include estrogen and progesterone, and endoscopic treatment with the neodymium yttrium-aluminum garnet (Nd:YAG) laser or argon plasma coagulator.202 Antrectomy may be required to control blood loss, and this operation is quite effective but carries increased morbidity in this elderly patient group. Patients with portal hypertension and antral vascular ectasia should be considered for transjugular intrahepatic portosystemic shunt (TIPSS).

**Dieulafoy’s Lesion**

Dieulafoy’s lesion is a congenital arteriovenous malformation characterized by an unusually large tortuous submucosal artery. If this artery is eroded, impressive pulsatile bleeding may occur. To the endoscopist or surgeon, this appears as a stream of arterial blood emanating from what appears grossly to be a normal gastric mucosa. The lesion typically occurs in middle-aged or elderly men and may be more common in patients with liver disease.203 Patients typically present with upper GI bleeding, which may be intermittent, and endoscopy can miss the lesion if it is not actively bleeding. Treatment options include endoscopic hemostatic therapy, angiographic embolization, or operation. At surgery, the lesion may be oversewn or resected.

**Bezoars/Diverticula**

Bezoars are concretions of indigestible matter that accumulate in the stomach. Trichobezoars are composed of swallowed hair (Fig. 26-62). Phytobezoars are composed of vegetable matter and, in the United States, are usually seen in association with gastroparesis or gastric outlet obstruction. They also are associated with persimmon ingestion. Most commonly, bezoars produce obstructive symptoms, but they may cause ulceration and bleeding. Perforation is rare. Asymptomatic diverticula do not require treatment, but symptomatic lesions should be removed. This can often be done laparoscopically.
large objects in the stomach should be removed. This can usually be done endoscopically, with an overtube technique. Recognized dangers include aspiration of the foreign body during removal and rupture of drug-containing bags in “body packers.” Both complications can be fatal. Surgical removal is recommended in body packers who ingest drug parcels for smuggling and in patients with large jagged objects that cannot be safely removed endoscopically. Corrosive objects (i.e., batteries) should be removed promptly usually endoscopically. Ingested magnets should be removed unless they are small and singular and without other ingested metal objects.

Mallory-Weiss Syndrome

The Mallory-Weiss lesion is a longitudinal tear in the mucosa of the GE junction. It is presumably caused by forceful vomiting and/or retching, and it is commonly seen in alcoholics. It presents with upper GI bleeding, often with hematemesis. Endoscopy confirms the diagnosis and may be useful in controlling the bleeding, but 90% of patients stop bleeding spontaneously. Other options to control the bleeding include balloon tamponade, angiographic embolization, or selective infusion of vasopressin, systemic vasopressin, and operation. Surgical treatment consists of oversewing the bleeding lesion through a long gastrotomy.

Volvulus

Gastric volvulus is a twist of the stomach that usually occurs in association with a large hiatal hernia. It also can occur in patients with an unusually mobile stomach without hiatal hernia. Typically, the stomach twists along its long axis (organoaxial volvulus), and the greater curvature flips up (Fig. 26-64C). If the stomach twists around the transverse axis, it is called mesentero-axial rotation (Fig. 26-64A and Fig. 26-64B). Often, volvulus is a chronic condition that can be surprisingly asymptomatic. In these instances, expectant nonoperative management is typically advised, especially in the elderly. The risk of strangulation and infarction has been overestimated in asymptomatic patients. Symptomatic patients should be considered for operation, especially if the symptoms are severe and/or progressive. Patients may present with symptoms of pain and pressure related to the intermittently distending and poorly emptying twisted stomach. Pressure on the lung may produce dyspnea, pressure on the pericardium may produce palpitations, and pressure on the esophagus may produce dysphagia. Symptoms are often relieved with vomiting or passage of a nasogastric tube. Gastric infarction is a surgical emergency, and the patient can be moribund. Gastric necrosis may be extensive or focal. Elective operation for gastric volvulus usually involves reduction of the stomach and gastropexy with or without repair of hiatal hernia. Gastropexy alone should be considered for high-risk patients since it can nearly always be performed laparoscopically and may be surprisingly effective in relieving mechanical symptoms.

GASTROSTOMY

A gastrostomy is performed either for alimentation or for gastric drainage/decompression. Gastrostomy may be done percutaneously, laparoscopically, or via open technique. Currently, percutaneous endoscopic gastrostomy is the most common method used. The open techniques include the Stamm method (Fig. 26-65), the Witzel method (Fig. 26-66), and the Janeway method. The Janeway gastrostomy, designed to create a permanent nondraining gastric stoma that can be intermittently intubated, is more complicated than the other open techniques, and is rarely necessary. By far the most common surgical technique is the Stamm gastrostomy, which can be performed open or laparoscopically.

Complications of gastrostomy include infection, dislodgment, leakage with peritonitis, and aspiration pneumonia. Although gastrostomy tubes usually do prevent tense gastric dilatation, they may not adequately drain the stomach, especially when the patient is bedridden, and they cannot always be relied upon to prevent pulmonary aspiration of gastric contents.
Dumping Syndrome

Dumping is a phenomenon caused by the destruction or bypass of the pyloric sphincter. However, other factors undoubtedly play a role because dumping can occur after operations that preserve the pylorus, such as parietal cell vagotomy. Also, an appropriate stimulus may provoke dumping symptoms, even in some patients who have not undergone surgery. Clinically significant dumping occurs in 5% to 10% of patients after pyloroplasty, pyloromyotomy, or gastrectomy, and consists of a constellation of postprandial symptoms ranging in severity from annoying to disabling. The symptoms are thought to be the result of the abrupt delivery of a hyperosmolar load into the small bowel due to ablation of the pylorus or decreased gastric compliance. Typically, 15 to 30 minutes after a meal, the patient becomes diaphoretic, weak, light-headed, and tachycardic. These symptoms may be ameliorated by recumbence or saline infusion. Crampy abdominal pain is not uncommon, and diarrhea often follows. This is referred to as early dumping and should be distinguished from postprandial (reactive) hypoglycemia, also called late dumping, which usually occurs later (2–3 hours following a meal) and is relieved by the administration of sugar. A variety of treatments have been advocated, including dietary modifications, anticholinergics, and epinephrine administration.
of hormonal aberrations have been observed in early dumping, including increased serum levels of VIP, CCK, neuropeptide, peripheral hormone peptide YY, renin-angiotensin-aldosterone, and decreased atrial natriuretic peptide. Late dumping is associated with hypoglycemia and hyperinsulinemia.

Medical therapy for the dumping syndrome consists of dietary modification and somatostatin analogue (octreotide). Often, symptoms improve if the patient avoids liquids during meals. Hyperosmolar liquids (e.g., milk shakes) may be particularly troublesome. There is some evidence that adding dietary fiber compounds at mealtime may improve the syndrome. If dietary manipulation fails, the patient is started on octreotide, 100 μg subcutaneously twice daily. This can be increased up to 500 μg twice daily if necessary. The long-acting depot octreotide preparation is useful. Octreotide not only ameliorates the abnormal hormonal pattern seen in patients with dumping symptoms, but it also promotes restoration of a fasting motility pattern in the small intestine (i.e., restoration of the MMC). The α-glucosidase inhibitor acarbose may be particularly helpful in ameliorating the symptoms of late dumping.

Only a very small percentage of patients with dumping symptoms ultimately require surgery. Most patients improve with time (months and even years), dietary management, and medication. Therefore, the surgeon should not rush to reoperate on the patient with dumping symptoms. Multidisciplinary nonsurgical management must be optimized first. Before reoperation, a period of inhospital observation is useful to define the severity of the patient’s symptoms and patient compliance with prescribed dietary and medical therapy.

The results of remedial operation for dumping are variable and unpredictable. There are a variety of surgical approaches, none of which work consistently well. Additionally, there is not a great deal of experience reported in the literature with any of these methods and long-term follow-up is rare. Patients with disabling refractory dumping after gastrojejunostomy can be considered for simple takedown of this anastomosis provided that the pyloric channel is patent. The reversed intestinal segment is rarely used today—and rightly so. This operation interposes a 10-cm reversed segment of intestine between the stomach and the proximal small bowel. This slows gastric emptying, but often leads to obstruction, requiring reoperation. Isoperistaltic interposition (Henley loop) has not been successful in ameliorating severe dumping over the long term. The Roux-en-Y gastrojejunostomy is associated with delayed gastric emptying, probably on the basis of disordered motility in the Roux limb. Taking advantage of this disordered physiology, surgeons have used this operation successfully in the management of the dumping syndrome. Although this is probably the procedure of choice in the small group of patients requiring operation for severe dumping following gastric resection, gastric stasis may result, particularly if a large gastric remnant is left. In the presence of significant gastric acid secretion, marginal ulceration is common after both jejunal interposition and Roux-en-Y procedures; thus, concomitant vagotomy and hemigastrectomy should be considered. The theoretical possibility of treating postpyloroplasty dumping with a Roux-en-Y to the proximal duodenum (the duodenal switch, a potentially reversible operation) has not yet been reported (Fig. 26-67). Because pyloric ablation seems to be the dominant factor in the etiology of dumping, it is not surprising that conversion of Billroth II to Billroth I anastomosis has not been successful in the treatment of dumping.

![Figure 26-67. Duodenal switch operation. (Reproduced with permission from Hinder RA: Duodenal switch: a new form of pancreatocobiliary diversion, Surg Clin North Am. 1992 Apr;72(2):487-499.)](image)

**Diarrhea**

Diarrhea following gastric surgery may be the result of truncal vagotomy, dumping, or malabsorption. Truncal vagotomy is associated with clinically significant diarrhea in 5% to 10% of patients. It occurs soon after surgery and usually is not associated with other symptoms, a fact that helps to distinguish it from dumping. The diarrhea may be a daily occurrence, or there may be significant periods of relatively normal bowel function. The symptoms tend to improve over the months and years after the index operation. The cause of postvagotomy diarrhea is uncertain. Possible mechanisms include intestinal dysmotility and accelerated transit, bile acid malabsorption, rapid gastric emptying, and bacterial overgrowth. The latter problem is facilitated by decreased gastric acid secretion and (even small) blind loops. Although bacterial overgrowth can be confirmed with the hydrogen breath test, a simpler test is an empirical trial of oral antibiotics. Some patients with postvagotomy diarrhea respond to cholestyramine, while in others codeine or loperamide may be useful. Octreotide should also be tried. Another theoretical cause of diarrhea following gastric surgery is fat malabsorption due to acid inactivation of pancreatic enzymes or poorly coordinated mixing of food and digestive juices. This can be confirmed with a qualitative test for fecal fat and treated with acid suppression. Postvagotomy diarrhea usually does not respond to treatment with pancreatic enzymes. In the rare patient who is debilitated by postvagotomy diarrhea unresponsive to medical management, operation might be considered, but outcomes can be problematic. The operation of choice is probably a 10-cm reversed jejunal interposition placed in continuity 100 cm distal to the ligament of Treitz. Another option is the onlay antiperistaltic distal ileal graft. Both operations can cause obstructive symptoms and/or bacterial overgrowth.

**Gastric Stasis**

Gastric stasis following surgery on the stomach may be due to a problem with gastric motor function or caused by an obstruction. The gastric motility abnormality could have been preexisting and unrecognized by the operating surgeon. Alternatively,
it may be secondary to deliberate or unintentional vagotomy, or resection of the dominant gastric pacemaker. An obstruction may be mechanical (e.g., anastomotic stricture, efferent limb kink from adhesions or constricting mesocolon, or a proximal small-bowel obstruction) or functional (e.g., retrograde peristalsis in a Roux limb). Gastric stasis presents with vomiting (often of undigested food), bloating, epigastric pain, and weight loss.

The evaluation of a patient with suspected postoperative gastric stasis includes EGD, upper GI and small bowel series, gastric emptying scan, and gastric motor testing. Endoscopy shows gastritis and retained food or bezoar. The anastomosis and efferent limb should be evaluated for stricture or narrowing. A dilated efferent limb suggests chronic stasis, either from a motor abnormality (e.g., Roux syndrome) or mechanical small bowel obstruction (e.g., chronic adhesion). If the problem is thought to be primarily a disorder of intrinsic motor function, newer techniques such as EGG and GI manometry should be considered, but chronic distal mechanical obstruction may result in disordered motility in the proximal organ confounding interpretation.

Once mechanical obstruction has been ruled out, medical treatment is successful in most cases of motor dysfunction following previous gastric surgery. This consists of dietary modification and promotility agents. Intermittent oral antibiotic therapy may be helpful in treating bacterial overgrowth, with its attendant symptoms of bloating, flatulence, and diarrhea.

Gastroparesis following V + D may be treated with subtotal gastrectomy but simple loop gastrojejunostomy (GJ) should be tried if previous drainage was pyloroplasty. Billroth II anastomosis with Braun enteroenterostomy may be preferable to Roux-en-Y reconstruction after subtotal gastrectomy for gastric stasis, but bile reflux can still occur. Initial operation for gastric stasis is often associated with persistent gastric emptying problems that may subsequently require near-total or total gastrectomy, a nutritionally unattractive option. Delayed gastric emptying following ulcer surgery (V + D or V + A) may represent an anastomotic stricture (often due to recurrent ulcer) or proximal small bowel obstruction. Recurrent ulcer may respond to medical therapy with PPI and abstinence from NSAIDs, aspirin, and smoking. And if necessary, endoscopic dilation is optionally helpful. However, when associated with symptomatic gastric stasis, reoperation is often necessary. Gastroparesis following subtotal gastric resection is best treated with near-total (95%) or total gastric resection and Roux-en-Y reconstruction. If total gastrectomy is performed, a jejunal reservoir should be considered. Gastric pacing is promising, but it has not achieved widespread clinical usefulness in the treatment of postoperative gastric atony.

**Bile Reflux Gastritis and Esophagitis**

Most patients who have undergone ablation or resection of the pylorus have bile in the stomach on endoscopic examination, along with some degree of gross or microscopic gastric inflammation. Therefore, attributing postoperative symptoms to bile reflux is problematic because most asymptomatic patients have bile reflux too. However, it is generally accepted that a small subset of patients have bile reflux gastritis syndrome. These patients present with nausea, bilious vomiting, and epigastric pain, and quantitative evidence of excess enterogastric reflux. Curiously, symptoms often develop months or years after the index operation. The differential diagnosis includes afferent or efferent loop obstruction, gastric stasis, and small-bowel obstruction. Plain abdominal X-rays, upper endoscopy, upper GI series, abdominal CT scan, and gastric emptying scans are helpful in evaluating these possibilities.

Bile reflux may be quantified with gastric analysis or esophageal impedance testing or with scintigraphy (bile reflux scan). Typically, enterogastric reflux is greatest after Billroth II gastrectomy or gastrojejunostomy, and least after vagotomy and pyloroplasty, with Billroth I gastrectomy giving intermediate values. Patients who are well into the abnormal range of bile reflux may be considered for remedial surgery if symptoms are severe. Remedial surgery will eliminate the bile from the vomitus and may improve the patient’s pain, but it is quite unusual to render these patients completely asymptomatic, especially if they are narcotic dependent.

Bile reflux gastritis after distal gastric resection may be treated by one of the following options: Roux-en-Y gastrojejunostomy; interposition of a 40-cm isoperistaltic jejunal loop between the gastric remnant and the duodenum (Henley loop); Billroth II gastrojejunostomy with Braun enterenterostomy; total gastrectomy with Roux esopha.gojejunostomy. To minimize reflux of bile into the stomach or the esophagus, the Roux limb should be at least 45 cm long (preferably 60 cm). The Braun enterenterostomy should be placed at a similar distance from the stomach. Excessively long limbs may be associated with obstruction or malabsorption. All of these operations can result in marginal ulceration on the jejunal side of the gastrojejunostomy and thus are combined with a generous distal gastrectomy. If this has already been done at a previous operation, the Roux or Braun operations may be attractively simple. Whether truncal vagotomy should be considered to decrease the risk of marginal ulceration is controversial because acid-suppressing medications may be equally effective. In addition, the benefits of decreased acid secretion following vagotomy may be outweighed by problems with vagotomy-associated dysmotility in the gastric remnant. The Roux operation may be associated with an increased risk of emptying problems compared to the other two options, but controlled data are lacking. Patients with debilitating bile reflux after gastrojejunostomy can be considered for simple takedown of this anastomosis provided that the pyloric channel is open.

Primary bile reflux gastritis (i.e., no previous operation) is rare, and may be treated with the duodenal switch operation, essentially an end-to-end Roux-en-Y to the proximal duodenum (see Fig. 26-68). The Achilles’ heel of this operation is, not surprisingly, marginal ulceration. Thus, it should be combined with highly selective vagotomy, and/or long-term acid suppressive medication.

Bile gastritis or esophagitis is a recognized complication after esophagogastrectomy with or without pyloroplasty. This can be effectively treated by division of the duodenum immediately distal to the pylorus with drainage of the prepyloric antrum into a Roux limb. Preservation of the right gastroepiploic pedicle is important. Proximal subtotal gastrectomy with esophagogastric anastomosis should be avoided, but when performed, the pylorus should be left intact.

**Roux Syndrome**

A subset of patients who have had distal gastrectomy and Roux-en-Y gastrojejunostomy will have great difficulty with gastric emptying in the absence of mechanical obstruction. These patients present with vomiting, epigastric pain, and weight loss. This clinical scenario has been labeled the
Roux syndrome. Endoscopy may show retained food or bezoars, dilation of the gastric remnant, and/or dilation of the Roux limb. Anastomotic inflammation and stricture from marginal ulceration is a confounding finding. An upper GI series confirms these findings and may show delayed gastric emptying. This is better quantified by a gastric emptying scan, which always shows delayed solid emptying and may show delayed liquid emptying as well.

GI motility testing shows abnormal motility in the Roux limb, with much of the propulsive activity toward, rather than away from, the stomach. Gastric motility also may be abnormal. Presumably, the disordered motility in the Roux limb occurs in all patients with this operation. Why only a subset develops the Roux syndrome is unclear. Perhaps patients with disordered gastric motility are at most risk. The disorder seems to be more common in patients with a generous gastric remnant. Truncal vagotomy also has been implicated.

Medical treatment consists of promotility agents. Surgical treatment consists of paring down the gastric remnant. Care should be taken to preserve adequate blood supply to the new gastric pouch. If the left gastric artery is intact, a vertically oriented lesser curvature based pouch (similar to gastric bypass) with excision of the fundus can be considered. If gastric motility is severely disordered, 95% gastrectomy or total gastrectomy should be considered. The Roux limb should be resected if it is dilated and flaccid, unless doing so puts the patient at risk for short bowel.

Gallstones
Gallstone formation following gastric surgery generally is thought to be secondary to vagal denervation of the gallbladder with attendant gallbladder dysmotility and stasis. Although prophylactic cholecystectomy is not justified with most gastric surgery, it should be considered if the gallbladder appears abnormal, especially if subsequent cholecystectomy is likely to be difficult. If preoperative evaluation reveals sludge or gallstones, or if intraoperative evaluation reveals stones, incidental cholecystectomy should be considered.

Weight Loss
Weight loss is common in patients who have had a vagotomy and/or gastric resection. The degree of weight loss tends to parallel the magnitude of the operation. It may be insignificant in the obese but devastating in the asthenic patient. The surgeon should always consider the possible nutritional consequences before performing a gastric resection for benign disease in a thin patient. The causes of weight loss after gastric surgery generally fall into one of two categories: altered dietary intake or malabsorption. If a stool stain for fecal fat is negative, it is likely that decreased caloric intake is the cause. This is the most common cause of weight loss after gastric surgery, and it may be due to small stomach syndrome, postoperative gastroparesis, anorexia due to loss of ghrelin, or self-imposed dietary modification because of dumping and/or diarrhea. Consultation with an experienced dietitian may prove invaluable.

Anemia
Iron absorption takes place primarily in the proximal GI tract, and it is facilitated by an acidic environment. Intrinsic factor, essential for the enteric absorption of vitamin $\text{B}_{12}$, is made by the parietal cells of the stomach. Vitamin $\text{B}_{12}$ bioavailability also is facilitated by an acidic environment. Folate-rich vegetables may be poorly tolerated if gastric emptying is delayed or if gastric capacity is limited. Since iron, $\text{B}_{12}$, and folate play vital roles in hematopoiesis, it is easy to understand why patients who have had a gastric operation are at risk for anemia. Anemia is the most common metabolic side effect in patients who have had a gastric bypass for morbid obesity. It also occurs in up to one-third of patients who have had a vagotomy and/or gastric resection. Iron deficiency is the most common cause, but vitamin $\text{B}_{12}$ or folate deficiency also occurs, even in patients who have not had total gastrectomy. Of course, patients who have had a total gastrectomy will all develop $\text{B}_{12}$ deficiency without any type of regular nonenteral vitamin $\text{B}_{12}$ administration. Gastric bypass patients should be given oral iron supplements and monitored for iron, $\text{B}_{12}$, and folate deficiency. Patients who have had a vagotomy and/or gastrectomy should be similarly monitored with periodic determination of hematocrit, red blood cell indices, iron and transferrin levels, $\text{B}_{12}$, and folate levels. Marginal nutrient status should be corrected with oral and/or parenteral supplementation.

Bone Disease
Gastric surgery sometimes disturbs calcium and vitamin D metabolism. Calcium absorption occurs primarily in the duodenum, which is bypassed with gastrojejunostomy. Fat malabsorption may occur because of blind loop syndrome and bacterial overgrowth or because of inefficient mixing of food and digestive enzymes. This can significantly affect the absorption of vitamin D, a fat-soluble vitamin. Both abnormalities of calcium and vitamin D metabolism can contribute to metabolic bone disease in patients following gastric surgery. The problems usually manifest as pain and/or fractures many years after the index operation. Musculoskeletal symptoms should prompt a study of bone density. Dietary supplementation of calcium and vitamin D may be useful in preventing these complications. Routine skeletal monitoring of patients at high-risk (e.g., elderly males and females and postmenopausal females) may prove useful in identifying skeletal deterioration that may be slowed or stopped with appropriate treatment after gastric surgery.

LAPAROSCOPIC GASTRIC OPERATIONS

The most common laparoscopic gastric operations performed today are for GERD and obesity. However, all conventional gastric operations can be performed with minimal access techniques. Some are more technically challenging (e.g., partial or total gastric resection) and are of debatable advantage over conventional open approaches. Certainly, highly selective vagotomy, vagotomy and gastrojejunostomy, and gastrostomy lend themselves to a minimal access approach. Laparoscopic local excision is often feasible for GI stromal tumors, leiomyomas, or gastric diverticula. Difficult to access lesions near the GE junction or pylorus may be removed through an anterior gastrotomy; more recent approaches utilizing transgastric ports or combined laparoscopic and endoscopic approaches show promise in allowing removal of practically any small gastric lesion with limited incisions.

In Japan and Korea, laparoscopic and robotic assisted approaches have been applied increasingly in the management of gastric cancer. Indeed, laparoscopic subtotal gastrectomy has supplanted the traditional open operation as the preferred operation for patients with earlier stage tumors, and laparoscopic total gastrectomy for proximal tumors is performed with regularity and excellent outcomes. The Asian experience...
has firmly established the feasibility of safe laparoscopic D2 gastrectomy. Translation of this experience to the United States, however, is not easily accomplished. Studies from Asia suggest that expertise in the laparoscopic approach require upwards of several high volume centers in the United States have reported excellent outcomes after laparoscopic gastrectomy. As robotic technology that facilitates dissection and anastomosis with articulated instrumentation and enhanced visualization becomes increasingly ubiquitous, the pendulum will likely swing toward increasing utilization of minimal access approaches for all gastric operations. 220

REFERENCES

Entries highlighted in bright blue are key references.


This page intentionally left blank
INTRODUCTION

Despite the global pandemic of obesity, there has been little progress in nonsurgical treatment approaches, especially among patients with severe obesity. In addition, the evidence base for bariatric procedures has grown rapidly over the last 10 years, yielding important short- and long-term data on the safety and efficacy of the surgical treatment for obesity and related metabolic disorders. Therefore, the approach for patients considering bariatric and metabolic surgery has now shifted to a well-informed and shared decision-making process as there are significant tradeoffs between the potential risks and benefits of these procedures.1

History

During the 1950s, operations were first performed to treat severe hyperlipidemia with associated obesity.2 These were ileocolic bypass operations to limit absorption and were associated with severe nutritional complications and liver failure postoperatively. A more modest jejunoileal bypass was performed next, also a malabsorptive operation, but it bypassed only a portion of the small intestine. Complications after this procedure included severe diarrhea, electrolyte disturbances, protein-calorie malnutrition, renal stones, and liver failure.

In 1969, Mason and Ito performed the first gastric bypass, describing a loop of jejunum connected to a transverse proximal gastric pouch.3 Bile reflux esophagitis was severe postoperatively, causing Griffin and colleagues to describe the Roux-en-Y modification of the gastric bypass in 1977.4 The gastric pouch was also altered from transverse to vertical using the upper lesser curvature at this time (Fig. 27-1).

In 1980, Mason5 first performed the vertical banded gastroplasty (VBG), which was a restrictive procedure using a stapled proximal gastric pouch of the upper lesser curvature of the stomach with a restrictive band for its outlet to the rest of the stomach. This operation produced excellent initial weight loss (50% of excess weight or more) with low morbidity and mortality. It rapidly became the most commonly performed bariatric operation in the United States during the 1980s. However, by the early 1990s, it became clear that patients who underwent VBG modified their diets to high-calorie soft foods and liquids and some regained weight.6 A significant incidence of stenosis at the cuff and staple line...
disruptions was also problematic. Long-term weight loss was poor, and by the 1990s in the United States, Roux-en-Y gastric bypass (RYGB) became the procedure of choice for bariatric surgery.

In the meantime, in Italy Scopinaro had developed and popularized the biliopancreatic diversion (BPD) in the early 1980s. This procedure was also modified to include duodenal switch (DS), the only major malabsorptive operation currently in use.

The laparoscopic approach to bariatric surgery became available in the 1990s, and Belachew performed the first laparoscopic adjustable gastric banding (LAGB) operation in 1994. Wittgrove and Clark performed the first laparoscopic RYGB the same year. LAGB was commonly performed in Europe and Australia during the late 1990s, and in 2001 it was approved for use in the United States. Sleeve gastrectomy (SG) as a primary bariatric operation has grown rapidly in use since 2008.

The laparoscopic approach to bariatric surgery became available in the 1990s, and Belachew performed the first laparoscopic adjustable gastric banding (LAGB) operation in 1994. Wittgrove and Clark performed the first laparoscopic RYGB the same year. LAGB was commonly performed in Europe and Australia during the late 1990s, and in 2001 it was approved for use in the United States. Sleeve gastrectomy (SG) as a primary bariatric operation has grown rapidly in use since 2008.

**Figure 27-1.** History of bariatric surgery. (Reproduced with permission from Arterburn DE, Courcoulas AP: Bariatric surgery for obesity and metabolic conditions in adults, BMJ. 2014 Aug 27;349:g3961.)
State of the Field

There has been an ongoing major shift in bariatric procedures both in the United States and worldwide\textsuperscript{15} with the rapid adoption of the laparoscopic sleeve gastrectomy and the simultaneous decreasing utilization of the laparoscopic adjustable gastric banding procedure (Fig. 27-2). International trends in the utilization of bariatric surgical procedures have also been published. These show that the total number of bariatric surgical procedures performed in 2014 was 579,517. The three most commonly performed procedures in the world were SG at 46%, followed by RYGB (40%), and LAGB (7%). The annual percentage changes from 2013 show an increased utilization of SG and a decreased use of RYGB in the United States, Canada, Europe, and Asia and Pacific countries. In Central and South America, however, the use of SG decreased, and RYGB was most commonly used.\textsuperscript{14}

Extension of the indication from bariatric surgery for weight loss to metabolic surgery to treat type 2 diabetes (T2DM) even in patients with less than severe obesity has been another recent development, driven by the availability of more level 1 data.\textsuperscript{15,16}

Also, considerable effort is now being devoted to the study of the basic physiologic mechanisms underlying weight loss and, perhaps more importantly, the resolution of comorbid medical problems associated with obesity. Despite the classic “restrictive” and “malabsorptive” anatomic conceptualizations of bariatric surgical procedures (see Fig. 27-1), there is much ongoing research in animal and human models towards understanding the specific underlying mechanisms of action, which may be more physiologic in nature.\textsuperscript{17} Some of the potential candidates for the mechanisms of action of bariatric procedures include alterations in ghrelin, leptin, glucagon-like peptide-1 (GLP-1), cholecystokinin, peptide YY (PYY), gut microbiota, and bile acids. In the future, bariatric procedures will not be described by anatomic surgical similarities but by how they affect key physiological variables, which will provide greater mechanistic insight into how the procedures actually work.

THE DISEASE OF OBESITY

Overview

Worldwide obesity has more than doubled since 1980. In 2014, 1\textsuperscript{9} 39% of adults age 18 years and over (38% of men and 40% of women) were overweight, and 13% of the world’s adult population (11% of men and 15% of women) were obese. In 2014, an estimated 41 million children under the age of 5 years were overweight or obese.\textsuperscript{16} Overweight and obesity are now on the rise in low- and middle-income countries, particularly in urban settings. Currently, 65% of the world’s population live in countries where overweight and obesity are linked to more deaths than underweight and malnutrition. Obesity is the second leading cause of preventable death in adults in the United States, after tobacco use.\textsuperscript{19}

The degrees of obesity are defined by body mass index (BMI = weight [kg]/height [m]\textsuperscript{2}), which correlates body weight with height. The World Health Organization international classification of overweight and obesity is shown in Table 27-1. It should be noted that for Asian populations, classifications remain the same as the international classification, but public health action points for interventions are set at a lower BMI threshold. For children, age needs to be considered when defining overweight and obesity, so for children age 5 to 19 years,

<table>
<thead>
<tr>
<th>CLASSIFICATION</th>
<th>PRINCIPAL CUTOFF POINTS</th>
<th>ADDITIONAL CUTOFF POINTS\textsuperscript{a}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal range</td>
<td>18.50–24.99</td>
<td>18.50–22.99, 23.00–24.99</td>
</tr>
<tr>
<td>Overweight</td>
<td>≥25.00</td>
<td>≥25.00</td>
</tr>
<tr>
<td>Preobese</td>
<td>25.00–29.99</td>
<td>25.00–27.49, 27.50–29.99</td>
</tr>
<tr>
<td>Obese</td>
<td>≥30.00</td>
<td>≥30.00</td>
</tr>
<tr>
<td>Obese class I</td>
<td>30.00–34.99</td>
<td>30.00–32.49, 32.50–34.99</td>
</tr>
<tr>
<td>Obese class II</td>
<td>35.00–39.99</td>
<td>35.00–37.49, 37.50–39.99</td>
</tr>
<tr>
<td>Obese class III</td>
<td>≥40.00</td>
<td>≥40.00</td>
</tr>
</tbody>
</table>

\textsuperscript{a}For Asian populations, classifications remain the same as the international classification, but public health action points for interventions are set at 23, 27.5, 32.5, and 37.5 kg/m\textsuperscript{2}.

overweight is BMI-for-age greater than 1 standard deviation above the World Health Organization (WHO) growth reference median, and obesity is greater than 2 standard deviations above the WHO growth reference median.

The fundamental cause of obesity and overweight is an energy imbalance between calories consumed and calories expended. Globally, there has been an increased intake of energy-dense foods that are high in fat and a decrease in physical inactivity due to the increasingly sedentary nature of many forms of work, changing modes of transportation, and increasing urbanization. These changes in dietary and physical activity patterns are the result of both societal and environmental changes associated with development. There is also a dearth of supportive policies to counteract these forces in sectors such as health, agriculture, transport, urban planning, environment, food processing, marketing, and education. Obesity as a disease was recognized by the American Medical Association in 2013. It is multifactorial in its etiology, and the components of the disease likely include a combination of both environmental and genetic factors.

U.S. Prevalence of Obesity

According to the 2013–2014 National Health and Nutrition Examination Survey (NHANES) data, in the United States the overall age-adjusted prevalence of obesity was 37.7% (95% confidence interval [CI], 35.8–39.7%); among men, it was 35.0% (95% CI, 32.8–37.3%); and among women, it was 40.4% (95% CI, 37.6–43.3%). The corresponding prevalence of class III obesity overall was 7.7% (95% CI, 6.2–9.3%); among men, it was 5.5% (95% CI, 4.0–7.2%); and among women, it was 9.9% (95% CI, 7.5–12.3%). Changes over the decade from 2005 through 2014, adjusted for age, race, smoking status, and education, showed significantly increasing trends among women for overall obesity and for class III obesity, but not among men.

For children in the United States, obesity is defined as a BMI at or above the sex-specific 95th percentile, and extreme obesity is defined as a BMI at or above 120% of the sex-specific 95th percentile on the U.S. Centers for Disease Control and Prevention (CDC) BMI-for-age growth charts. In the most recent NHANES study of U.S. children and adolescents age 2 to 19 years, the prevalence of obesity from 2011 to 2014 was 17.0%, and extreme obesity was 5.8%.

Causes of Obesity

Both genetic and environmental factors contribute to the development of obesity. Not everyone exposed to the prevailing environment becomes obese, suggesting that genetic mechanisms are operating at the individual level. Estimates vary, but twin, family, and adoption studies show that the rate of heritability of BMI is high, ranging from 40% to 70%. Eleven rare and monogenic forms of obesity are now recognized, including a deficiency of the leptin and melanocortin-4 receptors, which are expressed in the hypothalamus and are involved in regulating energy homeostasis. Heterozygous mutations in the melanocortin-4 receptor gene are currently the most common cause of monogenic obesity, causative in 2% to 5% of children with severe obesity.

Genes and environment interact in a complex process that regulates energy balance and weight. Reducing food intake or increasing physical activity leads to a negative energy balance and a cascade of compensatory adaptive mechanisms that preserve vital functions and are associated with reductions in resting energy expenditure, food preoccupation, and many other changes that depend on the amount and duration of caloric restriction. There is also a counterregulatory increase in appetite and food intake that limits the degree of expected weight loss that is associated with interventions such as exercise programs.

Individuals with obesity have excessive adipose cells, both in size and number. The number of such cells often is determined early in life; adult-onset obesity is largely a product of increase in adipose cell size. Weight gain results from increase in both adipose cell size and number. Adipose tissue may be deposited in large quantities in the subcutaneous layer of the abdominal wall or the visceral. Generally, males tend to have central visceral fat distribution, whereas females more often have a peripheral fat distribution. Central or visceral fat distribution is associated with metabolic diseases such as diabetes, hypertension, and the metabolic syndrome.

Concurrent Medical and Social Problems

Raised BMI is a major risk factor for diseases such as cardiovascular disease (mainly heart disease and stroke), which were the leading cause of death in 2012, diabetes, osteoarthritis, some cancers (including endometrial, breast, ovarian, prostate, liver, gallbladder, kidney, and colon). The risk for these conditions increases with increases in BMI. Childhood obesity is associated with a higher chance of obesity, premature death, and disability in adulthood. In addition to increased future risks, children with obesity experience sleep apnea, increased risk of fractures, hypertension, early markers of cardiovascular disease, insulin resistance, and psychological effects.

The severely obese patient typically presents with multiple chronic and weight-related problems or comorbidities/comorbid conditions. These include degenerative joint disease, low back pain, hypertension, obstructive sleep apnea, gastroesophageal reflux disease (GERD), cholelithiasis, T2DM, dyslipidemia, asthma, hypoventilation syndrome of obesity, right-sided heart failure, migraine headaches, pseudotumor cerebri, venous stasis ulcers, deep venous thrombosis (DVT), fungal skin rashes, skin abscesses, stress urinary incontinence, infertility, dysmenorrhea, depression, and large abdominal wall hernias.

There are anatomical, metabolic, and physiological effects of obesity through which this excess adiposity leads to disease risk factors and chronic diseases themselves (Fig. 27-3). Subcutaneous adipose tissue holds most of the stored lipid at a variety of anatomical sites while visceral adipose tissue is a smaller storage compartment with omental and mesenteric fat mechanismically linked to some of the metabolic disturbances and adverse consequences outcomes associated with obesity. Adipose tissue surrounds and compresses the renal parenchyma and may contribute to the hypertension frequently observed in patients who are obese. Obesity is frequently accompanied by an increase in pharyngeal soft tissues, which can lead to obstructive sleep apnea. Excess adiposity also imposes a significant mechanical load on joints, making it a risk factor for the development of osteoarthritis. An increase in intraabdominal pressure likely accounts for the elevated risks of gastroesophageal reflux disease (GERD), Barrett’s esophagus, and esophageal adenocarcinoma among those with obesity. Chronic overactivity of the sympathetic nervous system is present in some patients with obesity and may contribute to pathophysiological processes, including high blood pressure. Obesity is also associated with an increased prevalence of mood, anxiety, and other psychiatric disorders, especially...
among persons with severe obesity and in those seeking bariatric surgery. Decreased quality of life also results due to severe obesity. Most patients seeking surgical treatment of severe obesity do so because of the medical issues they face from comorbid conditions or the decreased quality of life they are experiencing as a result of severe obesity.

**MEDICAL MANAGEMENT OF OBESITY**

Treatments should be aligned with the severity of obesity, associated comorbid conditions, and the individual’s functional limitations. There are guidelines available to evaluate an individual’s health risks and potential treatment options. Three main treatment options exist with sufficient evidence-based support: lifestyle intervention, pharmacotherapy, and bariatric surgery.

**Lifestyle Intervention**

Lifestyle interventions designed to modify eating behaviors and physical activity are the first option for weight management, given their low cost and low risk. Behavioral therapy, the core of any lifestyle intervention, provides patients with techniques for adopting dietary and activity recommendations. Among these recommendations are regular recording of food intake, physical activity, and weight. Patients review their progress approximately weekly with a trained interventionist.
SPECIFIC CONSIDERATIONS

who provides support and encouragement, help setting goals, and problem-solving instructions. This type of comprehensive program results in a mean weight loss of 5% to 8%, and approximately 60% to 65% of patients lose 5% or more of initial weight. The Look AHEAD study randomized 5145 adults with obesity to either an intensive lifestyle intervention (ILI) or to a diabetes support group and education group (DSE) to assess the impact on weight loss, T2DM, and cardiovascular outcomes. At 1 year, the intensive intervention group lost an average of 8.6% initial weight compared to 0.7% in the support and education program. As shown in Fig. 27-4, 68% of participants in the Look AHEAD study lost at least 5% of their initial weight, and 37% of these participants lost at least 10%. Also at 1 year, participants undergoing the more intensive program experienced improved cardiovascular risk factors and glycemic control.

At 4 years, participants in the intensive intervention group experienced more weight loss (−6.15% ILI compared to −0.88% DSE), better glycemic control, fitness, and an improvement in cardiovascular risk factors. Nevertheless, the beneficial clinical effects of the improved weight loss achieved with intensive lifestyle intervention did not reduce morbidity and mortality associated with cardiovascular disease after 9.6 years when the Look AHEAD study was stopped due to futility for that primary endpoint. Figure 27-4 shows a comparison of several lifestyle intervention trials (Look AHEAD, the Diabetes Prevention Program [DPP] trial, and the trial reported by Teixeira et al) for >5% and >10% weight loss outcomes. These trials, specifically, were selected because they were judged to be good quality by the Guidelines (2013) for the Management of Overweight and Obesity in Adults and because the trial data were reported as categorical weight loss. Categorical weight loss data from the DPP trial were provided by the DPP Research Group to the authors of the review.

In summary, multidisciplinary lifestyle intervention and weight-management programs are viable and potentially cost-effective treatment options in overweight or obese patients with or without T2DM. Such approaches, however, often fail to achieve durable weight loss of more than 5% to 10%, so they are not effective enough for the severely obese. Importantly, lifestyle and medical approaches do not appear to improve cardiovascular outcomes in studies so far. Thus, further research is needed to evaluate the role for current medical and lifestyle therapeutic regimens for obesity and T2DM, including comparisons to surgical interventions. Certainly, and at the very least, these approaches are important adjuncts to bariatric surgery.

**Pharmacotherapy**

Medications may be considered as an adjunct to lifestyle modification in adults who have a BMI of 30 or higher or a BMI of 27 to 29 with at least one obesity-related condition. Phentermine and lifestyle intervention together lead to additive weight losses and should be used together and may also be helpful in facilitating the maintenance of reduced weight. Phentermine, the most widely prescribed weight-management medication in the United States, is a sympathomimetic amine that was approved by the FDA in 1959 for short-term use of fewer than 3 months long. There are now five newer FDA-approved medications for long-term weight management that include three single drugs and two combination drugs. In 1-year pivotal trials, total weight losses for the three single therapies (orlistat, lorcaserin, and liraglutide), the effects of which are mediated by different mechanisms, ranged from 5.8% to 8.8% of initial body weight. The two combination medications (phentermine–topiramate and naltrexone–bupropion) include drugs that act on neural weight-loss mechanisms. In 1-year pivotal trials, total weight loss for these combination drugs ranged from 6.4% to 9.8% of initial body weight.

These medications, when prescribed with lifestyle interventions, produce additional weight loss relative to placebo ranging from approximately 3% of initial weight for orlistat and lorcaserin to 9% for the higher-dose phentermine plus topiramate–extended release at 1 year. The proportion of patients achieving clinically meaningful (at least 5%) weight loss ranges from 37% to 47% for lorcaserin, 35% to 73% for orlistat, and 67% to 70% for higher-dose phentermine plus topiramate–extended release. All three of these medications produce greater improvements in cardiometabolic risk factors than placebo,
but none has been shown to reduce cardiovascular morbidity or mortality. There is limited data for the long-term safety and efficacy of these medications, and some of these drugs may increase heart rate or attenuate expected blood pressure reductions. In addition, completed trials of hard cardiovascular disease outcomes (heart attack and stroke) in patients treated with these medications have yet to be published, except in the case of liraglutide. Figure 27-4 shows a comparison of weight loss outcomes for these medications. The median percentages of participants who had a weight loss of at least 5% or 10% with each of five medications approved for long-term weight management are from a meta-analysis by Khera et al.

In summary, medications approved for long-term obesity treatment, when used as an adjunct to lifestyle intervention, lead to greater mean weight loss and an increased likelihood of achieving clinically meaningful 1-year weight loss compared to placebo. Yet weight loss medications are underutilized, likely due to several factors. First, patients are often disappointed by moderate weight loss. Second, there are requirements to pay a substantial portion of costs, which may lead to short-term rather than longer-term use. Third, there remain concerns about medication safety. Finally, weight regain is common after termination of drug treatment, which is discouraging to both patients and their providers.

**Barriers to Treatment**

Only a small fraction of patients for whom these medical treatments or bariatric surgery are indicated actually pursue and receive them. Past studies have estimated that 1% or fewer of those people with severe obesity who could consider bariatric surgery ever do so. Barriers to general obesity care include the slow recognition among providers that obesity requires long-term management, inadequate physician training in nutrition and obesity, limited reimbursement for the full range of treatments, lack of more effective and accessible lifestyle programs, and limited referrals of patients with severe obesity to experienced surgeons. Lack of knowledge about the more recent outcomes of bariatric surgery may also play a contributing role.

**CANDIDATES FOR BARIATRIC SURGERY**

**Indications**

There has been significant procedure evolution over the last several years indicating an ongoing major shift in bariatric procedures both in the United States and worldwide. According to a 2016 report from the American Society of Metabolic and Bariatric Surgery (ASMBS), the two most common procedures in the United States are RYGB and SG, accounting for approximately 25% to 30% and 50% to 60%, respectively, of annual cases. The utilization of LAGB has declined dramatically to under 10% of cases, and the malabsorptive procedure BPD with or without DS is utilized in less than 1% to 2% of cases. All of these procedures were defined by the Centers for Medicare & Medicaid Services (CMS) as standard approved procedures, noting that SG coverage is based on the discretion of regional carriers throughout the United States. The indications for performing bariatric surgery in class II and class III obesity still remain as described in the National Institutes of Health (NIH) Consensus Conference of 1991, and a summary of the broad selection criteria are shown in Table 27-2. In 2016, the second Diabetes Surgery Summit (DSS-II) published guidelines indicating that metabolic surgery should also be considered for patients with T2DM and BMI of 30 to 34.9 kg/m² (class I obesity) if blood sugar is inadequately controlled despite optimal medication treatment. In addition, these guidelines recommended that the BMI threshold for metabolic surgery (surgery for diabetes as the indication) should be reduced by 2.5 kg/m² for Asian populations at risk.

The NIH criteria for bariatric surgery do not set guidelines or limits for age, and surgical practice varies widely. The pediatric obesity epidemic is both increasing and also driving the adult epidemic, and a growing proportion of younger patients are potentially eligible for bariatric surgery. For young patients, there are concerns about assent to surgery and compliance with and adherence to postoperative lifestyle changes, but there are also some emerging data that suggest intervening earlier in the disease process may lead to improved reversal of comorbid conditions compared to adults. In addition, there is a longer period of postoperative benefit in terms of improved quality of life and prevention of or reduction in the emotional, social, and physical consequences of obesity (see “Bariatric Procedures in Adolescents”). Alternatively, older patients are more likely to have more numerous and debilitating comorbid conditions and thus have an immediate benefit in quality of life but not necessarily enhanced longevity. There is also some concern that recovery from potential complications is impaired in patients over the age of 65. Most studies in older patients have focused on RYGB and older restrictive procedures with limited follow-up. The results of more recent studies in older patients are generally equivocal in terms of any increased risk of morbidity and mortality or any difference in weight outcomes compared to younger adults.

One study has shown that the older patient population, especially those few patients older than age 70 undergoing bariatric surgery, did have an increased risk of mortality and morbidity after RYGB. A second study from the Utah Obesity group found that RYGB is protective against mortality even for older patients and also reduces the age-related increase in mortality observed in severely obese individuals not undergoing surgery.

**Contraindications**

Medical issues that preclude patients from being good surgical candidates include American Society of Anesthesiologists (ASA) class IV disease of a nature that makes surgical therapy extraordinarily high risk. Psychological instability or the inability to understand the implications of the proposed operation and what changes will result from it in terms of the patient’s lifestyle are also contraindications. Known and documented active drug or alcohol addiction is a contraindication to surgery (see Table 27-2). Tobacco use should be completely avoided by bariatric patients at all times, and smoking cessation should occur 6 weeks prior to surgery. After surgery smoking increases risks of poor wound healing, anastomotic ulcers, and impaired health. A poorly controlled eating disorder, especially bulimia, is also a contraindication to surgery. Nonambulatory status is a relative contraindication to surgery and is associated with increased surgical risk, especially if the obesity is so severe that the patient cannot normally do self-care or would not likely be able to do so after surgery. In addition to excessive morbidity, the placement of these individuals in care facilities postoperatively for recovery is often impossible due to their size and limitations of physical ability. Finally, lack of sufficient social support or an extremely poor or unsupportive home environment can be contraindications to surgical care, since such environmental factors are important to optimize outcomes once discharged from the hospital.
Figure 27-5. Bariatric surgery procedure evolution. A. Horizontal gastroplasty; B. vertical banded gastroplasty; C. Roux-en-Y gastric bypass; D. transected Roux-en-Y gastric bypass; E. laparoscopic adjustable gastric band; F. biliopancreatic diversion; G. biliopancreatic diversion with duodenal switch; H. vertical sleeve gastrectomy. (Modified with permission from Arterburn DE, Courcoulas AP: Bariatric surgery for obesity and metabolic conditions in adults, BMJ. 2014 Aug 27;349:g3961.)
MECHANISM OF ACTION OF BARIATRIC AND METABOLIC SURGERY

Overview
There is not yet a clear understanding as to how various bariatric procedures exert their effects on weight loss, metabolism, and glycemic control. Much effort is currently being devoted to gaining a better understanding of these specific mechanisms. A review of what is known from published animal and human studies about mechanisms related to the three most common surgical procedures is shown in Fig. 27-6. A few interim, summary statements can be drawn from this available data. First, neither LRYGB nor SG can be thought of as primarily “restrictive procedures,” and there are changes in behavior and physiology that likely help to maintain the new reduced body weight that are not observed after nonsurgically induced weight loss. LAGB appears to be more dependent on gastric restriction as both the behavioral changes and changes in gut hormone secretion are much less dramatic. Second, both LRYGB and SG are associated with metabolic improvements that are different from those that are caused by weight loss alone, and these mechanisms remain under current active study. For LAGB, the metabolic effects are mostly due to the impact of the resulting weight loss. This growing understanding of the physiology of these procedures points away from the older, classic anatomic classifications of “restrictive” versus “malabsorptive” procedures. This new conceptual approach has important implications for future studies of how bariatric surgery exerts its effects. Earlier hypotheses for the mechanism of action of LRYGB have been classified into either the “foregut hypothesis” or “hindgut hypothesis.” The foregut hypothesis states that improvements after LRYGB come from the bypassing of the upper small intestine that results in the reduction of nutrient-dependent hormonal actions that would normally impair glucose tolerance. The hindgut hypothesis states instead that the key events are the result of more rapid delivery of nutrients to the distal small intestine causing effects such as increased GLP-1/PYY secretion and the ileal brake. Now, the more recently recognized and common metabolic effects of SG and LRYGB may indicate directions for study away from this foregut/hindgut distinction, as the SG does not bypass the foregut or induce nutrients further down in the intestine.

Mechanisms of Bariatric Surgery (Weight Loss)
Certainly, one component mechanism by which RYGB produces weight loss is related to reduced caloric intake and malabsorption brought about by a smaller gastric volume and bypass of the proximal small bowel, so weight loss following LAGB and SG may be explained, at least in part and early on, by gastric restriction and resulting reduced food intake. Aside from anatomic changes induced by surgery, there are also changes in physiology that may affect food preferences and energy expenditure. Gastric emptying and insulin sensitivity increases following RYGB and SG. Functional magnetic resonance imaging has demonstrated a decreased neuronal activation of the food reward-related centers in response to high-calorie foods following RYGB. Neural signaling may be altered and stretch sensitive vagal endings in the new pouch and Roux limb, resulting in a feeling of early satiety. Hormonal changes are also evident, and in a number of studies looking at the effect of RYGB on ghrelin levels, results are conflicting. Changes in intestinal microbiota is another area of active study. Individuals with obesity have different gut flora compared to nonobese subjects. The Firmicutes (mainly Lactobacillus and Clostridium species) to Bacteroidetes ratio (Bacteroides or Prevotella species) is elevated in obese subjects. Following gastric bypass, the Firmicutes group decrease while Bacteroides/Prevotella increase at 3 and 6 months intervals. Bacteria transplant provides some of the benefits of gastric bypass surgery without the surgery. It has also been shown that administering oral lactobacillus post- RYGB leads to increased weight loss; this indicates benefits of changing gut microbiota to induce weight loss.

Serum bile acid levels also increase following gastric bypass. Gastric bypass diverts undiluted bile acids to the distal bowel. Bile acids activate protein-coupled receptor TGR5 present in L cells responsible for GLP-1 secretion. They also activate FXR (farsenoid-X receptor) in the jejunum, which regulates lipid and glucose metabolism. Furthermore, bile acids lead to rapid clearance of triglycerides. Bariatric surgery alters bile acid enterohepatic circulation in favor of weight loss and resolution of nonalcoholic steatohepatitis (NASH). Ryan et al demonstrated in a study on mice that the therapeutic value of SG is not limited to mechanical restriction but to an increase in circulating bile acids and associated changes to gut microbiota. Hollanda et al studied two cohorts of patients: those who lost more than 50% of their excess weight compared to those who did not. This group suggested that ghrelin and GLP-1 may be mediators of successful weight loss as those levels increased, while PYY and GLP-2,
<table>
<thead>
<tr>
<th></th>
<th>LRYGB</th>
<th>LAGB</th>
<th>SG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lipid</td>
<td>Elevated HDL</td>
<td>Elevated HDL</td>
<td>Elevated HDL</td>
</tr>
<tr>
<td></td>
<td>Reduced triglycerides</td>
<td>Reduction in triglycerides not as dramatic as LRYGB or SG</td>
<td>Reduced triglycerides</td>
</tr>
<tr>
<td>Glucose homeostasis</td>
<td>Improved fasting blood glucose and insulin sensitivity, prior to weight loss</td>
<td>Improvements are slower and not as dramatic as after SG or LRYGB</td>
<td>Improved fasting blood glucose and insulin sensitivity, prior to weight loss</td>
</tr>
<tr>
<td>Role of gastric restriction</td>
<td>Has not yet been directly tested</td>
<td>Failure of band leads to less gastric restriction and less weight loss</td>
<td>Gastric restriction is not the critical factor preventing hyperphagia</td>
</tr>
<tr>
<td>Gastric emptying</td>
<td>Few published studies</td>
<td>No overall change in gastric emptying rate; Emptying rate of proximal pouch created by band is enhanced</td>
<td>Most papers show increase</td>
</tr>
<tr>
<td>Energy expenditure</td>
<td>Controversial</td>
<td>Not reported</td>
<td>Unchanged, but only reported in one study</td>
</tr>
<tr>
<td>Leptin</td>
<td>Circulating leptin levels lower than expected for body weight Changes to leptin sensitivity not tested</td>
<td>Plasma leptin reduced, as expected for body weight; Changes to leptin sensitivity not tested</td>
<td>Circulating leptin levels lower than expected for body weight; Body weight changes not driven by changes to leptin sensitivity</td>
</tr>
<tr>
<td>Ghrelin</td>
<td>Reduced total ghrelin; Controversial, but no change in acyl-ghrelin levels</td>
<td>Increased circulating ghrelin</td>
<td>Reduced total ghrelin; Controversial, but no change in acyl-ghrelin levels</td>
</tr>
<tr>
<td>CCK</td>
<td>No change</td>
<td>No change</td>
<td>Not measured</td>
</tr>
<tr>
<td>GLP-1 (postprandial)</td>
<td>Weight loss-independent postprandial increase</td>
<td>Increased circulating GLP-1 but much less than RYGB or SG</td>
<td>Weight loss-independent increase comparable to LRYGB</td>
</tr>
<tr>
<td>PYY (postprandial)</td>
<td>Increased postprandial PYY levels; Reduced body weight loss in PYY knockout mice</td>
<td>No change</td>
<td>Increased postprandial PYY levels, comparable to levels after LRYGB</td>
</tr>
<tr>
<td>Bile acids</td>
<td>Increased plasma bile acids</td>
<td>Not reported</td>
<td>Increased plasma bile acids</td>
</tr>
<tr>
<td>Diet change</td>
<td>Decreased fat intake, more fruits and vegetables</td>
<td>Decrease bread intake and increase in caloric liquids; Greater fat intake and fewer fruits/vegetables than RYGB</td>
<td>Decreased fat intake, similar to LRYGB</td>
</tr>
<tr>
<td>Food Intolerance</td>
<td>Some dumping syndrome, usually well-tolerated</td>
<td>More persistent and problematic than LRYGB; Mainly vomiting</td>
<td>Little or none</td>
</tr>
</tbody>
</table>

**Figure 27-6.** Mechanisms of effect: comparison of LRYGB, LAGB, and SG. (Reproduced with permission from Stefater MA, Wilson-Pérez HE, Chambers AP, et al: All bariatric surgeries are not created equal: insights from mechanistic comparisons, Endocr Rev. 2012 Aug;33(4):595-622.)
gut mass and hypertrophy (citrulline), and the bile acid effect on fibroblast growth factor-19 (FGF-19) appear to have no effect on weight loss. In summary, there exists a complex relationship between dietary changes, bile flow changes, altered hormonal milieu and the gut microbiota that is not yet completely well characterized as it relates to weight loss after bariatric surgery.

**Mechanisms of Metabolic Surgery (Diabetes Improvement)**

Understanding the basic mechanism(s) of diabetes improvement following bariatric surgery is an important area of intensive study. Recently published data on worldwide trends in diabetes indicate that the number of adults with diabetes has increased from 108 million in 1980 to 422 million in 2014, the majority with T2DM. In the United States, diabetes is the number two cause of hospitalizations in adults age 18 years or older, accounting for approximately 11% of all hospital admissions. With no medical cure, the natural course of diabetes is characterized by progressive β-cell failure and development of microvascular and macrovascular complications, leading to renal failure, blindness, amputation, and death due to cardiovascular disease (CVD). Bariatric surgery has been renamed **metabolic surgery** for T2DM treatment and has emerged as an effective tool for control of hyperglycemia.

More than 20 years ago Pories et al found that bariatric surgery rapidly normalized blood glucose levels in people with obesity and T2DM, and 10 years later the majority remained disease free. He suggested that caloric restriction played a role but that there were likely other factors such as proximal intestinal nutrient exclusion, rapid distal gut nutrient delivery, and the role of gut hormones that would require further investigation. The findings of T2DM improvement and remission after bariatric surgery have now been widely replicated by others, and there is evidence that bariatric surgery prevents or delays incident cases of T2DM. Much work has been done to investigate these specific physiological mechanisms underlying the beneficial glycemic effects of bariatric surgery, but they remain incompletely understood. Candidate hypotheses include changes in bile acid metabolism, nutrient sensing and glucose utilization, intestinal adaptation, incretins, possible anti-incretin(s), and the intestinal microbiome. These physiologic and molecular changes lead to reduced hepatic glucose production, increased glucose uptake in tissues, improved insulin sensitivity, and enhanced β-cell function. A schematic of these potential mechanisms of improved glycemic control is shown in Fig. 27-7.

It is likely that several of these individual factors, acting together and with different impact based on the specific surgical procedure, are responsible for postoperative glycemic improvement. Work in this area is actively ongoing, and genomic, metabolomic, and gut microbiome studies will likely enhance the understanding of these changes. This may potentially lead to identifying novel pathways and potential therapeutic targets to replace bariatric procedures by equally effective, but less invasive, new treatments for obesity-related T2DM. In other words, understanding mechanisms of glycemic improvement after bariatric surgery may allow for the development of treatments to “bypass the bypass.”

**PREOPERATIVE ISSUES**

**Preoperative Preparation**

Patient selection for surgery should be based on a multidisciplinary team approach. All patients should undergo preoperative evaluation for obesity-related comorbidities and causes of obesity, with special attention directed to factors that could affect candidacy for bariatric surgery (Table 27-3).

The preoperative assessment of the patient for bariatric surgery must include input from the nutritionist as an important independent evaluation. Careful assessment of the patient’s eating habits, knowledge, self-awareness, and insight are important. An estimation of the patient’s motivation to change eating habits is important. The nutritionist should have at least one assessment session with the patient and an educational session preoperatively once the decision to proceed with surgery has been determined. The operation to be performed requires specific nutritional counseling and education. Psychological assessment is required by most programs and many insurance carriers with a goal of identifying potential contraindications to surgical intervention, such as poorly controlled psychiatric illness or active substance abuse, and identifying strategies to help with long-term weight management. There are published recommendations regarding the content of a mental health evaluation for bariatric surgery, but no consensus guidelines have been published. These evaluations are carried out by interview and questionnaires, which rely on clinical interviews including tests of personality or psychological conditions. More comprehensive evaluations also assess bariatric surgery knowledge, weight history, lifestyle habits, and potential barriers. Psychological assessment in clinical practice may be inaccurate compared to independent evaluations for research purposes, as patients present themselves in the most favorable light in order to gain access to surgery.

Obstructive sleep apnea (OSA) is prevalent in over 90% of bariatric surgery candidates with approximately one-third undiagnosed. The Epworth Sleepiness Scale, a standard set of questions evaluating daytime sleepiness, is often used as a screening tool for OSA. As OSA is associated with increased risk of mortality and in bariatric surgery patients, with adverse outcomes, routine preoperative screening with polysomnography should be considered. In addition, standard preoperative management of obese patients with OSA using continuous positive airway pressure (CPAP) is recommended. Asthma and hypoventilation syndrome of obesity are other significant pulmonary diseases often requiring preoperative management. Hypoventilation syndrome of obesity is defined as resting arterial partial pressure of oxygen less than 55 mmHg and arterial pressure of carbon dioxide greater than 47 mmHg, with accompanying pulmonary hypertension and polycythemia. Pulmonary consultation is indicated for patients with hypoventilation syndrome. Postoperative intensive care unit hospitalization, rarely used after bariatric surgery, may be indicated for these patients.

Preoperative weight loss can reduce liver volume/size and may help improve the technical aspects of surgery in those people with extreme central obesity and an enlarged liver, and it is sometimes utilized as a practice-specific recommendation or requirement. Ten percent total body weight loss (TBWL) with energy-restricted diets has been associated with a reduction in hepatic volume, variable perceived and measured improved facility in operative technique, variable effects on short-term complication rates, and weight loss. Cirrhosis has been associated with poor outcomes following bariatric surgery, including progression to liver transplantation.

Preoperative glycemic control should be optimized using diet, physical activity, and medications, as needed. Reasonable targets for preoperative glycemic control include a hemoglobin A1c value of 6.5% to 7.0% or less, a fasting blood glucose
level of ≤110 mg/dL, and a 2-hour postprandial blood glucose concentration of ≤140 mg/dL. More liberal preoperative targets with higher A1c should be considered in patients with advanced comorbid conditions or long-standing diabetes where lower targets are not attainable. For patients with active GERD on medication, a preoperative screening upper endoscopy to rule out Barrett’s esophagus and to rule out intrinsic lesions of the stomach or duodenum is recommended. This is especially true for patients planning LRYGB, where the distal stomach and duodenum will be precluded from easy inspection postoperatively. In addition, the presence of Barrett’s esophagus is a contraindication to SG, which is a reflux-inducing operation. The presence of a hiatal hernia detected on preoperative esophagogastroduodenoscopy will alert the surgeon for the need to perform intraoperative repair.

Patients with a history of DVT or cor pulmonale should undergo a diagnostic evaluation for DVT. A prophylactic vena caval filter may present a greater risk than benefit in patients

Figure 27-7. Schematic of potential mechanisms of improved glycemic control after LRYGB and SG. A. Immediate effects of RYGB and SG due to anatomical changes. B. Potential mediators/mechanisms involved. Cross talk occurs among these factors. C. Effects on glucose homeostasis. (Reproduced with permission from Batterham RL, Cummings DE: Mechanisms of Diabetes Improvement Following Bariatric/Metabolic Surgery, Diabetes Care. 2016 Jun;39(6):893-901.)
with a history of prior pulmonary embolism (PE) or DVT given the risks of filter-related complications including thrombosis. The overall risk of venous thromboembolism (VTE) after surgery is 0.42%, and over 70% of these events occur after hospital discharge, most within 30 days after surgery.116 The risk of VTE is greater in patients undergoing RYGB than in those undergoing LAGB and is more frequent following open surgery. Patients with a VTE event tend to be male, older, and have higher BMIs; they are also more likely to have a history of VTE.116 The risk of VTE is greater in patients with an inferior vena cava filter (hazard ratio [HR] 7.66, 95% CI 4.55–12.91).116 and there is evidence suggesting that prophylactic inferior vena cava (IVC) filter placement before bariatric surgery does not prevent PE and may lead to additional morbidity, which may outweigh its use.117,118

Candidates for bariatric surgery should avoid pregnancy preoperatively and for 12 to 18 months postoperatively and women who become pregnant after bariatric surgery should be counseled and monitored for appropriate weight gain, nutritional supplementation, and for fetal health.119 All women of reproductive age should be counseled on contraceptive choices following bariatric surgery as utilization, absorption, and effectiveness are inconsistent.120,121 Patients should be provided with educational materials and access to preoperative educational sessions. Multimedia tools for patient education and consent show promise for improving understanding.122-124 There should be a thorough and dynamic consent discussion regarding the risks and benefits, procedural options, and the need for long-term follow-up and vitamin supplementation (including costs required to maintain appropriate follow-up). Consent should include the experience of the surgeon with the specific procedure and whether the program participates in national quality improvement initiatives and certification.

Anesthesiology Issues
Two major challenges that face the anesthesiologist when performing a general anesthetic for the severely obese patient are vascular access and airway management. Fiberoptic laryngoscopy is often used for the most difficult class IV or even class III airways should standard laryngoscopy be determined to provide an inadequate view. Videotelescopic intubation systems are successfully used as well. Significant preoxygenation for 3 minutes or longer prior to intubation is used for the severely obese patient to provide a longer safe duration for intubation should difficulties be encountered. However, desaturation must be immediately addressed with reestablishment of oxygenated ventilation because this patient group does not tolerate any prolonged desaturation without potential adverse cardiopulmonary consequences.

The anesthesiologist must also manage alterations in cardiopulmonary function from the use of a pneumoperitoneum during laparoscopic bariatric procedures. These include the effects of carbon dioxide absorption on required minute ventilation, the potential for bradyarrhythmias, and the potential for decreased systemic pH with longer procedures in patients with preexisting cardiopulmonary disease. Arterial monitoring of the latter group of patients may be necessary by the anesthesiology team, and a radial arterial line is standard for such patients.125 Drug pharmacokinetics differ in severely obese patients as well. Changes in volume of distribution include smaller-than-normal fraction of total body water, greater adipose tissue content, altered protein binding, and increased blood volume. Possible changes in renal function and hepatic function must be considered when administering drugs. Specific anesthetic drug metabolic alterations in the severely obese include a larger volume of distribution of thiopentone, resulting in a prolonged effect of the drug. Calculation of the dosage should be done by lean body mass. Benzodiazepines also exhibit a prolonged elimination phase, causing prolongation of their effects. Increased pseudocholinesterase activity is present in the severely obese patient, requiring increased dosages of pancuronium. Enflurane metabolism is increased over the average-sized person, requiring a lower dosage of this agent.

Enhanced Recovery After Surgery
Enhanced recovery after surgery (ERAS) protocols have been initiated in bariatric surgery and have demonstrated promise to decrease surgical morbidity. Additionally, a recent meta-analysis has identified a significant decrease in length of stay (standard mean difference = −2.40 [−33.89, −0.89], P = 0.002).126 In 2016, the ERAS Society published evidence-based guidelines

---

**Table 27-3**

**Preoperative checklist for bariatric surgery**

- Complete for History & Physical (H&P) (obesity-related comorbidities, causes of obesity, weight/BMI, weight loss history, commitment, and exclusions related to surgical risk)
- Routine labs (including fasting blood glucose and lipid panel, kidney function, liver profile, lipid profile, urine analysis, prothrombin time/INR, blood type, CBC)
- Nutrient screening with iron studies, B12 and folic acid (RBC folate, homocysteine, methylmalonic acid optional), and 25-vitamin D (vitamins A and E optional); consider more extensive testing in patients undergoing malabsorptive procedures based on symptoms and risks
- Cardiopulmonary evaluation with sleep apnea screening (ECG, CXR, echocardiography if cardiac disease or pulmonary hypertension suspected; DVT evaluation if clinically indicated)
- GI evaluation (H pylori screening in high-prevalence areas; gallbladder evaluation and upper endoscopy if clinically indicated)
- Endocrine evaluation (A1c with suspected or diagnosed prediabetes or diabetes); TSH with symptoms or increased risk of thyroid disease; androgens with PCOS suspicion (total/bioavailable testosterone, DHEAS, D4-androstenedione); screening for Cushing’s syndrome if clinically suspected (1 mg overnight dexamethasone test, 24-hour urinary free cortisol, 11 PM salivary cortisol)
- Clinical nutrition evaluation by registered dietitian
- Psychosocial-behavioral evaluation
- Document medical necessity for bariatric surgery
- Informed consent
- Provide relevant financial information
- Continue efforts for preoperative weight loss
- Optimize glycemic control
- Pregnancy counseling
- Smoking cessation counseling
- Verify cancer screening by primary care physician

for perioperative care in bariatric surgery.\textsuperscript{127} The guidelines include recommendations in preoperative, intraoperative, and postoperative care. These include shorter acting and lower absorption anesthetic agents and opioid minimization as important intraoperative recommendations.

**Special Equipment and Infrastructure**

The special needs of the bariatric patient and program extend from the entry to the hospital and clinic, to the operating room, and throughout the inpatient and outpatient experience. The program needs infrastructure and support at all levels including support staff, physicians and surgeons, administrators, program directors, psychologists, and nutritionists. The physical plant needs to include extra-wide doorways, special seating, a scale that weighs up to 800 lb (363 kg), larger patient gowns, large blood pressure cuffs, and floor-mounted toilets. In the operating room, the table must accommodate 600 to 800 lb (272 to 363 kg) and must position in steep reverse Trendelenburg position. Larger lower extremity compression devices, extra padding, safety belts, and a footboard are required. An angled (30° or 45°) telescope, extra-long graspers and staplers, and a liver retractor system are all standard equipment. Staff sensitivity training for the care of the obese as well as regular education about the complications of bariatric surgery are program requirements.

**BARIATRIC SURGICAL PROCEDURES**

VBG shown in Figs. 27-1 and 27-5, although still listed as one of the approved operations for the surgical treatment of severe obesity based on the NIH Consensus Conference of 1991,\textsuperscript{51} is not currently performed due to poor long-term weight loss and technical complications, so it is of historic interest only, and the surgical technique will not be described here.\textsuperscript{8} The other procedures described in this section will be articulated using a laparoscopic approach as that is the dominant method. RYGB, BPD, and DS may still be performed by some surgeons using an open approach, but this has now become the exception. In this text LRYGB will refer specifically to RYGB performed by the laparoscopic approach, while RYGB will indicate procedures performed by the open approach or by both approaches as is the case for many studies of outcomes. Minimization of the morbidity of the open incision, especially incisional hernias and wound complications, as well as earlier hospital discharge and lower 30-day complication rates have all been clearly shown to favor using a laparoscopic approach when feasible.\textsuperscript{128-130} Laparoscopy begins with the safe creation of a pneumoperitoneum, often a difficult step in the bariatric patient. A tracheostomy hook can be inserted through a trocar-sized incision to elevate the fascia in the left subcostal region to facilitate the insertion of a Veress needle into an appropriate location for pneumoperitoneum creation. The use of a Hasson approach for creating a pneumoperitoneum in the bariatric population may be limited by the thick body wall. In the patient with an extremely thick body wall, extra-long trocar ports can be used for laparoscopic surgery. The pneumoperitoneum pressure that is used when performing bariatric surgical procedures is generally in the 15 to 18 mmHg range. A high-flow insufflator is mandatory to maintain the pneumoperitoneum for adequate and safe visualization.

When an open surgical approach is used for any of these procedures, an upper midline incision with table mounted retractors is the most commonly used approach. The robotic approach to bariatric procedures is also now utilized with purported advantages of reduction of the use of the open technique, improved surgical, length of stay, cost outcomes, and potentially improving ergonomics and resultant surgeon fatigue and injury. A meta-analysis involving 27 studies and over 25,000 patients concluded that there were no significant differences between robotic bariatric surgery and laparoscopic bariatric surgery with respect to overall complications, length of stay, reoperation, conversion, and mortality.\textsuperscript{131} Another study utilizing University Consortium data demonstrated no difference in hospital mortality, major complications, readmissions, or length of stay between the robotic and laparoscopic approach.\textsuperscript{132} In both studies, robotic surgery did increase significantly operative time and hospital costs (>20%) compared to laparoscopic approaches. As yet, larger prospective cohort studies and/or randomized trials have not yet been published, so the role of this technique is still to be defined and further studies are needed.\textsuperscript{133,134}

**Laparoscopic Roux-en-Y Gastric Bypass**

**Background and Patient Selection.** Figure 27-8 depicts the configuration of the LRYGB. It is an appropriate operation for consideration for most patients eligible for bariatric surgery. Relative contraindications specifically for LRYGB include previous gastric surgery, previous antireflux surgery, severe iron deficiency anemia, distal gastric or duodenal lesions that require ongoing future surveillance, and Barrett’s esophagus with severe

![Figure 27-8. Configuration of laparoscopic gastric bypass. (Reprinted with permission from Cleveland Clinic Center for Medical Art & Photography © 2005-2009. All Rights Reserved.)](image-url)
dysplasia. The major feature of the operation is a proximal gastric pouch of small size (<20 mL) that is totally separated from the distal residual stomach. A Roux limb of proximal jejunum is brought up and anastomosed to the pouch. The pathway of that limb can be anterior to the colon and stomach, posterior to both, or posterior to the colon and anterior to the stomach. The length of the biliopancreatic limb from the ligament of Treitz to the distal enterenterostomy is 20 to 50 cm, and the length of the Roux limb is 75 to 150 cm.

Creating the proximal gastric pouch by totally dividing it from the distal stomach is superior to simply stapling and partitioning the stomach, since the latter is associated with a high incidence of staple line breakdown.\textsuperscript{135} The size of the proximal gastric pouch must be small to create adequate restriction and should be based on the lesser curvature of the stomach to prevent dilation over time. Length of the Roux limb was associated with higher short-term weight loss for longer length limbs,\textsuperscript{136} but this difference becomes less meaningful on long-term follow-up and has not been demonstrated in more recent studies.\textsuperscript{137} Gastric pouch size and caliber of the gastrojejunostomy have not, in any studies, been shown to be related to weight loss. The gastrojejunal anastomosis can be constructed in a variety of ways, including hand sewn techniques and linear and circular staplers. Smaller diameter circular staplers are associated with a higher incidence of postoperative stenosis, and linear stapling is associated with a lower incidence of stenosis compared to circular stapling.\textsuperscript{138,139}

**Technique.** The operation generally is performed using five ports plus a liver retractor as shown in Fig. 27-9. Both the surgeon, who stands on the patient’s right, and the first assistant, who stands on the patient’s left, have two ports for instruments. The telescope requires a port, usually in the supraumbilical region. The assistant’s ports are in the left subcostal and flank areas, while the surgeon may have both ports in the right upper quadrant or one on each side of the camera. Division of the proximal jejunum at 40 to 50 cm distal to the ligament of Treitz is performed with the linear stapler, using a vascular stapler cartridge. Further division of the mesentery at that location is performed either with the stapler or harmonic scalpel, such that adequate mobilization of the Roux limb is achieved. A Penrose drain or a marking suture is placed on the proximal Roux limb for identification and facilitation of advancement to the gastric pouch (Fig. 27-10). The length of the Roux limb (usually 100–150 cm) to be created is measured. A jejunojejunostomy is then created to the proximal end of the biliopancreatic limb at the previously determined location along the Roux limb. A side-to-side stapled anastomosis is performed (Fig. 27-11). Either single- or double-fired staple technique (the latter using a stapler

---

**Figure 27-9.** Port scheme for laparoscopic gastric bypass. (Reprinted with permission from Cleveland Clinic Center for Medical Art & Photography © 2005-2009. All Rights Reserved.)

**Figure 27-10.** Creating Roux limb during laparoscopic gastric bypass.
fired in each direction) is used. The stapler defect is optimally closed with sutures but can be closed with a stapler if great care is taken not to narrow the lumen of the alimentary tract at this location. Once the stapler defect is closed, the mesenteric defect is then also closed with running permanent suture.

Passage of the Roux limb toward the stomach is now performed. If an antecolic route is to be used, the end of the Roux limb is brought up so as to confirm its ability to reach the stomach (Fig. 27-12). If a retrocolic route is to be used, a defect is made in the transverse colon mesentery just to the left and slightly above the ligament of Treitz. The proximal end of the Roux limb is placed into the retrogastric space. The left lobe of the liver is now retracted using any one of several retractor types. The patient is moved to a reverse Trendelenburg position. The harmonic scalpel divides the peritoneum in the area of the angle of His, and then it is used to open an area along the lesser curvature of the stomach approximately 3 cm down from the gastroesophageal junction. Another approach for creating access to the lesser curvature of the stomach is to use a white or gray load (vascular load) of the stapler and divide the lesser curvature vessels up to the surface of the stomach. Then a blue load of the stapler is fired one time transversely from the lesser curvature side partially across the stomach, followed by multiple subsequent firings of the stapler upward in the direction of the angle of His, to completely separate the proximal gastric pouch from the remainder of the stomach (Fig. 27-13). Optionally, use of an Ewald tube passed by the anesthesiologist and maneuvered to lie against the lesser curvature of the proximal stomach can help calibrate the pouch size.

Once the pouch is created, the Roux limb is brought up to the proximal gastric pouch. For the linear stapled anastomosis, the proximal end of the Roux limb is aligned with the distal gastric pouch end, and the sides of the organs are sutured together to maintain their side-by-side position. A stapler is introduced through a gastrotomy and an enterotomy for the two legs of the stapler, and the anastomosis is created (Fig. 27-14). The stapler defect is closed with sutures and often reinforced with a second
Figure 27-14. Gastrojejunostomy in laparoscopic Roux-en-Y gastric bypass. (Reprinted with permission from Cleveland Clinic Center for Medical Art & Photography © 2005-2009. All Rights Reserved.)

Layer of sutures. The gastrojejunostomy may be tested for security by using either methylene blue injected under pressure through the Ewald tube or a flexible upper endoscopy intraoperatively to test for air leakage from the anastomosis. The final step of the operation involves suture closure of all mesenteric defects using permanent suture.

Alternatively, a completely hand-sewn gastrojejunostomy can be created using two layers of absorbable suture to anastomose an approximately 1-cm gastrotomy and enterotomy. The circular anastomosis technique is another approach to complete the gastrojejunostomy and is also a particularly useful technique for “salvage” anastomosis if the gastric pouch is very small and/or high. This is done through placement of the anvil of the stapler through the anterior wall of the proximal gastric pouch. This is accomplished by pulling the anvil transorally via an endoscopically placed guidewire (Fig. 27-15), making a gastrotomy in the pouch that is later closed, or making a gastrotomy in the lower stomach before completing gastric division to create the pouch, allowing the anvil to be placed into the lumen of the stomach and then be brought through the anterior stomach in an area that is subsequently included in the proximal gastric pouch (Fig. 27-16).

Procedure-Specific Complications. Mortality after LRYGB is now consistently less than 0.5% in most large reported series. Data from several national data sets/studies find a mortality rate of approximately 0.3%, 0.14%, and 0.2% at 30 days overall. Overall morbidity after LRYGB has also been low. In the Longitudinal Assessment of Bariatric Surgery (LABS) study, a composite endpoint including death, deep-vein thrombosis or venous thromboembolism, reintervention, or failure to be discharged by 30 days after surgery occurred in 4.8% of those who had undergone LRYGB. Overall morbidity alone was 14.87% for 30,864 gastric bypass procedures. Complications that do occur after LRYGB include a 0.3% incidence of anastomotic leak, 0.33%...
incidence of venous thromboembolism, a 3% to 5% incidence of wound infections or problems, a 3% to 15% incidence of marginal ulcers, an approximately 7% incidence of bowel obstruction, a 4% incidence of postoperative transfusion, and a 1% to 19% incidence of anastomotic stenosis, based on the type of anastomosis created. Postoperative nutritional complications after LRYGB include a 66% incidence of iron deficiency, a 5% incidence of iron deficiency anemia, a 50% incidence of vitamin B$_12$ deficiency, and an at least 15% incidence of vitamin D deficiency, which usually is present preoperatively. Both early and late dumping occur in an unspecified number of postoperative cases, as the symptoms are difficult to document and overlap with other problems such as hypoglycemia.

Several complications that are specific to LRYGB must be emphasized. One of the most important is small bowel obstruction. This complication must be treated differently than in the average general surgery patient, whose complication is usually from adhesions and often will resolve with conservative, nonoperative therapy. Patients who have had LRYGB who present with obstructive symptoms generally require surgical therapy on an emergent basis. This is because the etiology of the bowel obstruction after LRYGB is often an internal hernia from inadequate or nonclosure of the mesenteric defects by the surgeon at the time of operation. Thus, treatment for these patients differs from most patients with small bowel obstruction. One of the most important points of this chapter is to emphasize to general surgeons to be aware of the need to emergently operate on patients after LRYGB who present with small bowel obstruction. Currently, centers that perform small bowel transplantation are seeing patient referral for that procedure after small bowel obstruction after LRYGB, where patients developed infarction of most of the bowel from an internal hernia and have short gut syndrome. Other patients, for whom surgery is delayed and the bowel infarcts, do not survive. When the surgeon does encounter bowel obstruction after LRYGB, he or she can expect to see proximally dilated bowel. Cutoff of passage of contrast on CT scan at the enteroenterostomy is particularly suggestive of this diagnosis (Fig. 27-17). The surgical treatment of this particular problem can, if addressed early in the course of the obstruction, be treated laparoscopically. The surgeon must place a trocar for the telescope low enough in the abdomen to adequately survey most of the small intestine. The cecum and terminal ileum are identified, and the bowel is followed retrograde from the terminal ileum to determine the anatomy. Often much of the small bowel is herniated through a mesenteric defect, and only this technique allows the surgeon to reliably identify the bowel and decompress it appropriately. If the bowel is viable,
suturing the mesenteric defect is all that is needed for treatment. It should be emphasized that either an antecolic or retrocolic placement of the Roux limb can result in this complication, as internal hernias can arise from either approach.

Marginal ulcers are another complication relatively specific to LRYGB. The patient presents with pain in the epigastric region that is not altered by eating. Diagnosis is by endoscopy. Treatment is medical with proton pump inhibitors, which are effective in 90% of cases. Only those with a gastrogastric fistula to the distal stomach, severe stenosis of the lumen of the gastrojejunostomy, or acute perforation require surgical therapy. Treatment of a perforated marginal ulcer is a laparoscopic Graham patch. Stenosis of the gastrojejunostomy has been reduced by the use of the linear stapling technique. Stenosis symptoms usually appear from 6 to 12 weeks postoperatively, but less commonly can occur later. Diagnosis is by upper endoscopy. Treatment is circumferential balloon dilatation. Resolution normally occurs with one to two treatments. Less than 10% of patients require reoperation, and those are almost always associated with concurrent marginal ulcers.

In the immediate postoperative period, anastomotic leak is the single serious complication after RYGB, either open or laparoscopic. Careful vigilance and a high index of suspicion for this problem are important since its presentation may be insidious and the patient’s demise, if untreated, may be sudden and complete. Tachycardia, tachypnea, fever, and oliguria are the most common symptoms that should arouse suspicion for this problem. The treatment is surgical, except in rare circumstances where a drain is already in place, no hemodynamic or clinical deterioration is present, and the leak is contained. Usual surgical treatment involves repair as feasible, drainage, and creation of a reliable feeding access through a distal Stamm gastrostomy.

In the first few hours or day after surgery, hematemesis indicates bleeding from the gastrojejunostomy unless proven otherwise. The dangers to the patient include aspiration, life-threatening hemorrhage, or more commonly intraluminal hematoma of the Roux limb and enterenterostomy, which then causes an obstruction of the biliopancreatic limb leading to distal gastric staple line rupture. In fact, any obstructive symptoms in the first few weeks after surgery or any signs of obstruction of the biliopancreatic limb on postoperative swallow studies due to stenosis of the enterenterostomy require immediate surgical intervention to prevent rupture of the distal gastric staple line. Some reports show that percutaneous decompression of the distal stomach can help to ameliorate the problem, but operative therapy to decompress the stomach and treat the obstruction is first-line therapy.

Laparoscopic Sleeve Gastrectomy

Background and Patient Selection. SG was originally introduced as the first of a two-stage operative treatment for patients with super obesity (BMI >60 kg/m²). Its currently utilization is as a primary single-stage operation, but the possibility of a second-stage treatment remains, especially for the super obese patients, depending on the effectiveness of it as the primary operation. In addition, patients who have longstanding severe GERD may not be good candidates for SG as GERD is worsened by the anatomical configuration of the SG. Barrett’s esophagus is also a contraindication for performing SG, since the potential for future esophageal dysplasia and the need for an available intact stomach for esophageal reconstruction outweigh the potential advantages of the procedure.

Figure 27-18. Port scheme for laparoscopic sleeve gastrectomy.

Technique. The patient is positioned supine, with foot support to allow reverse Trendelenburg positioning. The surgeon stands to the patient’s right along with the camera driver, while the assistant stands to the patient’s left. Port placement may vary, but a recommended port placement schema is shown in Fig. 27-18. The 15-mm port, helpful for removal of the stomach, is located in either the camera (just to the patient’s left of the umbilicus) or surgeon’s right hand (right upper quadrant near the midline) location. The other of these ports is a 12-mm port. The assistant has two 5-mm ports available in the left upper quadrant laterally, and the surgeon’s left-hand port is a 5-mm port more lateral and superior in the right upper quadrant. A liver retractor is placed in the epigastric region.

The operation begins by devascularizing the greater curvature of the stomach, beginning 3 to 5 cm proximal to the pylorus. The division of all vessels adjacent to the greater curvature is continued up to the left crus of the diaphragm. A complete mobilization of the fundus in this area and division of posterior fibrous attachments to the antrum and body of the stomach are then performed such that the stomach is attached solely by the lesser curvature blood supply and the pyloric and esophageal regions. Stapled division of the stomach now follows. The first firing of the stapler occurs from the point of devascularization of the greater curvature at an angle pointing toward a point about 2 cm lateral to the incisura. The antrum of the stomach is at its thickest here, and so it is important to be certain the stapler load used is sufficiently large enough to allow good approximation and closure of the divided stomach. Two staple firings are performed, which takes the gastric division to past the incisura. After the first staple firing, some surgeons will engage the anesthesiologist to pass a 32- to 40-French bougie and position it along the lesser curvature of the stomach. This bougie then serves as a guide for further gastric division. Alternatively, some surgeons will insert the endoscope instead of the bougie as a guide for gastric division. It can also be used to test for air leaks, bleeding, or obstruction as it is withdrawn after gastric division. Dividing the stomach adjacent to the bougie or endoscope will produce the desired diameter of the gastric sleeve. It is most important not to narrow the stomach lumen at the incisura. During the second and third firing of the stapler to divide the stomach, it is critical to confirm by visualization of both the anterior and posterior surfaces of the stomach that the incisura area is not narrowed. By the third firing of the stapler, usually the angle of the gastric division is now pointed directly toward the angle of His, parallel to the bougie (Fig. 27-19).
At this point, changing staple loads to lower staple height is advisable.

Once the stomach is completely divided up to the angle of His, the staple line is inspected for hemostasis and integrity. Some surgeons will reinforce the staple line with a buttress material, while others will invaginate the staple line with a running serosa to serosa suture. Some surgeons will exchange the bougie at this point for a 32-French Ewald tube and perform a methylene blue leak test. Alternatively, if an endoscope is used, it is withdrawn with insufflation, and the staple line is inspected for air leaks while submerged in saline. The specimen is removed through the 15-mm trocar site, usually with only slight enlargement of the site. Figure 27-20 shows the completed operation. Controversy still exists as to the optimal size of the bougie used during the procedure and the relative utility of methods used to oversew or reinforce the staple line.

Procedure-Specific Complications. The major factor unique to SG is that it creates a high-pressure gastric tube. This increased intraluminal pressure places the staple line at risk for leakage and increased risk for GERD. As noted previously, controversy exists about both bougie size and staple line reinforcement/oversew as they relate to clinical outcomes. One summary of the literature shows that the stenosis rate is lower if a 40-French bougie is used, and the leak rate may also be lower without compromising weight loss. However, individual institutional experiences with smaller-sized bougies have shown the potential for good weight loss and no increased incidence of stenosis. Another controversial area is that of staple line reinforcement with staple buttressing material or reinforcement with oversewing. The overall bleeding rate for the staple line after SG is generally cited as about 2% in collected series. There have been no studies that have shown a definitive decrease in this bleeding rate with the use of buttress materials; however, a panel of experts has voiced support for a decreased incidence of bleeding from the staple line if buttress material is used.

One meta-analysis did show that there is evidence to suggest buttress materials may decrease the staple line leak rate. Other prospective randomized studies have failed to show a benefit of buttress materials for leak prevention. A more recent study of over 180,000 SG procedures in the Metabolic and Bariatric Surgery Accreditation and Quality Improvement Program (MBSAQIP) national database showed that staple line reinforcement cases were associated with higher leak rates (0.96% vs. 0.65%, odds ratio [OR] 1.20 95% CI 1.00–1.43) and lower bleeding rates (0.75% vs. 1.00%, OR 0.74 95% CI 0.63–0.86) compared to no reinforcement, at the patient level. At this time, there is no preponderance of data to support one approach as being superior to others for both staple line bleeding or leak rates. Given this state of the literature, a surgeon should consider the risks, benefits, and costs of these surgical techniques and utilize those that, in their hands, minimize morbidity while maximizing clinical outcomes.

If there is a relative obstruction or stenosis of the sleeve, which most often occurs at the incisura because of narrowing there during formation of the sleeve, pressure above the level of the obstruction will be even more elevated and create an increased risk for staple line leak. Leaks of the proximal staple line are the most frequent type seen after SG and often are felt to be related to increased intraluminal pressure distally. They may also be related to stapling too close to the angle of His, with resultant instability of the tissue directly adjacent to the esophagus in this area. It is important not to staple too close to the angle of His during the final stapling division portion of the stomach so as to not further weaken the staple line in this area. Proximal staple line leaks may also present as late leaks, 6 weeks to months following the procedure. Late leaks are rare.
following other bariatric procedures, but are seen with SG, and index of suspicion should remain high.

Distal staple line leaks are different from proximal staple line leaks and are usually associated with earlier presentation and related to mechanical failure of the staple line to securely approximate the thicker distal gastric tissue. These leaks are more amenable to successful repair with a reoperation, whereas proximal leaks may not improve with oversewing at a reoperation unless the mechanics of the relative distal obstruction and high intraluminal pressure of the sleeve are also treated. Endoscopic intervention to dilate stenotic areas as well may be beneficial in the setting of a stenosis with or without a proximal leak. Care must be taken by the endoscopist to not excessively dilate the tract beyond the original size of the bougie used. Another factor that may influence stenosis at the incisura is that there may be a relative twisting of the stomach at this location, with the antrum being partially twisted away from the upper portion of the sleeve. Endoscopic treatment can help straighten and markedly alleviate the obstruction in such cases. Thus, relatively early endoscopic intervention is appropriate in the patient with a stenosis at the incisura. One study has shown that endoscopic dilation is usually successful in treating stenosis after SG, with a mean of 1.6 dilations being done an average of 48 days postoperatively.163

The patient with the proximal gastric staple line leak because of mechanical factors may experience persistence of the leak for months. Nonsurgical treatment with drainage and stenting can be used initially. Some now advocate for conversion of the patient with a longstanding leak after SG to a RYGB to provide a low-pressure anastomosis above the site of the stenosis.164,165 Similarly, persistent stenosis of the sleeve despite conservative therapy and endoscopic dilatation also is an indication for conversion to RYGB.

Laparoscopic Adjustable Gastric Banding

Background and Patient Selection. LAGB involves placement of an inflatable silicone ring around the proximal stomach. The band is attached to a reservoir system that allows adjustment of the tightness of the band. This reservoir system is accessed through a subcutaneously placed port, similar in concept to ports used for chemotherapy via central venous catheters. Figure 27-21 shows the LAGB apparatus in place. Patients who have had previous upper gastric surgery, such as a Nissen fundoplication, and those with severe GERD are relatively poor candidates for LAGB due to altered proximal gastric anatomy interfering with proper band placement or worsening of GERD symptoms. Two major types of bands have been used for this procedure. The original Lap-Band has been used most frequently. The Swedish Band, remarked as the Realize Band in the United States, is slightly wider and larger in circumference than the Lap-Band but is no longer being manufactured. The port systems have differences as to profile and methods of attachment to the fascia.

Technique. Port placement for LAGB has varied among surgeons. Usually some combination of two ports for the surgeon’s hands, one or two for the assistant, a port for the telescope, and a liver retractor site are needed. With the patient placed in reverse Trendelenburg position, the procedure begins with division of the peritoneum at the angle of His and then division of the gastrohepatic ligament in its avascular area (the pars flaccida) to expose the base of the right crus of the diaphragm. If a hiatal hernia is present, it must be repaired at this point, using a standard posterior esophageal dissection to expose the crura and perform suture repair. A grasper is inserted along the base of the anterior surface of the diaphragmatic crura, from right to left, emerging at the angle of His in the area of the divided peritoneum (Fig. 27-22). The device is then used to pull the band...
underneath the posterior surface of the gastroesophageal junction. This technique, by passing the band through some fibrous tissue in this plane, serves to anchor the band more securely posteriorly. During the initial years of band placement, a retrogastric location of the posterior half of the band in the free space of the lesser sac caused an unacceptably high incidence of slippage and prolapse of the band. The adoption of the *pars flaccida* technique decreased the incidence of such slippage.166

Once the band is passed around the proximal stomach, it is locked into its ring configuration through its own self-locking mechanism. This involves the tubing end being passed through the orifice of the buckle for the Lap-Band and the suture on the end of the flanged end of the band site being passed through for the Realize Band. Once the band is securely locked in place, the buckle portion of the band is located on the lesser curvature of the stomach (Fig. 27-23A,B). Now the anterior surface of the fundus and proximal stomach is imbricated over the band using several sutures (Fig. 27-24). The tubing of the band system is brought out through the desired site for placement of the port portion of the system. Usually this is a trocar site near the upper abdomen or xiphoid region to place the port most superficially such that it can be palpated postoperatively. The port is secured to the anterior abdominal wall fascia. Access to the port for subsequent addition of fluid to the band system is percutaneously achieved using a Huber or noncutting type needle. The band is initially placed empty of fluid, except priming, in most circumstances.

**Procedure-Specific Complications.** The complications that may occur after LAGB include gastric prolapse, band slippage, band erosion, and port and tubing complications. In addition, failure to lose clinically significant weight is more common following this procedure compared to others. Acute gastric prolapse is the most common emergent complication that requires reoperation after LAGB. Acute, severe pain with immediate dysphagia, vomiting, and inability to take oral food or liquid is the typical presentation. Vomiting may predispose or exacerbate this problem. Either anterior or posterior prolapse may occur.167 The initial evaluation for prolapse involves obtaining a plain film radiograph. If the band is in a horizontal position instead of its normal oblique position, prolapse must be strongly suspected. Initial treatment for an acute or chronic prolapse is to remove all the fluid from the system. This often allows reduction of the prolapse and resolution of symptoms. If symptoms do not resolve after this, an upper gastrointestinal (UGI) series

![Figure 27-23. A. Lap-Band in place around stomach. (Reprinted with permission from Cleveland Clinic Center for Medical Art & Photography © 2005-2009. All Rights Reserved.) B. Realize* (Swedish) Band around stomach.](image-url)
is indicated, and if prolapse persists, then reoperation laparoscopically to reduce the prolapse and resuture the band in place is indicated. Chronic gastric prolapse is more subtle. The band retains its normal oblique angle, but there is symmetric dilation of the gastric pouch above the band. These are initially managed with fluid removal and monitoring of symptoms. Follow-up evaluation can be performed in 4 to 8 weeks, and if the chronic prolapse resolves on UGI, slow refilling of the band may begin.

Band erosion is uncommon, reported in 1% to 2% of most series. The patient usually develops either a port site infection or systemic fever and a low-grade abdominal inflammatory sepsis. Endoscopy can be diagnostic, visualizing the white band material within the stomach. The presence of otherwise unexplained free air on computed tomography (CT) scan should alert the surgeon to this diagnosis as well. Laparoscopic removal of the band is indicated, with repair of the gastric perforation. Often the perforation is already sealed by an inflammatory process, but if not, appropriate management of a gastric perforation must be followed.167

Port and tubing problems occur in at least 5% to 15% of patients undergoing LAGB. These require revision of the port/tubing system due to perforation, leaking, or kinking of the tubing or turning of the port such that access to the surface of the port for adding fluid is precluded. Usually a procedure under local anesthesia is all that is required to repair or realign the tubing or port. The incidence of band removal for patient dissatisfaction or lack of weight loss has been difficult to assess completely, but this number is increasing annually. Angrisani et al.168 reported a 40.9% incidence of band removal after 10-year follow-up. In the Longitudinal Assessment of Bariatric Surgery (LABS) study, 18 subsequent reoperations occurred for every 100 participants with LAGB who were followed up for 3 years.169 Overall, these numbers are expected to increase as follow-up increases.

Biliopancreatic Diversion and Duodenal Switch Background and Patient Selection. BPD was first described by, and remains championed by, Scopinaro in Italy.9 The operation, which is shown in Fig. 27-25, involves resection of the distal half to two-thirds of the stomach and creation of an alimentary tract of the most distal 200 cm of ileum, which is anastomosed to the stomach. The biliopancreatic limb is anastomosed to the alimentary tract at approximately 100 cm proximal to the ileocecal valve. This operation is limited in its utilization due to both technical difficulty and the significant percentage of nutritional complications that arise postoperatively.

One early problem with BPD was the development of a high incidence of marginal ulcers postoperatively. Hess and
Hess and Marceau and colleagues separately described the adaptation of the DS operation, originally proposed by DeMeester and colleagues for treatment of bile reflux gastritis, to replace the gastric portion of the BPD. This procedure was originally called BPD with DS. For ease of description, it is now simply called the duodenal switch (DS) (Fig. 27-26). Currently, BPD and DS represent together less than 1% to 2% of bariatric operations performed in the United States. Patients who undergo either BPD or DS must be prepared for the consequences of a malabsorptive operation. Frequent and large-quantity bowel movement after any large amount of oral intake is common. Also, patients must agree to close follow-up and a large number of vitamin and mineral supplements to avoid nutritional problems. Given the increased incidence of postoperative nutritional and other complications, BPD and DS usually are recommended only for patients who have higher BMIs or for patients who have failed another operation for weight loss or metabolic control. Contraindications to the procedure include geographic distance from the surgeon, lack of financial means to afford supplements, and preexisting calcium, iron, or other nutrient deficiencies.

**Technique.** The technique for BPD and DS is the same for the open and the laparoscopic approach, and they are very technically challenging operations. The BPD operation begins with performance of a distal subtotal gastrectomy with a residual 200-mL gastric pouch. The terminal ileum is identified and divided 250 cm proximal to the ileocecal valve. The distal end of that divided ileum is then anastomosed to the stomach, creating a 2- to 3-cm stoma. The proximal end of the ileum is then anastomosed side-to-side to the terminal ileum approximately 100 cm proximal to the ileocecal valve. Prophylactic cholecystectomy is performed due to the high incidence of gallstone formation with the malabsorption of bile salts.

The DS procedure differs from the BPD procedure only in the proximal gut portion of the operation. Instead of a distal gastrectomy, resection of all the stomach except for a narrow lesser curvature tube of the stomach (SG) is performed. The diameter of this tube is calibrated with a bougie of 32- to 40-French size. The duodenum is then divided in its first portion, leaving an approximately 2-cm length of duodenum intact beyond the pylorus. This end of the duodenum is then anastomosed to the distal 250 cm of ileum. This anastomosis is often done in an end-to-end fashion with a circular stapler. This is the most difficult portion of the DS procedure, and leak rates are slightly higher than with other anastomoses. The distal bowel configuration and cholecystectomy are similar to BPD.

**Procedure-Specific Complications.** Complications that occur after BPD include those seen after RYGB, where intestinal anastomoses and gastric division create potential problems with bleeding and leakage. Scopinaro and colleagues reported GI obstruction in 1.2%, wound infections in 1.2%, and marginal ulcers in 2.8% of patients. However, others found the incidence of marginal ulcer to be higher after BPD, leading to the adoption of the DS. Preservation of the pylorus reduces the incidence of dumping (poorly quantitated in most series) after BPD. The duodenoejejunostomy of DS also has a very low rate of stomal ulcer, unlike the gastrojejunostomy of BPD.

Nutritional complications are by far the most frequent and concerning after both of these operations, particularly on long-term follow-up. Scopinaro and colleagues reported a protein malnutrition rate of 7%, iron deficiency anemia rate of less than 5%, and bone demineralization at 5 years of 53%. Other problems that may arise include alopecia from inadequate protein absorption, night blindness from a lack of vitamin A, and gallstones if the gallbladder is not removed. However, of all these nutritional complications, protein-calorie malnutrition is the most severe and life-threatening. When it is diagnosed, the treatment is parenteral nutrition. Two episodes of required parenteral nutrition are usually considered adequate indication to lengthen the “common channel” of ileum—the ileum between the ileo-jejunostomy of the biliopancreatic limb to the alimentary tract and the ileocolic valve. The amount of length that the surgeon should increase the common channel is poorly documented.

**Investigational Bariatric Procedures**

There is continuous evolution of the approaches to and procedures for bariatric surgery. The goals of this dynamic process are to minimize risk, reduce invasiveness, and maximize clinical effectiveness. This same benefit-risk approach/paradigm has also been adopted by the FDA for the design of clinical trials for obesity devices to help facilitate product development and approval. In the past, a variety of medical devices to assist with weight reduction have been studied, but only a few have been commercially available. In 2012, to address the need for more intermediate treatment options with devices, the FDA initiated a new paradigm based on a benefit-risk determination to suggest appropriate levels of benefit for devices with different risk levels. In other words, it became more feasible to trial less invasive obesity treatment devices, as the threshold for weight loss was lower if the risk of the device or procedure was lower as well. Since that time, several new devices have been approved. These intermediate devices are intended to provide...
tools in the middle of the spectrum of care between lifestyle modification and bariatric surgery and are offered to people with BMIs between 30 and 40 kg/m².

The vagus nerve is known to play a role in satiety, metabolism, and autonomic control in the upper gastrointestinal tract. Studies have been conducted to determine the efficacy of vagal nerve block therapy with a treatment device that consistently delivers at least 12 hours of therapy a day and a sham control device that has no possibility of delivering therapy. A laparoscopic abdominal procedure is performed to attach two electrodes to the anterior and posterior vagal trunks at the level of the gastroesophageal junction. Customized electrodes are placed around the nerves and then secured with sutures. These electrodes are then connected to a transcutaneously rechargeable neuroregulator placed in a subcutaneous pocket on the thoracic side wall. Published results show modest weight loss in the vagal nerve blockade group of 9.2% compared to 6.0% TBWL in the sham group at 12 months and 8.8% and 3.8% TBWL in vagal nerve blockade and sham groups, respectively, at 18 months. More weight loss was sustained than 18 months in the vagal blockade group, and the device was shown to be safe, as there was a low rate of serious complications.

Endoscopically placed intragastric balloons (IGBs) are once again an option for overweight and obese patients with a BMI greater than 27 kg/m². The original Garren-Edwards bubble (GEB) from the late 1980s was an endoscopically placed and removed balloon filled with 220 mL of air that was left in the stomach for 3 months. Adverse events related to the GEB reported in the medical literature included small-bowel obstruction secondary to unplanned deflation, gastric ulcers with GI hemorrhage, and gastric perforation, so its use was abandoned. A multidisciplinary conference that followed recommended that future IGBs should (a) be effective at promoting weight loss, (b) be filled with liquid (not air), (c) be capable of adjustment to various sizes, (d) have a smooth surface with low ulcerogenic and obstructive potential, (e) contain a radiopaque marker, and (f) be constructed of durable materials.

Newer IGBs have undergone evaluation and approval by the FDA. These include both a single and a double lumen balloon, both placed endoscopically and filled with saline. Results from these two pivotal trials show weight losses of 7.6% and 10.2% TBWL at 6 months in the device group that exceeded weight loss in the control or sham groups. There were some early removals in 9% to 18% of subjects for failure to tolerate symptoms, early deflations without migration in 6%, and gastric ulcers in 10%. The precise role for these devices is yet to be determined, and they must be paired with a diet and exercise plan to maximize effectiveness. Repeat or sequential balloon therapy may be effective in enhancing and sustaining weight loss, and it is being studied in Europe. Finally, ensuring proper follow-up is important to reduce adverse events related to ulcers, spontaneous deflation, or migration of the balloon.

An endoscopically placed percutaneous gastrostomy tube is approved for weight loss. It facilitates drainage of approximately 30% of the calories consumed in a meal, in conjunction with lifestyle counseling. In a randomized trial, participants lost 12.1% ± 9.6% TBWL compared to 3.5% ± 6.0% TBWL in the lifestyle-only control group. The most frequent complication was abdominal pain and discomfort in the perioperative period and peristomal granulation tissue and peristomal irritation in the postoperative period. Serious adverse events were reported in 3.6% of participants in the device group.

A duodenal-jejunal bypass liner is an endoscopic device that mimics the duodenal-jejunal exclusion component of an RYGB and is undergoing trials in the United States. Prior studies assessing the efficacy of the DJBL have shown modest weight loss and improvements in glycemic control. There are associated adverse events of migration, obstruction, and epigastric pain. One study demonstrated a high (29%) early device removal rate due to these events. A more recent meta-analysis showed that the DJBL was associated with significant mean differences in TBWL for the device (12.6%) compared with lifestyle modification. The mean differences in glycated hemoglobin and fasting plasma glucose among subjects with T2DM in this meta-analysis did not reach statistical significance.

Various endoscopic and endoluminal procedures are also being utilized as less invasive approaches for bariatric surgical procedures. These include procedures to decrease gastric pouch size and to limit gastrojejunoscopy anastomotic size after “failed” LRYGB. Overall, reports have been disappointing for effectiveness. Gastric plication is also being approached both laparoscopically and endoscopically to mimic results of an SG but without requiring stapling or gastric resection. Further studies with long-term safety and efficacy data are required before these investigational procedures can be considered for routine clinical use.

**FOLLOW-UP AND POSTOPERATIVE CARE**

Postoperative follow-up is required following bariatric surgery to detect and treat postoperative short- and longer-term complications. Weight regain, internal hernias, ulcers, and important nutrient deficiencies can occur years after bariatric surgery. These specific problems are detailed in the “Procedure-Specific Complications” and overall “Complications” sections. The frequency of follow-up varies by surgical procedure and to some extent by surgical practice, but continues, hopefully, for life. Postoperative follow-up is defined as short-term (0–2 years), medium (2–5 years), and long term (≥5 years). Recommendations are that at least 75% of patients are followed for 5 years for LAGB, SG, and LRYGB operations, and 90% are followed closely for 5 years and longer if they have malabsorptive operations (BPD and DS). Although a clinical follow-up system may be in place, it still requires patient compliance, which is generally low for long-term follow-up. In a systematic review, Puzziferri et al also identified that less than 3% of bariatric studies included >80% long-term follow-up. Vigorous efforts can help to improve follow-up, but these require significant staffing and funding. In the NIH-funded prospective, longitudinal bariatric study, more complete follow-up data and weight measurements were obtained for 79% of RYGB patients in the longer term with the use of these resources.

The goals of short-term follow-up are to maximize care of the patient in the postoperative period; assist in adjustment to new eating, exercise, and lifestyle patterns; be on the alert for and treat postoperative complications; and recommend measures to limit such complications. The goals of long-term follow-up are similar, but focus more on weight regain, the management of comorbid condition relapse, and the emergence of recurrent depression, substance and alcohol misuse, and nutritional complications. Vitamin and mineral supplements must be taken regularly for life, including oral supplements for iron, calcium, and vitamin B12, and a multivitamin. Evidence indicates that vitamin and mineral deficiencies, including deficiencies of
calcium, vitamin D, iron, zinc, and copper, are common after bariatric surgery. Guidelines suggest screening patients for iron, vitamin B₁₂, folic acid, and vitamin D deficiencies preoperatively, as well. Patients should also be given daily nutritional supplementation postoperatively, including two adult multivitamin plus mineral supplements (each containing 18 mg of iron, 400 to 800 μg of folic acid, and 50 mg of thiamine), 1200 to 1500 mg of elemental calcium (1800 to 2400 mg for BPD/DS), at least 3000 IU of vitamin D, and vitamin B₁₂, the dose of which varies by route of administration. In addition, all patients should undergo annual screening for specific deficiencies (Table 27-4).

Objective data that should be obtained after all bariatric operations include weight loss, change in BMI, resolution or improvement in medical comorbidities, and any complications that occur. Assessment of quality of life can help gauge efficacy as well, with the Short Form-36 (SF36) questionnaire being one frequently used example. In a retrospective review based on the bariatric outcomes longitudinal database (BOLD) dataset by Spaniolas et al, the effect of postoperative follow-up on 12-month weight loss was studied in 51,081 patients. Complete follow-up was independently associated with excess weight loss ≥50% and total weight loss ≥30%. To identify the relationship between regular follow-up and resolution of comorbidities, the same group studied a cohort of 46,381 patients (31% RYGB patients) who had minimum of 12-month follow-up. After adjusting for baseline characteristics, the group determined that complete follow-up in the first year after RYGB was independently associated with a higher rate of improvement or remission of comorbid conditions (T2DM, hypertension, and dyslipidemia). Frequent and protocolized band adjustments and postoperative support individual/group sessions were shown to be important for longer-term outcomes following LAGB. Finally, the 12-month postoperative visit, which coincides with the plateauing of weight loss for most procedures, presents an opportunity to intervene while bariatric surgery patients are still engaged. Engaging patients and the use of technology to maintain contact with a medical provider are important tools to maintain follow-up after bariatric surgery.

### RESULTS OF BARIATRIC SURGERY

#### Short-Term Outcomes

The short-term (1–2 year) outcomes for bariatric surgical procedures are shown in Table 27-5, which summarizes of the majority of the literature from 2009 to 2017. Average 30-day mortality is low (<1.0%) for all procedures except BPD/DS. Mortality after LRYGB is now consistently less than 0.3% to 0.5% in most large reported series. Morbidity varies by procedure, but it is the lowest for LAGB, followed by SG and then LRYGB, and highest for the malabsorptive procedure BPD/DS. In the Longitudinal Assessment of Bariatric Surgery (LABS) study, a composite endpoint including death, deep-vein thrombosis or venous thromboembolism, reinsertion, or failure to be discharged by 30 days after surgery occurred in 4.8% of those who had undergone LRYGB. Short-term results of the SG have been reported from large national databases. These data show that SG is positioned between LAGB and LRYGB for efficacy of weight loss and resolution of comorbid medical problems and for morbidity and mortality. Few longer-term results with SG have been published.

In the past, large institutional series of LAGB results have been published from centers in Europe and Australia, showing better results for weight loss than those that have been observed in the United States (13–22% TBWL) (see Table 27-5). Weight loss results with BPD or DS are both excellent and comparable but come with higher surgical morbidity. The results from malabsorptive procedures are also very durable for the small percentage of people who undergo them. One 18-year follow-up study after BPD showed a mean excess weight loss of 70% persisting for that duration of time. Although most of the results of BPD or DS are after open operations, one report of laparoscopic DS at an experienced center showed that for 40 patients with an average BMI of 60 kg/m² the mean hospital stay was 4 days, average operation room time was 3.5 hours, and mean excess weight loss at 9 months was 58%. Buchwald and colleagues showed that the average weight loss after BPD and DS in the literature was over 70%, with a mortality rate of 1.1%, a complication rate of 27% to 33%, and a nutritional complication rate of 40% to 77% (see Table 27-5).

<table>
<thead>
<tr>
<th>Table 27-4</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recommended postoperative nutritional monitoring</strong></td>
<td>LAGB</td>
<td>SG</td>
<td>LRYGB</td>
</tr>
<tr>
<td><strong>RECOMMENDATION</strong></td>
<td><strong>LAGB</strong></td>
<td><strong>SG</strong></td>
<td><strong>LRYGB</strong></td>
</tr>
<tr>
<td>Bone density (DXA) at 2 years</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>24 hour urinary calcium excretion at 6 months and annually</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Vitamin B₁₂ annually (methylmalonic acid and homocysteine optional) then every 3–6 months if supplemented</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Folic acid (red blood cell folic acid optional), iron studies, vitamin D, intact parathyroid hormone</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Vitamin A initially and every 6–12 months thereafter</td>
<td>No</td>
<td>No</td>
<td>Optional</td>
</tr>
<tr>
<td>Copper, zinc, and selenium evaluation with specific findings</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Thiamine evaluation with specific findings</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

DXXA = dual energy X-ray absorptiometry; LAGB = laparoscopic adjustable gastric banding; SG = sleeve gastrectomy; LRYGB = laparoscopic Roux-en-Y gastric bypass; BPD/DS = biliopancreatic diversion with duodenal switch.

Further detail on surgical morbidity for each procedure is addressed in both the “Procedure-Specific Complications” and overall “Complications” sections.

**Effectiveness of Bariatric Surgery Compared to Nonsurgical Treatment**

The following section summarizes important findings of studies that compare bariatric procedures with nonsurgical management of obesity. The results of these studies concerning remission from T2DM will be discussed in more detail in “Results of Surgery for Diabetes.” A systematic review and meta-analysis by Gloy summarized all randomized controlled trials (RCTs) that compared bariatric surgery with nonsurgical treatments for obesity.193 The review analyzed 11 trials comprising nearly 800 people with a BMI of 30 to 52. These studies generally focused on cohorts with T2DM and 1 to 2 years of follow-up. They provided good evidence of the effectiveness of bariatric procedures, including LRYGB,78-80 LAGB,77,194 BPD,79 and SG.78 These procedures resulted in greater short-term (1–2 years) weight loss (mean difference −26 kg; 95% CI −31 to −21; \( P < 0.001 \)) and greater remission of T2DM (complete case analysis relative risk of remission: 22.1, 3.2–154.3; \( P = 0.002 \); conservative analysis: 5.3, 1.8–15.8; \( P = 0.003 \)) compared with various nonsurgical treatments.77-80,194 After this meta-analysis, two additional RCTs were published that show similar short-term results for both weight loss and T2DM.83,84

In addition, serum triglycerides and high-density lipoproteins were significantly reduced by bariatric procedures, but blood pressure and other lipoproteins were not (although some studies showed reduced medication use for these conditions).193 The Gloy review also noted a lack of evidence from RCTs beyond 10 years for some endpoints (including mortality). The SOS investigators have published widely on health outcomes beyond 10 years and up to 20 years, including: weight loss, mortality, T2DM remission and incidence, cardiovascular events, incident cancer, psychosocial outcomes, and health care use and costs. Weight loss among surgical subjects in SOS was greater than in control subjects (mean changes in body weight at 2, 10, 15, and 20 years were −23%, −17%, −16%, and −18% in the surgery group and 0%, 1%, −1%, and −1% in the control group). After 15 years, the mean percent weight loss by procedure type was 27 + 12% for RYGB, 18 + 11% for VBG, and 13 + 14% for gastric banding.

The SOS study also showed major improvements in obesity-related comorbidities. In the surgical group, there was a 72% remission of T2DM after 2 years (OR for remission: 8.4) and 36% durable remission after 10 years (OR for remission: 3.5). In spite of the considerable relapse of T2DM over time, bariatric surgery was associated with a lower incidence of myocardial infarction and other T2DM complications. The SOS study demonstrated that bariatric surgery also reduced the risk of incident T2DM by 96%, 84%, and 78% after 2, 10, and 15 years among subjects without the condition at baseline. The SOS study also

### Table 27-5

**Short-term bariatric surgical outcomes**

<table>
<thead>
<tr>
<th>OUTCOME&lt;sup&gt;a&lt;/sup&gt;</th>
<th>ALL</th>
<th>LRYGB</th>
<th>SG</th>
<th>LAGB</th>
<th>BPD/DS</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Mortality 30-day</td>
<td>&lt;1.0%</td>
<td>0.3–0.5%</td>
<td>0.11%</td>
<td>0.05%</td>
<td>1.1%</td>
</tr>
<tr>
<td>% Morbidity 30-day</td>
<td>NA, depends on procedure</td>
<td>12–21%</td>
<td>3–6%</td>
<td>2–4%</td>
<td>27–33%</td>
</tr>
<tr>
<td>% Total body weight loss (TBWL)</td>
<td>NA, depends on procedure</td>
<td>31–36%</td>
<td>25–30%</td>
<td>13–22%</td>
<td>36–38%</td>
</tr>
<tr>
<td>% Excess body weight loss (EBWL)</td>
<td>NA, depends on procedure</td>
<td>48–77%</td>
<td>49–81%</td>
<td>29–50%</td>
<td>&gt;70%</td>
</tr>
<tr>
<td>% Diabetes remission</td>
<td>77%</td>
<td>60–80%</td>
<td>60%</td>
<td>35%</td>
<td>75%</td>
</tr>
<tr>
<td>% Dyslipidemia&lt;sup&gt;b&lt;/sup&gt; remission</td>
<td>70%</td>
<td>63–91%</td>
<td>72–82%</td>
<td>78%</td>
<td>80%</td>
</tr>
<tr>
<td>% Hypertension remission</td>
<td>62%</td>
<td>61–81%</td>
<td>60–92%</td>
<td>43%</td>
<td>60%, few reports</td>
</tr>
<tr>
<td>% Sleep apnea remission&lt;sup&gt;c&lt;/sup&gt;</td>
<td>84%</td>
<td>80%</td>
<td>80%</td>
<td>68%</td>
<td>80%</td>
</tr>
</tbody>
</table>

<sup>a</sup>Outcomes are the averages found in the literature at 1 to 2 years postoperatively, unless otherwise noted.

<sup>b</sup>Denotes any component lipid remission.

<sup>c</sup>Denotes clinical remission as repeat sleep studies are uncommonly performed.

### Longer-Term Studies

A summary of studies with long-term outcomes are shown in Table 27-6. The following section describes these studies and other larger studies that have contributed data to the growing body of evidence with respect to some short-term and now much longer-term outcomes.

**Swedish Obese Subjects Study.** Much of what is currently known about the long-term results of bariatric surgery come from the Swedish Obese Subjects (SOS) study, which was initiated in 1987 as a prospective trial of 2010 subjects undergoing bariatric surgery compared to a usual care control group (\( n = 2037 \)) that were matched on 18 clinical and demographic variables. The most common bariatric procedure performed in SOS was the VBG (68%), followed by gastric banding (19%), and RYGB (13%). Follow-up rates are high and reported at 99% for some endpoints (including mortality). The SOS investigators have published widely on health outcomes beyond 10 years and up to 20 years, including: weight loss, mortality, T2DM remission and incidence, cardiovascular events, incident cancer, psychosocial outcomes, and health care use and costs. Weight loss among surgical subjects in SOS was greater than in control subjects (mean changes in body weight at 2, 10, 15, and 20 years were −23%, −17%, −16%, and −18% in the surgery group and 0%, 1%, −1%, and −1% in the control group). After 15 years, the mean percent weight loss by procedure type was 27 + 12% for RYGB, 18 + 11% for VBG, and 13 + 14% for gastric banding.

The SOS study also showed major improvements in obesity-related comorbidities. In the surgical group, there was a 72% remission of T2DM after 2 years (OR for remission: 8.4) and 36% durable remission after 10 years (OR for remission: 3.5). In spite of the considerable relapse of T2DM over time, bariatric surgery was associated with a lower incidence of myocardial infarction and other T2DM complications. The SOS study demonstrated that bariatric surgery also reduced the risk of incident T2DM by 96%, 84%, and 78% after 2, 10, and 15 years among subjects without the condition at baseline. The SOS study also
### Table 27-6: Long-term studies of bariatric surgery outcomes

<table>
<thead>
<tr>
<th>Study Design</th>
<th>Populations and Procedures</th>
<th>Follow-Up Duration</th>
<th>Published Outcomes</th>
<th>Mortality and Survival</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prospective observational</td>
<td>2010 Surgical cases (13%</td>
<td>10–20 years,</td>
<td>Surgery was associated with: greater weight loss at 2 years (−23% vs. 0%) and at 20</td>
<td>Bariatric surgery treatment: 16 years, 29% lower risk of death from any cause (hazard</td>
<td>Not randomized; includes mostly procedures that are no longer in use;</td>
</tr>
<tr>
<td>with matched controls</td>
<td>RYGB, 19% banding; 68%</td>
<td>depending on the report</td>
<td>y (−18% vs. −1%)16; greater remission of T2DM after 2 y (OR for remission, 8.4; P &lt;.001) and 10 y (OR, 3.5; P &lt;.001); lower incidence of T2DM (HR, 0.17; P &lt;.001)</td>
<td>ratio 0.71, 0.54 to 0.92; P = 0.01) vs. usual care; common causes of death: myocardial infarction (HR, 0.71; P = .02), stroke (HR, 0.66; P = .008), and cancer (in women only; HR, 0.58; P &lt;.001)</td>
<td>patients from a single country with little racial/ethnic diversity.</td>
</tr>
<tr>
<td></td>
<td>VBG) and 2037 matched</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retrospective observational</td>
<td>7925 RYGB cases and 7925</td>
<td>Mean, 7.1 years</td>
<td>Only mortality outcomes reported</td>
<td>Bariatric surgery treatment: average 7.1 years post treatment, 40% reduction in all cause mortality (hazard ratio HR 0.60, 0.45 to 0.67; P &lt;.0001), 49% (HR 0.51, 0.36 to 0.73; P &lt;.0001), and 92% (HR 0.08, 0.01 to 0.47; P = 0.005), cardiovascular mortality, and T2DM mortality, respectively</td>
<td>Not randomized; matching based on self-reported height and weight from driver's license database; includes only RYGB procedures; patients from a single state.</td>
</tr>
<tr>
<td>with matched controls</td>
<td>weight-matched controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prospective observational</td>
<td>418 RYGB cases; 417</td>
<td>6 years, 12 years</td>
<td>6 years: RYGB group lost 27.7% body weight compared with 0.2% weight gain in control group 1 and 0% change in control group 2; T2DM remission in 62% of RYGB patients and 8% and 6% in each of the control groups (P &lt;.001); incident T2DM was observed in 2% of RYGB patients but 17% and 15% of each of the control groups at 6 years (P &lt;.001); surgery associated with greater improvements in blood pressure, cholesterol, and quality of life (P &lt;.01)</td>
<td>Deaths at 6 years: 12 (2.8%), 14 (3.3%), and 3 (0.93%) for bariatric surgery, control 1, and control 2, respectively</td>
<td>Not randomized; includes only RYGB procedures; patients from a single state.</td>
</tr>
<tr>
<td>with two matched control</td>
<td>bariatric-surgery seekers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>groups</td>
<td>who did not undergo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>operation (control 1);</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>321 population-based</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>matched controls (control 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source</td>
<td>Study Design</td>
<td>Participants</td>
<td>Follow-up</td>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>-------------------------------</td>
<td>------------------------------------------------------------</td>
<td>-----------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Maciejewski et al., 2011, 2012, 2015, 2016 (Department of Veterans Affairs)</td>
<td>Retrospective observational with matched controls</td>
<td>847 to 1787 to 2500 surgical cases and their matched controls</td>
<td>6.7 years</td>
<td>Patients undergoing RYGB lost 28.6% (95% CI, 19.5%–37.6%) of their baseline weight at 10 years, whereas nonsurgical matches lost 7.3% (95% CI, 1.4%–13.3%) of their baseline weight at 10 years. Patients undergoing RYGB lost 21% (95% CI, 11%–31%) more of their baseline weight at 10 years than nonsurgical matches. A total of 405 of 564 patients undergoing RYGB (71.8%) had more than 20% body weight loss and 224 of 564 (39.7%) had more than 30% estimated weight loss at 10 years. At 4 years, patients undergoing LRYGB lost 27.5% (95% CI, 23.8%–31.2%) of their baseline weight, patients undergoing LAGB lost 10.6% (95% CI, 0.6%–20.6%), and patients undergoing SG lost 17.8% (95% CI, 9.7%–25.9%). Surgery was not significantly associated with lower health expenditures 3 years after the procedure, in first study.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>First study, in 2011, bariatric surgery not significantly associated with reduced mortality. Later study, at the end of 14 years, 263 deaths in the surgical group (mean follow-up, 6.9 years) and 1277 deaths in control group (mean follow-up, 6.6 years). Mortality rates were 2.4% at 1 year, 6.4% at 5 years, and 13.8% at 10 years for surgical patients; for matched control patients, 1.7% at 1 year, 10.4% at 5 years, and 23.9% at 10 years. Significantly lower mortality after 1 to 5 years (HR, 0.45 [95% CI, 0.36–0.56]) and 5 to 14 years (HR, 0.47 [95% CI, 0.39–0.58]).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Not randomized; includes older (mean age, 55 years), primarily male (74%) veterans; mortality studies mostly RYGB procedures (Continued)</td>
<td></td>
</tr>
</tbody>
</table>

12 years: RYGB 26.9% body weight loss, 2.0% and 0.9% in control groups 1, 2, respectively; T2DM remission in 51% RYGB group; odds ratio for the incidence T2DM 0.08 (95% CI, 0.03 to 0.24) for RYGB vs. control group 1 and 0.09 (95% CI, 0.03 to 0.29) RYGB vs. control group 2 ($P <0.001$ for both comparisons); RYGB group had higher remission rates and lower incidence rates of hypertension and dyslipidemia than did control group 1 ($P <0.05$ for all comparisons).
<table>
<thead>
<tr>
<th>AUTHOR STUDY DESIGN</th>
<th>POPULATIONS AND PROCEDURES</th>
<th>FOLLOW-UP DURATION</th>
<th>PUBLISHED OUTCOMES</th>
<th>MORTALITY AND SURVIVAL</th>
<th>LIMITATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Courcoulas et al, 2009, 2013, 2017 (Longitudinal Assessment of Bariatric Surgery [LABS])</td>
<td>4776 in LABS-1, 30 day safety study and 2458 in LABS-2, effectiveness study (70.7% RYGB; 24.8% LAGB; and 5% other procedures)</td>
<td>30 days, 3 years, 7 years</td>
<td>3 year: 31.5% for RYGB and 15.9% for LAGB; T2DM remission in 67.5% of RYGB cases and 28.6% for LAGB; incidence of T2DM was 0.9% after RYGB and 3.2% after LAGB. Dyslipidemia remission in 61.9% RYGB cases and 27.1% AGB cases; HTN remission in 38.2% RYGB cases and 17.4% AGB cases; other procedures’ results not reported</td>
<td>30 days: 0.3% overall; 0% LAGB, 0.2% LRYGB, and 2.1% open RYGB. 3 years: 0.9 per 300 person-years for RYGB and 0.8 per 300 person-years for LAGB, i.e., number of events if 100 people were followed for 3 years 7 years: 3.7 per 700 person-years for RYGB and 2.7 per 700 person years for LAGB, i.e., number of events if 100 people were followed for 7 years</td>
<td>Not randomized; lacks nonsurgical control population; primarily RYGB and LAGB procedures; high follow-up rates (&gt;80% for weight) but some not in-person</td>
</tr>
<tr>
<td>Arterburn et al, 2013 (HMO Research Network)</td>
<td>4434 RYGB cases with T2DM</td>
<td>Median, 3.1 years</td>
<td>68% of patients (95% CI, 66–70) experienced an initial T2DM remission within 5 years after RYGB; among these, 35.1% (95% CI, 32–38) redeveloped T2DM within 5 years; median duration of T2DM remission, 8.3 years</td>
<td>Not reported</td>
<td>Not randomized; lacks nonsurgical control population; only RYGB procedures</td>
</tr>
<tr>
<td>Carlin et al, 2013 (Michigan Bariatric Surgery Collaborative)</td>
<td>8847 to 35,477, varies depending on publication</td>
<td>30 days to 3 years, varies depending on publication</td>
<td>Complication rates for SG (6.3%) were significantly lower than for RYGB (10.0%; P &lt;.001) but higher than LAGB (2.4%; P &lt;.001). Excess body weight loss at 1 year was 13% lower for SG (60%) than for RYGB (69%; P &lt;.001) but was 77% higher for SG than for LAGB (34%; P &lt;.001).</td>
<td>Not reported</td>
<td>Not randomized; lacks nonsurgical control; patients from a single state</td>
</tr>
</tbody>
</table>

found that bariatric surgery was associated with a reduced incidence of fatal or nonfatal cancer among women but not in men. Finally, at 16 years follow-up, surgery was associated with a 29% lower risk of death (the primary endpoint of the study) from any cause compared to usual care, with the most common causes of death being cancer and myocardial infarction\textsuperscript{196-201} (see Table 27-6).

**Utah Obesity Studies.** Another important long-term observational study performed in Utah from 1984 to 2002 included 7925 people who had undergone RYGB and 7925 weight, age, and sex matched controls. This study showed a 40% reduction in all-cause mortality (hazard ratio 0.60, 0.46 to 0.77; \( P < 0.001 \)) and a 49% (0.51, 0.36 to 0.73; \( P < 0.001 \)) and 92% (0.08, 0.01 to 0.47; \( P = 0.005 \)) reduction in death from cardiovascular disease and death related to T2DM, respectively, at an average of 7.1 years later.\textsuperscript{202}

A separate ongoing prospective Utah Obesity Study involving over 400 RYGB cases and two nonrandomized, matched control groups—each with over 400 and 300 severely obese subjects (one group were people seeking surgery that did not undergo operation; the other was a population-based severely obese control group)—has found that the surgery group lost 27.7% of their initial body weight compared to 0.2% weight gain in control group 1 (surgery seekers), and 0% change in control group 2 (population-based control) at 6 years. Diabetes was in remission in 62% of RYGB patients and only 8% and 6% in each of the control groups, while incident T2DM was observed in 2% of RYGB patients and in 17% and 15% of the control groups at 6 years.\textsuperscript{203} The 12-year follow-up results were also recently published and showed long-term durability of weight loss and effective remission and prevention of T2DM, hypertension, and dyslipidemia after RYGB.\textsuperscript{204} Follow-up rates in this study were high at over 90% at 12 years. The mean percent change from baseline in body weight in the RYGB group was \(-26.9\% \) at 12 years compared to \(-2.0\% \) and 0.9% in each of the two control groups. There was remission of T2DM in 43 of 84 patients (51%) at 12 years. The OR for the incidence of T2DM at 12 years was 0.08 (95% CI, 0.03–0.24) for the RYGB group versus control group 1 and 0.09 (95% CI, 0.03–0.29) for the RYGB group versus control group 2 (\( P < 0.001 \) for both comparisons). The RYGB group had higher remission rates and lower incidence rates of hypertension and dyslipidemia than did control group 1 (\( P < 0.05 \) for all comparisons) (see Table 27-6).

**Veteran’s Administration Study.** A retrospective observational study involving 847 U.S. veterans who were older and more high risk than in typical bariatric studies, found no significant association between bariatric surgery and survival compared to usual care at mean 6.7 years follow-up.\textsuperscript{205} When matched to control subjects, the outpatient, inpatient, and total expenditures were higher for bariatric surgical cases in the 3 years leading up to the procedure and then went back to the lower cost levels of nonsurgical controls in 3 years after the procedure. The conclusion from this study was that bariatric surgery did not appear to be associated with reduced health care expenditures 3 years after the procedure.\textsuperscript{206}

In a follow-up retrospective cohort study, 2500 U.S. veterans (74% men) who underwent bariatric surgery (74% gastric bypass, 15% sleeve gastrectomy, 10% adjustable gastric banding, and 1% other) were matched to 7462 control patients. The primary outcome was all-cause mortality. In this study, surgical patients (\( n = 2500 \)) had a mean age of 52 years and a mean BMI of 47. Matched control patients (\( n = 7462 \)) had a mean age of 53 years and a mean BMI of 46. At the end of the 14-year study period, there were a total of 263 deaths in the surgical group (mean follow-up, 6.9 years) and 1277 deaths in the matched control group (mean follow-up, 6.6 years). Mortality rates were 2.4% at 1 year, 6.4% at 5 years, and 13.8% at 10 years for surgical patients; for matched control patients, 1.7% at 1 year, 10.4% at 5 years, and 23.9% at 10 years. So, in this later study, there was significantly lower all-cause mortality at longer follow-up.\textsuperscript{207}

Ten-year weight change in 1787 veterans who underwent RYGB compared to controls, and separately, 4-year weight change in veterans who underwent RYGB (\( n = 1785 \)), SG (\( n = 379 \)), and AGB (\( n = 246 \)) were reported. Patients undergoing RYGB lost 21% more of their baseline weight at 10 years than nonsurgical matches. A total of 405 of 564 patients undergoing RYGB (71.8%) had more than 20% weight loss, and 224 of 564 (39.7%) had more than 30% weight loss at 10 years compared with 134 of 1247 (10.8%) and 48 of 1247 (3.9%), respectively, for nonsurgical matches. At 4 years, patients undergoing LRYGB lost 27.5% of their baseline weight, patients undergoing LAGB lost 16.6%, and patients undergoing SG lost 17.8%. Patients undergoing RYGB lost 16.9% more of their baseline weight than patients undergoing AGB and 9.7% more than patients undergoing SG\textsuperscript{208} (see Table 27-6).

**The Longitudinal Assessment of Bariatric Surgery Study.** The Longitudinal Assessment of Bariatric Surgery (LABS-1) study, a multicenter observational surgical cohort, prospectively assessed 30-day safety among 4776 severely obese patients who underwent a first bariatric surgical procedure (25% AGB, 62% laparoscopic RYGB, 9% open RYGB, and 3% another procedure) between 2005 and 2007.\textsuperscript{209} The 30-day mortality in the LABS study was 0.3% for all procedures with a major adverse outcome rate (a predefined composite endpoint that included death, venous thromboembolism, reintervention [percutaneous, endoscopic, or operative], or failure to be discharged from the hospital in 30 days) of 4.1%. These results did vary by procedure and approach, with no mortality in the 1198 patients who had undergone LAGB, 0.2% of the 2975 patients who had undergone LRYGB, and 2.1% of the 437 patients who had undergone open RYGB. Similarly, the rate of adverse outcomes (morbidty) occurred in 4.1% of patients overall; 1.0% for LAGB, 4.8% for LRYGB, and 7.8% for open RYGB.\textsuperscript{209} The Longitudinal Assessment of Bariatric Surgery (LABS-2) study is another large prospective multicenter observational bariatric cohort study that was not randomized and did not include a nonsurgical control group. LABS-2 assessed weight change and comorbid conditions in 2458 participants (1738 RYGB—both open and laparoscopic, 610 LAGB, and 110 other procedures) recruited between 2005 and 2009 who were followed for 7 years.\textsuperscript{209} At baseline, 33% had diabetes, 63% had dyslipidemia, and 68% had hypertension. In the LABS-2 cohort, median weight change was 31.5% for RYGB and 15.9% for adjustable gastric banding after 3 years, with much variability in response to each surgical treatment. Remission of T2DM was noted in 67% and 28% of those who had undergone RYGB and LAGB, respectively. The incidence T2DM was 0.9% and 3.2%, respectively, over the 3 years (see Table 27-6). LABS-2 looked at both pre- and postoperative predictors of weight change and found that very few of many baseline variables studied (Black race, T2DM) were associated
with 3-year weight change, and the effects were small overall. Postoperatively, for RYGB only, three behaviors explained most of the variability (16%) in 3-year weight change: weekly self-weighing, continuing to eat when feeling full more than once a week, and eating continuously during the day. If a person started weekly self-weighing, stopped eating when feeling full, and stopped eating continuously during the day, they lost 14% more weight than those who made no positive changes (38.8% vs. 24.6% TBWL).\(^{200,211}\)

At 7 years of follow-up, in LABS, data completeness for weight was high (83%), and the median weight change was 28.4% for RYGB and 14.9% for LAGB. Weight regain between years 3 and 7 was 3.9% of baseline weight for RYGB and 1.4% for LAGB, but 75% of RYGB participants maintained at least 20% total body weight loss, and 50% of LAGB participants maintained at least 16% through 7 years. Remission of T2DM was reported in 60.2% of RYGB cases and 20.3% for LAGB. Mortality was reported as 3.7 per 700 person-years for RYGB and 2.7 per 700 person years for LAGB, i.e., number of events if 100 people were followed for 7 years. Reoperations were also much more common after LAGB compared to RYGB at 7 years (see Table 27-6).

**HMO Research Network.** Arterburn and colleagues have leveraged the integrated health network system to study bariatric outcomes using the electronic health/medical record. They studied clinical predictors of diabetes remission and relapse among patients undergoing gastric bypass. Theirs was a retrospective cohort study of adults with uncontrolled or medication-controlled T2DM who underwent gastric bypass in three integrated health care delivery systems in the United States. Remission and relapse events were defined by diabetes medication use and clinical laboratory measures of glycemic control. Of 4434 adults with T2DM who underwent RYGB, 68.2% (95% CI, 66% and 70%) experienced an initial complete diabetes remission within 5 years after surgery. Among these, 35.1% (95% CI, 32% and 38%) relapsed back to T2DM within 5 years. The median duration of their remission was 8.3 years. Predictors of incomplete remission and relapse were poor preoperative glycemic control, insulin use, and longer diabetes duration\(^{212}\) (see Table 27-6).

In a second study, they compared rates of diabetes remission, relapse, and all-cause mortality at 2 years between severely obese adults with T2DM who underwent bariatric surgery or received usual medical care. There were 1395 adults with T2DM who had bariatric surgery and 62,322 who did not. Most procedures were RYGB (72.0% laparoscopic; 8.2% open); 44.4% were gastric banding, 2.4% were sleeve gastrectomy, and 13.2% were other procedures. At 2 years, bariatric subjects experienced significantly higher diabetes remission rates (73.7% [95% CI: 70.6, 76.5]) compared to nonsurgical subjects (69.9% [95% CI: 6.9, 7.1]). Age, site, duration of diabetes, hemoglobin A1c level, and intensity of diabetes medication treatment were significantly associated with remission. Bariatric subjects also experienced lower relapse rates than nonsurgical subjects (adjusted HR: 0.19; 95% CI: 0.15–0.23) with no higher risk of death (adjusted HR: 0.54; 95% CI: 0.22–1.30) (see Table 27-6).\(^{213}\) This group also studied short-term comparative effectiveness outcomes between procedures (LRYGB and LAGB) and found that LRYGB resulted in much greater weight loss than LAGB but had a higher risk of short-term complications and long-term subsequent hospitalizations.\(^{214}\)

**Michigan Bariatric Surgery Collaborative.** The Michigan Bariatric Surgery Collaborative is a statewide consortium of hospitals and surgeons that maintains an externally audited prospective clinical registry. The comparative effectiveness of SG, LRYGB, and LAGB procedures was studied in this dataset. Nearly 3000 SG patients with equal numbers of RYGB and LAGB patients were matched on 23 baseline characteristics. Outcomes assessed included 30-day complications, weight loss, quality of life, and comorbid remission up to 3 years after bariatric surgery. Overall complication rates for SG (6.3%) were significantly lower than for RYGB (10%) but higher than for LAGB (2.4%). Serious complication rates were similar for SG (2.4%) and LRYGB (2.5%) but higher than for LAGB (1.0%). Excess body weight loss at 1 year was 13% lower for SG (60%) than for RYGB (69%), but was 77% higher for SG than for LAGB (34%). Remission of comorbid conditions was similar between SG and LRYGB\(^{215}\) (see Table 27-6).

This group also developed a risk prediction model for serious 30-day complications after bariatric surgery. Overall, 2.5% of patients experienced a serious complication. Significant risk factors included prior venous thromboembolism (OR 1.90, CI 1.41–2.54); mobility limitations (OR 1.61, CI 1.23–2.13); coronary artery disease (OR 1.53, CI 1.17–2.02); age over 50 (OR 1.38, CI 1.18–1.61); pulmonary disease (OR 1.37, CI 1.15–1.64); male gender (OR 1.26, CI 1.06–1.50); smoking history (OR 1.20, CI 1.02–1.40); and procedure type.\(^{216}\) Further, to assess the relationship between IVC filter insertion and complications while controlling for differences in baseline patient characteristics and medical venous thromboembolism prophylaxis, this group published an additional study, and 35,477 patients from 32 hospitals in Michigan were included. Patients receiving IVC filters had higher rates of pulmonary embolism, deep vein thrombosis, venous thromboembolism, serious complications, and death.\(^{217}\)

Recently, this group has also been evaluating the effect of surgical skill and operative technique on complications following bariatric surgery.\(^{218,219}\)

**Other Studies**

**Metabolic and Bariatric Surgery Quality Improvement Program (MBASQIP).** This is a prospective, multi-institutional, national database that has been used to compare SG to RYGB and LAGB. The study from 2011 was short term and compared 30-day, 6-month, and 1-year outcomes including morbidity and mortality, readmissions, and reoperations as well as reduction in BMI and weight-related comorbid conditions. The findings were that SG has higher risk-adjusted morbidity, readmission and reoperation/intervention rates compared to LAGB, but lower reoperation/intervention rates compared to RYGB either laparoscopic or open. There were no differences in mortality. Reduction in BMI and most of the weight-related comorbidities after SG was also between LAGB and RYGB rates.\(^{140}\)

A later study addressed the impact of various SG techniques on short-term (30-day) outcomes. Using the MBASQIP data registry, 189,477 SG operations that were performed at over 700 centers in the United States were analyzed. Cases in which staple line reinforcement was used were associated with higher leak rates and lower bleeding rates. Bougie size ≥38 French was associated with significantly lower leak rates compared to <38 French.\(^{162}\) Longer-term data will eventually be available from this national dataset, but the completeness of follow-up has not yet been determined.

**Geisinger Health System.** This is an electronic medical record database in a large rural integrated health system. They
have published long-term results (7–12 years) of the percentage of TBWL and preoperative predictors for LRYGB in approximately 700 patients. Over 200 preoperative clinical factors were studied. At a median of 9.3 postoperative years following surgery, the mean (SD) percentage TBWL was 22.5% (13.1%). Preoperative insulin use, history of smoking, and use of 12 or more medications before surgery were associated with greater long-term weight loss; 6.8%, 2.8%, and 3.1%, respectively. Preoperative hyperlipidemia, older age, and higher body mass index were associated with poorer long-term weight loss (−2.8%, −8.8%, and −4.1%, respectively). Again, there were only a few preoperative clinical factors associated with differences in long-term weight loss after LRYGB.220

This group also developed a method to predict the probability of T2DM remission after RYGB surgery on the basis of preoperative clinical criteria in a retrospective cohort study. Over 200 clinical variables were used to identify independent predictors of remission within 5 years and to produce a score (DiaRem) to assess this likelihood. Records were available for 690 patients in the primary cohort, of whom 463 (63%) had achieved partial or complete T2DM remission. Four preoperative clinical variables were included in the final model: insulin use, age, HbA1c concentration, and type of antidiabetic drugs. The DiaRem score was developed from that, and it ranges from 0 to 22 with the proportion of patients achieving remission being highest for the lowest scores.221

**Comparisons Between Procedures**

There have been many systematic reviews of bariatric surgery attempting to summarize and quantify differences in the efficacy and safety of the different surgical procedures. A major challenge in summarizing this literature from the last 10 years is the fact that no single randomized trial has included all of the most common procedures (RYGB, LAGB, SG, and BPD/DS), so inference must be made through pooled analysis of data from many disparate randomized and non-randomized studies of bariatric surgery with different lengths and completeness of follow-up. There are also no studies that have examined differences in long-term survival, incident cardiovascular events, and quality of life across bariatric procedures.1 Still one of the most comprehensive systematic reviews by Buchwald included 136 studies and a total of 22,094 bariatric patients. Only 5 of the included studies were randomized trials (28 non-RCTs and 101 uncontrolled case series) and the review did not include any data on the SG procedure, so will need to be updated. This review found a strong trend towards different weight loss outcomes across procedures: weighted mean percentage of excess weight loss (%EWL) 50% for LAGB; 68% for RYGB; 69% for VBG; and 72% for BPD/DS. The rate of T2DM remission also appeared to differ across procedures: 48% for LAGB; 84% for RYGB; 72% for VBG; and 99% for BPD/DS. A similar pattern of disease remission was observed for dyslipidemia, hypertension, and obstructive sleep apnea, with the greatest rates of remission observed among BPD/DS patients, followed by RYGB patients, with the least disease remission among LAGB patients.192

There is still ongoing debate regarding the comparative effectiveness of the three most common procedures currently in use: LRYGB, SG, and LAGB. Several other systematic reviews have concluded that LRYGB is more effective for weight loss than LAGB; however, there have been only two small RCTs with follow-up at 4 and 5 years addressing this issue specifically.65,222,223 There is evolving data from a number of smaller RCTs to examine differences between LRYGB, LAGB, and SG for comorbidity improvement (addressed in “Results of Surgery for Diabetes”), but systematic reviews of nonrandomized studies indicate greater remission of T2DM, dyslipidemia, hypertension, and sleep apnea with LRYGB compared to LAGB. Two recent systematic reviews have compared the outcomes of the SG with other procedures.224,225 One review identified 15 RCTs involving 1191 patients. The percent excess body weight loss (%EBWL) ranged from 49% to 81% for SG, from 62% to 94% for LRYGB, and from 29% to 48% for LAGB, with a follow-up ranging from 6 months to 3 years. The T2DM remission rate ranged from 27% to 75% for SG vs. 42% to 93% for RYGB. The second review only compared SG to RYGB and identified 6 RCTs and two nonrandomized controlled studies with follow-up ranging from 3 months to 2 years. They found that LRYGB achieved significantly greater improvement in BMI than SG (1.8 kg/m²) and greater improvements in metabolic factors. Longer-term comparative effectiveness data on SG are still needed, but the effectiveness of the SG procedure, again, appears to be positioned between the LRYGB and LAGB procedures.

**Resolution of Specific Comorbid Conditions**

Bariatric surgery can improve and induce remission of many obesity-related comorbid conditions. Nevertheless, the remission rates can decline over time due to relapse of disease, and as follow-up lengths, complete and more longer-term follow-up data is needed in some areas.

**Cardiovascular Disease.** A recent systematic review of long-term cardiovascular risk factor reduction after bariatric surgery involved 73 studies and 19,543 subjects with a mean age of 42 years; 76% of subjects were female, and 44%, 24%, and 44% had baseline hypertension, diabetes, and hyperlipidemia, respectively.226 At a mean follow-up of 57.8 months, the average excess weight loss for all bariatric procedures was 54%, and remission/improvement was 63% for hypertension, 73% for T2DM, and 65% for hyperlipidemia. Echocardiographic results from 713 subjects showed statistically significant improvements in hemodynamics. There are no long-term RCTs comparing bariatric surgery with nonsurgical medical treatment of obesity that specifically evaluate cardiovascular endpoints and cardiovascular mortality. However, 12 cohort-matched studies comparing bariatric surgery with nonsurgical controls have been reviewed.227 Collectively, all but two of these studies support a reduced cardiovascular event rate and all-cause mortality rate conferred by bariatric surgery. Of these studies, the Swedish Obesity Subjects (SOS) study still has the longest outcomes.

**Gastroesophageal Reflux Disease.** Patients with obesity and GERD have a higher chance of failing to obtain symptomatic relief from standard antireflux surgery. The recurrence of symptoms is higher, likely due to a higher incidence of wrap herniation into the mediastinum and other mechanical failure of the fundoplication, which in turn is likely affected by the increased intra-abdominal pressure of the obese condition. The patient with a BMI over 35 kg/m² who has GERD has a better chance of symptom improvement by undergoing LRYGB, which is effective for the treatment of GERD.228,229 LRYGB creates such a small gastric pouch that it has a very limited volume for acid production. LAGB may worsen or may improve GERD but to a considerably lesser extent than RYGB. A prospective analysis of 558 consecutive SG (n = 200) and LRYGB (n = 358) patients demonstrated significantly improved subjective GERD symptoms in the bypass cohort when compared to the SG patients at
1 year. Studies show that SG can increase GERD symptoms postoperatively.

**Obstructive Sleep Apnea.** A systematic review of 13,900 patients (69 studies) showed significant improvement or resolution of sleep apnea in more than 75% of bariatric surgery patients. Comparison of outcomes between procedures demonstrated the most benefit with BPD and RYGB and the least with LAGB. However, a randomized control trial comparing the effect of medical and surgical weight loss (LAGB) on sleep apnea found no significant difference in apnea events despite major differences in weight loss. The findings suggested that much of the improvement achieved was in the mild to moderate weight loss range, with little benefit of further weight loss.232

**Asthma.** Another pulmonary symptom that commonly occurs in severely obese patients is asthma. Dixon and colleagues233 studied 23 asthmatic patients who underwent bariatric surgery and found a significant improvement in asthma control (e.g., forced expiratory volume in 1 second, forced vital capacity), asthma-related quality of life, and responsiveness to methacholine. Boulet and colleagues234 found similar results in their cohort of 12 patients with asthma who experienced significant weight loss after bariatric surgery.

**Nonalcoholic Fatty Liver Disease.** Nonalcoholic Fatty Liver Disease (NAFLD) is a metabolically related problem associated with obesity. The disease is a spectrum of liver abnormalities including steatosis, steatohepatitis, fibrosis, and cirrhosis of the liver. It is estimated that 20% of U.S. adults have NAFLD, largely because of the high incidence of obesity. NAFLD is present in an estimated 85% of patients with severe obesity.235 Although further research is needed to accurately assess the role of bariatric surgery as a potential treatment for NAFLD, there are some reports that support its use. A systematic review of the available literature found many retrospective and prospective observational cohort studies, but no RCTs or case-control series.

**Musculoskeletal Disease.** Degenerative joint disease and low back pain are among the most common complaints and associated comorbid problems in the severely obese population. A prospective cohort of 50 obese females age 20 to 74 years were followed for 1 year after LRYGB using the timed-get-up-and-go (TGUG) and health survey SF-36.237 The results showed a significant improvement in musculoskeletal function and likely enhanced ability to progress in rehabilitation. Patients with osteoarthritis of the neck, shoulder, spine, hip, knee, ankle, wrist, and hand have been shown to have improved or resolved joint pain after bariatric surgery. Reduction in BMI values of 6.2 to 14.7 kg/m² has corresponded with back and knee pain resolution in 5% to 100% of patients, whereas pain severity was reduced in 31% to 94% of patients depending on the joint and study.238

The LABS-2 Study published data on pain and physical function in over 2200 participants. At year 1, clinically meaningful improvements were shown in 57.6% of patients for bodily pain, 76.5% for physical function, and 59.5% for walk time. Additionally, among participants with severe knee or disability (633), or hip pain or disability (500) at baseline, approximately three-fourths experienced joint-specific improvements in knee pain (77.1%) and in hip function (79.2%). But between year 1 and year 3, rates of improvement significantly decreased in knee pain (77.1%) and in hip function (79.2%). However, three studies of 6 to 10 years’ duration suggest that bariatric procedures are associated with greater improvements in overall and obesity specific measures of quality of life compared to medical treatment or care.240,241 Physical functioning aspects of quality of life seem to be more responsive to bariatric procedures than mental health domains, although more research is needed, especially in patients with less severe (class 1) obesity.

**Results of Surgery for Diabetes (Metabolic Surgery)**

Based on an abundance of recent observational studies and RCTs, bariatric surgery is increasingly used with the primary intent to treat T2DM or metabolic disease, hence the term metabolic surgery.242 Observational, nonrandomized studies first demonstrated profound improvements in hyperglycemia and other cardiovascular risk factors following metabolic surgery that were followed by RCTs. Resolution or remission of T2DM is typically defined as becoming “nondiabetic” with normal HbA1c, without medications. One meta-analysis (2009) of 19 mostly observational studies (n = 4,070 patients) reported an overall T2DM remission rate of 78% after bariatric surgery with 1 to 3 years follow-up.244 The patients all had BMI >35 and generally early/mild T2DM that likely increased remission rates. In the Swedish Obese Subjects study, the remission rate following surgery was 72% at 2 years and 36% at 10 years compared with 21% and 13%, respectively, for the nonsurgical controls (P < .001).241 Metabolic surgery was also significantly more effective than nonsurgical treatment in preventing new onset cases of T2DM, with a relative risk reduction of 78%.

A more recent systematic review (2012) evaluated long-term cardiovascular risk reduction after bariatric surgery in 73 studies and 19,543 patients.226 At a mean follow-up of 57.8 months, the average excess weight loss for all procedures was 54%, and rates of remission or improvement were 63% for hypertension, 73% for T2DM, and 65% for hyperlipidemia. Results from 12 cohort-matched, nonrandomized studies comparing bariatric surgery vs. nonsurgical controls demonstrated reduced cardiovascular events and death (30–88% reduction) in patients with and without T2DM.227 One of these studies involving male veterans, who were mostly at high cardiovascular risk, reported a 42% reduction in mortality at 10 years compared with medical therapy.207 Similarly, in the Swedish Obese Subjects study, the mortality rate from cardiovascular disease in the bariatric surgical group was lower than for control patients (adjusted hazard ratio, 0.47; P = .002).241 For patients with T2DM in this study, surgery was associated with a 50% reduction in microvascular complications (41.8 per 1000 person-years for control patients and 20.6 per 1000 person-years in the surgery group; hazard ratio, 0.44; P < .001).235 These observational, nonrandomized studies provide evidence that metabolic surgery is superior to medical management alone in improving glycemic control, reducing cardiovascular risk factors, and lowering long-term morbidity and mortality of T2DM, yet supporting RCTs have been lacking until recently.

During the past 10 years, 11 such RCTs have been published including 794 patients in total (Table 27-7). In these RCTs included obese patients with T2DM (n = 794; range 38–150 patients per study) with follow-up from 6 months to 5 years (Fig. 27-27). All common metabolic surgical procedures were represented including LRYGB (9 studies), LAGB (5 studies), SG (2 studies), and BPD (1 study). T2DM severity varied significantly from mild (mean HbA1c 7.7%, <2-year onset, no insulin) to severe (mean HbA1c 9.3%, duration 8.3 years, 48% on insulin). The BMI ranged from 25 to 53 kg/m², with 11 of 12 studies including patients with BMI <35 kg/m², also called class 1 obesity. Age, sex, and ethnic background were similar,
Table 27-7
Metabolic surgery randomized controlled trials for type 2 diabetes (n = 794)*

<table>
<thead>
<tr>
<th>STUDY</th>
<th>BMI (kg/m²), % OF PATIENTS</th>
<th>DESIGN</th>
<th>NO. OF PATIENTS RANDOMIZED</th>
<th>FOLLOW-UP (MONTHS)</th>
<th>REMISSION CRITERIA*</th>
<th>OUTCOME (REMISSION OR CHANGE IN HbA₁c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dixon¹⁷</td>
<td>&lt;35, 22%</td>
<td>LAGB vs. control</td>
<td>60</td>
<td>24</td>
<td>HbA₁c &lt;6.2%</td>
<td>73% vs. 13%, P &lt;0.001</td>
</tr>
<tr>
<td>Schauer²⁸,²⁴⁶,²⁵²</td>
<td>&lt;35, 36%</td>
<td>RYGB vs. SG vs. control</td>
<td>150</td>
<td>60</td>
<td>HbA₁c &lt;6.0%</td>
<td>22% vs. 15% vs. 0, P &lt;0.05</td>
</tr>
<tr>
<td>Mingrone⁷⁹,⁸²</td>
<td>&gt;35, 100%</td>
<td>RYGB vs. BPD vs. control</td>
<td>60</td>
<td>60</td>
<td>HbA₁c &lt;6.5%</td>
<td>42% vs. 68% vs. 0, P = 0.003</td>
</tr>
<tr>
<td>Ikramuddin ⁸⁰,²⁴⁷</td>
<td>&lt;35, 59%</td>
<td>RYGB vs. control</td>
<td>120</td>
<td>24</td>
<td>HbA₁c &lt;6%</td>
<td>44% vs. 9%, P &lt;0.001</td>
</tr>
<tr>
<td>Liang²⁴⁸</td>
<td>&lt;35, 100%</td>
<td>RYGB vs. control</td>
<td>101</td>
<td>12</td>
<td>HbA₁c &lt;6.5%**</td>
<td>90% vs. 0 vs. 0, P &lt;0.0001</td>
</tr>
<tr>
<td>Halperin⁸³</td>
<td>&lt;35, 34%</td>
<td>RYGB vs. control</td>
<td>38</td>
<td>12</td>
<td>HbA₁c &lt;6.5%</td>
<td>58% vs. 16%, P = 0.03</td>
</tr>
<tr>
<td>Courcoulas⁸⁴,⁸⁶</td>
<td>&lt;35, 43%</td>
<td>RYGB vs. LAGB vs. control</td>
<td>69</td>
<td>36</td>
<td>HbA₁c &lt;6.5%</td>
<td>40% vs. 29% vs. 0, P = 0.004</td>
</tr>
<tr>
<td>Wentworth²⁵⁰</td>
<td>≤30, 100%</td>
<td>LAGB vs. control</td>
<td>51</td>
<td>24</td>
<td>Fasting blood glucose &lt;7.0 mmol/L</td>
<td>52% vs. 8%, P = 0.001</td>
</tr>
<tr>
<td>Parikh²⁵¹</td>
<td>&lt;35, 100%</td>
<td>Bariatric surgery (RYGB, LAGB, SG) vs. control</td>
<td>57</td>
<td>6</td>
<td>HbA₁c &lt;6.5%</td>
<td>65% vs. 0, P = 0.0001</td>
</tr>
<tr>
<td>Ding⁸⁵</td>
<td>&lt;35, 34%</td>
<td>LAGB vs. control</td>
<td>45</td>
<td>12</td>
<td>HbA₁c &lt;6.5%***</td>
<td>33% vs. 23%, P = 0.46</td>
</tr>
<tr>
<td>Cummings⁸¹</td>
<td>&lt;35, 25%</td>
<td>RYGB vs. control</td>
<td>43</td>
<td>12</td>
<td>HbA₁c &lt;6.0%</td>
<td>60% vs. 5.9%, P = 0.002</td>
</tr>
</tbody>
</table>

*Remission defined as reaching HbA1c value without medication, unless otherwise specified
**Remission not precisely defined, extrapolated
***On or off medications


although three studies ⁸⁰,²⁴⁷,²⁴⁸ included a significant number of Asian patients. For most studies, the primary endpoint was remission, defined as an HbA1c target at or below 6.0% to 6.5% without use of diabetes medications.

Overall, these RCTs showed that surgery was significantly more effective than medical treatment in reaching remission and glycemic control (P <0.05) (Fig. 27-27). The one exception showing no superiority of surgery involved gastric banding and resulted in a diabetes remission for LAGB vs. medical treatment of 33% and 23%, respectively (P >0.05).⁸⁵ Overall, surgery decreased HbA1c by 2% to 3.5%, whereas medical treatment lowered it by 1% to 1.5%, as seen in Fig. 27-28. Most of these studies also showed superiority of surgery over medical treatment in achieving secondary endpoints such as weight loss, remission of metabolic syndrome, reduction in diabetes and cardiovascular medications, and improvement in triglycerides, lipids, and quality of life. Results were mixed in terms of improvements in systolic and diastolic blood pressure or low-density lipoproteins after surgery vs. medical treatment, but many studies did show a corresponding reduction in medication usage.

Although previous guidelines and payer coverage policies had limited metabolic surgery to severely obese patients (BMI ≥35 kg/m²), nearly all RCTs showed that the surgical procedures, especially LRYGB and SG, were equally effective in patients with BMI 30 to 35 kg/m². This is particularly important given that many patients with T2DM have a BMI <35 kg/m². The effect of surgery in these patients with a lower class of obesity is also durable out to at least 5 years.²⁴⁵,²⁵²

None of these RCTs were sufficiently powered to detect differences in macrovascular or microvascular complications or death, especially at the relatively short follow-up, and no such differences have been detected thus far. Four of the RCTs from Pittsburgh, Seattle, Boston, and Cleveland have combined their patient populations in a pooled study to assess 10-year outcomes. This study, Alliance of Randomized Trials of Medicine vs. Metabolic Surgery (ARMMS), aims to identify long-term risks and benefits of metabolic surgery.

The evidence, as previously summarized, was the basis for newly established international guidelines on the role of metabolic surgery in treating T2DM. In 2015, the 2nd Diabetes Surgery Summit (DSS-II) Consensus Conference generated
### Fixed-Effects Model

<table>
<thead>
<tr>
<th>Study Details</th>
<th>Glyc. Endp.</th>
<th>N</th>
<th>Weight</th>
<th>Peto, Fixed, 95% CI</th>
<th>Polo Odds Ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parikh 2014 (RYGB/LAGB/SG) [6 mo; ≤6.5% off meds] (18)</td>
<td>13</td>
<td>20</td>
<td>4.5%</td>
<td>21.15 [5.85, 76.51]</td>
<td></td>
</tr>
<tr>
<td>Courcoulas 2014 (RYGB/LAGB) [12 mo; ≤6.5% off meds] (14)</td>
<td>18</td>
<td>41</td>
<td>5.1%</td>
<td>7.51 [2.24, 25.21]</td>
<td></td>
</tr>
<tr>
<td>Ding 2015 (LAGB) [12 mo; ≤6.5% off meds] (22)</td>
<td>6</td>
<td>18</td>
<td>3.9%</td>
<td>1.68 [0.42, 6.66]</td>
<td></td>
</tr>
<tr>
<td>Halperin 2014 (RYGB) [12 mo; ≤6.0% off meds] (24)</td>
<td>11</td>
<td>19</td>
<td>4.4%</td>
<td>5.82 [1.59, 21.39]</td>
<td></td>
</tr>
<tr>
<td>Schauer 2014 (RYGB/SG) [36 mo; ≤6.0% off meds] (19)</td>
<td>28</td>
<td>57</td>
<td>12.5%</td>
<td>3.72 [1.72, 8.04]</td>
<td></td>
</tr>
<tr>
<td>Mingrone 2015 (RYGB/BPD) [60 mo; ≤6.5% off meds] (20)</td>
<td>34</td>
<td>99</td>
<td>10.4%</td>
<td>6.39 [2.74, 14.88]</td>
<td></td>
</tr>
</tbody>
</table>

### Heterogeneity

- Chi^2 = 45.43, df = 14 (P < 0.0001); I^2 = 69%
- Overall effect: Z = 15.36 (P < 0.00001)

### Figure 27-27

Glycemic endpoints of RCTs by length of follow-up. Forest plot of Peto odds ratios (ORs) of main glycemic end points, as defined in each trial, from published RCTs of bariatric/metabolic surgery compared with medical/lifestyle treatments for diabetes with data arranged in order of increasing length of follow-up. (Reproduced with permission from Cummings DE, Cohen RV: Bariatric/Metabolic Surgery to Treat Type 2 Diabetes in Patients With a BMI <35 kg/m^2, Diabetes Care. 2016 Jun;39(6):924-933.)
THE SURGICAL MANAGEMENT OF OBESITY

CHAPTER 27

Patients With Type 2 Diabetes

Obese
BMI ≥30 kg/m² or ≥27.5 for Asians

Class III Obese
BMI ≥40 kg/m² or ≥37.5 for Asians

Expeditied Assessment for Metabolic Surgery

Class II Obese
BMI 35.0–39.9 kg/m² or 32.5–37.4 for Asians

Optimal Lifestyle and Medical Rx

Class I Obese
BMI 30.0–34.9 kg/m² or 27.5–32.4 for Asians

Optimal Lifestyle and Medical Rx (including injectable meds and insulin)

Class II Obese With Poor Glycemic Control

Recommend Metabolic Surgery

Class II Obese With Adequate Glycemic Control

Consider Metabolic Surgery

Class I Obese With Poor Glycemic Control

Class I Obese With Adequate Glycemic Control

Nonsurgical Treatment

Nonobese
BMI <30 kg/m² or <27.5 for Asians


DSS-II guidelines, metabolic surgery should be recommended to treat T2DM in patients with class III obesity (BMI ≥40 kg/m²) regardless of glycemic control and in those with class II obesity (BMI 35.0–39.9 kg/m²) when hyperglycemia is inadequately controlled by lifestyle and optimal medical therapy (Fig. 27-29). Surgery should also be considered for patients with T2DM and BMI 30.0 to 34.9 kg/m² if hyperglycemia is inadequately controlled despite optimal treatment with either oral or injectable medications. These BMI thresholds should be reduced by 2.5 kg/m² for Asian patients. The new treatment algorithm from DSS-II incorporates appropriate use of all three treatment modalities: lifestyle intervention, drug therapy, and surgery (see Fig. 27-29). The 2017 Standards of Care for Diabetes from the American Diabetes Association include those key indications in its recommendations for metabolic surgery, as well.

COMPLICATIONS OF BARIATRIC SURGERY

Surgical Complications

None of the surgical procedures are without risks. The perioperative mortality for the average patient is low (<0.5%) and declining, but can vary significantly across subgroups with perioperative mortality rates of 2.0% or higher in some patient populations. The incidence of complications after the various surgical procedures varies from 4% to over 25% and depends on the definition of complication used, the type of bariatric procedure performed, and patient characteristics. The 2017 Standards of Care for Diabetes from the American Diabetes Association include those key indications in its recommendations for metabolic surgery, as well.
In the 11 RCTs (794 patients) that have compared bariatric surgery to nonsurgical treatment, rates of adverse events were higher among surgical subjects, with follow-up up to 5 years for two of the studies and up to 2 to 3 years for others.\textsuperscript{82,252,256} There were very few cardiovascular events or deaths in either the surgical or the nonsurgical groups, and the most common adverse events after surgery were iron deficiency anemia (15% with intestinal bypass operations) and reoperations (8%). These RCTs were not large enough to compare safety between procedure types, and most of the comparative data on procedure-to-procedure complications has and will come from larger observational studies.

The 30-day mortality in the LABS Study was 0.3% for all procedures with a major adverse outcome rate (a predefined composite endpoint that included; death, venous thromboembolism, reintervention [percutaneous, endoscopic, or operative], or failure to be discharged from the hospital in 30 days) of 4.1%.\textsuperscript{59} Major predictors of an increased risk of complications in LABS were a history of venous thromboembolism, a diagnosis of obstructive sleep apnea, impaired functional status defined as inability to walk 300 feet (91 m), extreme BMI, and undergoing an RYGB by the open technique. Other large observational studies, such as SOS, have shown higher rates of complications, with 14.5% having at least one nonfatal complication over the first 90 days, including pulmonary complications, vomiting, wound infection, hemorrhage, and anastomotic leak. However, the SOS included mostly open and VBG procedures, which are rarely performed today. Despite these older procedures and techniques, the 90-day mortality rate in SOS was low at 0.25%.\textsuperscript{196}

In a 2007 meta-analysis of 361 studies, mostly nonrandomized observational studies involving over 85,000 patients, Buchwald and colleagues reported important differences in ≤30-day mortality across different laparoscopic bariatric procedures: 0.06% for LAGB; 0.21% for VBG; 0.16% for RYGB; and 1.11% for BPD/DS.\textsuperscript{257} This review also found significantly higher mortality for open procedures compared with those performed laparoscopically. A U.S. study of over 300,000 patients in 12 states examined in-hospital complications before and after implementation of the CMS national policy restricting insurance coverage for bariatric surgery to “centers of excellence.” The study found no significant differences in complications before and after the policy was implemented, and overall complication rates were 7% to 8% with 3.3% to 3.6% being serious, including a 1% reoperation rate during both time periods.\textsuperscript{1,258} A clinically useful prognostic risk score was also developed and validated in 9382 patients to predict 90-day mortality after LRYGB surgery using five clinical characteristics: BMI ≥50 kg/m², male sex, hypertension, known risk factor for pulmonary embolism, and age ≥45 years. Patients with 4 to 5 of these characteristics are at much higher risk of death (4.3%) by 90 days than those with 0 to 1 characteristics (0.26%).\textsuperscript{259,260} A systematic review of 15 RCTs of SG found no deaths in just under 800 patients but a 9.2% mean complication rate (range 0–18%).\textsuperscript{234} In the American College of Surgeons Bariatric Surgery Network database, mortality 30 days after SG was 0.11%, positioning its overall complication profile between that for LAGB (0.05%) and RYGB (0.14%). The 30-day complication rate was similarly positioned at 5.6% for SG, 1.4% LAGB, and 5.9% for RYGB.\textsuperscript{140}

An established problem now is the frequent rate of reoperation, particularly among LAGB patients. In O’Brien and colleagues’ prospective cohort of 3227 LAGB patients, revisional procedures occurred in 1116 (35%) and were performed for the following reasons: proximal enlargement (26%), port and tubing problems (21%), and erosion (3.4%). The need for revision due to proximal enlargement decreased dramatically over a 17-year period from 40% to 6.4% as the surgical technique evolved; however, the band was ultimately removed in 5.6% of all individuals.\textsuperscript{262} Other long-term cohorts suggest that LAGB removal rates may be as high as 50%. The O’Brien systematic review of long-term studies indicates that the rate of revisional surgery for LRYGB may be similar to LAGB (22% revision for LRYGB, range 8% to 38%; 26% revision for LAGB, range of 8% to 60%). However, in the LABS Study, there was a higher rate of revision and reoperation for LAGB as compared to RYGB at both 3 and 7 years of follow-up.\textsuperscript{169} In general, more long-term data with more complete follow-up with standardized definitions and reporting of complications are needed to compare reoperation and complication rates of all bariatric procedures.

**Nonsurgical Complications**

Postgastric bypass hypoglycemia (PGBH) in a relatively uncommon but particularly challenging problem that affects an unknown number of patients in the longer term. Prevalence rates in the literature vary from 1% to 11%, depending on the definition.\textsuperscript{263,264} Hypoglycemia is characterized by documentation of Whipple’s triad (including both autonomic and neuroglycopenic symptoms or signs), at the time of a plasma glucose concentration <55 mg/dL with resolution of symptoms and signs after glucose administration. During these episodes, plasma insulin levels are inappropriately high, indicating dysregulation of β-cell function. When it was initially described in patients who had undergone partial pancreatectomy, nesidioblastosis was found in the specimens that are characterized by hyperplasia and/or dysplasia of the pancreatic islets.\textsuperscript{265} It was initially thought to be endogenous hyperinsulinemia from increased β-cell mass hyperfunctioning islet cells; however, the current thinking is that the recalcitrant symptoms of hyperinsulinemic hypoglycemia after RYGB are related to the anatomic and physiologic changes and not from an inherent change in β-cell mass. One candidate mediator of increased insulin secretion in PGBH is GLP-1, a peptide released from intestinal neuroendocrine L-cells in response to meals. Consistent with this hypothesis, postprandial GLP-1 levels are increased by over tenfold in post-RYG patients, are higher in those with hyperinsulinemic hypoglycemia and neuroglycopenia, and correlate inversely with postprandial glucose levels.\textsuperscript{266} PGBH needs to be distinguished from other forms of hypoglycemia such as other functional β-cell disorders such as noninsulinoma pancreatogenous hypoglycemia, insulinoma, reactive hypoglycemia, or early or late dumping syndrome. It is possible that PGBH is a spectrum of hypoglycemia with late dumping being on the end of the spectrum that is more responsive to dietary changes alone while more severe PGBH can be associated with severe symptoms.\textsuperscript{263} First-line therapeutic approaches to PGBH include medical nutrition therapy aimed at reducing intake of high glycemic index carbohydrates and premeal treatment with acarbose. Additional therapies that may be considered include octreotide, diazoxide, calcium channel blockers, GLP-1 receptor antagonists, and providing nutrition solely through a gastrostomy tube placed into the bypassed duodenum. Reversal of gastric bypass is not uniformly successful, suggesting the importance of underlying genetics and/or compensatory mechanisms that may persist after surgical reversal.\textsuperscript{266} Finally, although pancreatic resection was initially employed for patients with life-threatening hypoglycemia, this procedure
is not uniformly successful in remitting hypoglycemia and should not be considered for the majority of patients, who can experience improvement in their symptoms with a combination of medical approaches. A 2017 American Society of Metabolic and Bariatric Surgery (ASMB) position statement provides a comprehensive summary of this topic and also recommends this multimodal medical approach.267

There is data to suggest that babies born to women following bariatric surgery are at risk for certain complications. In a Swedish study, bariatric surgery was originally associated with reduced risks of gestational diabetes and excessive fetal growth, shorter gestation, and an increased risk of small-for-gestational-age infants. In a later follow-up report, this same group reported a significant association between a history of bariatric surgery and an increased risk of preterm birth and spontaneous preterm birth, in particular.199,268

There is also emerging data from observational studies that some bariatric procedures may be associated with a greater long-term risk of substance and alcohol use disorders, suicide, and nutritional deficiencies. Pharmacokinetic studies indicate that after LRYGB and SG, the anatomic changes lead to very rapid absorption of alcohol and marked increases in blood alcohol concentrations for a single small.269,270 In the SOS study, RYGB was associated with increased alcohol consumption and an increase in alcohol abuse events (HR: 4.9) over 20 years.196 Similarly in the LABS study, alcohol use disorders were found to be more common in the second postoperative year (9.6%) in those undergoing RYGB compared to before surgery (7.6%). Risk factors for alcohol use disorders included male gender, younger age, and preoperative smoking or alcohol use.271 At 7-year follow-up in LABS, there was a progressive and significant increase over time in the prevalence of regular alcohol consumption for both RYGB and LAGB. In addition, alcohol use disorders, illicit drug use, and treatment for substance use disorders, increased over the 7-year period for RYGB only.272 Overall, these rates were high, with 20% of RYGB participants reporting incident alcohol use disorder symptoms within 5 years of surgery.

In addition, there may be an increased risk for suicide273-275 following bariatric surgery, although the etiology is unclear and the data is varied and complex to interpret.276 The Utah mortality study showed a 58% increase in all non-disease-related causes of death in the RYGB group compared to the matched control population, including a small but significant increase in suicides, accidents, and intentional poisonings.202 Similar findings were observed in the second Utah Obesity Study.203 An observational study using Pennsylvania state data found that suicide rates were 13.7 per 10,000 among men and 5.2 per 10,000 among women among postbariatric surgery patients in Pennsylvania over 10 years, which were both significantly higher than age and sex-matched rates in the United States. In addition, the majority (70%) of these deaths occurred in the first 3 years following surgery when clinical follow-up is incomplete.277

Finally, there is evidence that vitamin and micronutrient deficiencies are common following bariatric surgery including calcium, vitamin D, iron, zinc, and copper, and others. Guidelines suggest that all patients should be screened for deficiencies preoperatively as some deficiencies predate the surgical procedure (see “Follow-Up Postoperative Care”). After surgery, patients must be provided daily nutritional supplementation and undergo routine long-term monitoring for deficiencies (see Table 27-4). Data continue to suggest that the prevalence of micronutrient deficiencies is increasing, while monitoring is decreasing. Aside from these recommendations, there is insufficient evidence currently regarding optimal dietary and nutritional management following bariatric procedures, including how to treat some of the specific complications of bariatric operations such as chronic nausea and vomiting, hypoglycemic episodes, failed weight loss, and anastomotic ulcers and strictures.1,278-280

REOPERATIVE (REVISION AND CONVERSION) BARIATRIC SURGERY

Introduction

Surgical treatments for chronic diseases such as obesity often require additional or revisional surgical procedures when the primary procedure did not sufficiently treat the underlying disease. This is true in joint replacements when treating osteoarthritis, coronary-artery bypass graft surgeries when treating coronary artery disease, and bariatric surgery in the treatment of obesity.281 Additionally, we suspect that obesity is a heterogeneous disorder282 and is therefore being treated with a variety of procedures with different mechanisms of action.63 Given this, it is not surprising that some patients are “treatment failures” with respect to improvements in weight, comorbidities, and quality of life. Also, older bariatric surgical procedures such as the jejunoileal bypass, the VBG, and early gastric bypass procedures (which utilized a horizontal and/or partitioned stomach) require revision because of a higher complication rate.283,284 Initial bariatric surgery cases have averaged greater than 150,000 cases per year for the last 15 years.285 For these reasons, reoperative bariatric surgery has become increasingly prevalent over the last decade. Despite its increasing prevalence, there are challenges in assessing the frequency and effectiveness of these procedures. Reoperative bariatric surgery has been difficult to categorize meaningfully and to quantify due to the multiple procedure codes, many with little specificity. Reports in the literature range from 5% to 50% depending on the primary procedure.286

Multiple retrospective, as well as case-matched and case-controlled studies of revisional bariatric surgery, demonstrate they are effective with benefits to weight loss and overall health,281,282 although this is not without some controversy. The literature supports reoperative bariatric surgery in two situations: treatments of insufficient weight loss or weight regain and the treatment of acute and chronic complications.281,283 There are many revisional procedures and approaches that are effective. There is no data-driven evidence to guide in the selection of which patient will benefit most from revisional bariatric surgery. Additionally, there is no evidence-based consensus as to which revisional surgical approach is most optimal in any given situation.283 Currently, there is also little evidence as to which bariatric surgical procedure will be efficacious for any specific given patient.286,287

Principles and Preoperative Evaluation

As with any other decisions for surgery, revisional or additional bariatric surgery requires evaluating the risks and the benefits of the procedure for specific patient situations. It has been observed that the weight loss following revisional procedures is less than with a primary procedure.288 Reoperations in general are associated with morbidity and mortality that is higher than with primary bariatric procedures281,288,289 but are acceptably low if careful selection of patients is coupled with
adequate surgeon experience.\textsuperscript{281,290} Reoperative bariatric surgery should be undertaken by experienced bariatric surgeons in centers with the wide range of medical resources to manage these complex patients.\textsuperscript{281} In some situations, it is reasonable to consider a two-stage or an open vs. laparoscopic approach.\textsuperscript{283} When evaluating a patient for revisional bariatric surgery, it is most important to establish clearly the reason for revision.\textsuperscript{283,291} Is the issue a surgical complication, insufficient weight loss, or weight regain? Is there a discernable anatomic cause for the patient’s symptoms (e.g., abdominal pain, nausea, vomiting, heartburn or reflux, nutritional deficiency?)? In situations of noninittal weight loss or weight regain, what is the best justification that can be made? Are there behavioral, or other nonsurgical modifications that might significantly help to attain further weight loss? Will the patient be able to set into place the behavioral modification required of the revisional surgical approach? With consideration for revision, it is reasonable to place less emphasis on absolute weight loss and focus on the comorbidity and quality of life states. A comprehensive preoperative evaluation is required to fully answer these and other questions.

The preoperative evaluation for bariatric surgical revision should include:

- Review of the initial surgical operative note to understand the exact initial procedure
- Anatomic evaluation of the GI track utilizing upper gastrointestinal endoscopy and radiology to identify known bariatric surgical complications
- An extensive nutritional evaluation
- A behavioral health evaluation with a focus on the adaption to the initial bariatric procedure and potential adaptation to a revisional procedure
- A full medical evaluation to determine the patient’s suitability to undergo anesthesia and a surgical procedure\textsuperscript{288,291}

The results from these evaluations are used to formulate a hypothesis to explain the patient’s symptoms and outcome from the initial procedure and to assess their suitability for and the potential benefits from the potential revisional surgical options. It will come down to a risk-benefit discussion between the surgeon, other members of the multidisciplinary support team, and the patient.\textsuperscript{281}

Treatment for Insufficient Weight Loss or Weight Regain

**Vertical Banded Gastroplasty.** Reversal of VBG is associated with significant weight gain. Revision of the VBG to a re-VBG was associated with poor outcomes.\textsuperscript{292} VBGs have been safely converted with open and laparoscopic approaches to RYGB and SG.\textsuperscript{293-295} Conversion of VBG to BPD/DS procedures has limited data with higher leak rate and mortality.\textsuperscript{296}

**Adjustable Gastric Band.** There are increasing reports of LAGB failure and disappointing weight loss over the last decade.\textsuperscript{298,297,298} Removal of the band without an additional bariatric procedure has been associated with significant weight regain.\textsuperscript{298} Repositioning or replacement of the band is technically possible; however, the long-term weight loss outcomes are mixed.\textsuperscript{299-301} When converting an LAGB to another bariatric procedure, there is some evidence to suggest possibly fewer complications with a two-step approach when significant adhesions or a thick gastric capsule are present. This two-step approach would include removal of the band, allowing 3 to 6 months for gastric tissue healing, and then completing the conversion.\textsuperscript{302,303} However, one-stage conversions have been reported with acceptable outcomes.\textsuperscript{297} Good outcomes have been demonstrated with conversion of LAGBs to SG, RYGB, and BPD/DS.\textsuperscript{285,297} Several investigators advocate that a failed restrictive bariatric surgery due to poor weight loss should include a revision option with a malabsorptive component.\textsuperscript{285}

**Sleeve Gastrectomy.** Approximately 5% to 10% of primary SG procedures have been reported to require revision for poor weight loss outcomes.\textsuperscript{304,305} The literature supports conversion to RYGB and BPD/DS.\textsuperscript{306} There is controversy regarding resleeve gastrectomy.\textsuperscript{307,308}

**Roux-en-Y Gastric Bypass.** Approximately, 10% to 20% of patients after a primary RYGB will have inadequate weight loss or weight regain at 2 years, and a subset of these will require a revisional surgical procedure.\textsuperscript{309} Options for revision include banding over the Roux-en-Y bypass, gastric pouch and gastrojejunal revisions, RYGB limb lengthening, and conversion to a duodenal switch.\textsuperscript{281,285} Each of these options has strengths and weaknesses in specific situations. Endoscopic revisions to reduce the gastric pouch and/or gastrojejunal stomal size have been shown to arrest weight gain with short-term weight loss, but the studies have been small and are noncontrolled.

Treatment of Surgical Complications

**Vertical Banded Gastroplasty.** Several complications have been indications for revision/conversion of VBGs: wide outlet, pouch dilation, staple line erosion, stoma stenosis, band erosion, band dehiscence, and GERD. Most often conversion is undertaken, and VBGs have been safely converted to RYGB.\textsuperscript{281,285}

**Adjustable Gastric Band.** The following LAGB complications may require additional or revisional surgery: early band obstruction, severe or chronic gastric prolapse or symmetrical gastroesophageal dilatation, band erosion, port and tubing problems, severe or persistent esophageal dysmotility, or psychological intolerance to restriction of band. The literature supports repair of tubing and port problems and revision of LAGB to SG, RYGB, BPD/DS in one- or two-stage revisional operations.\textsuperscript{285,297}

**Sleeve Gastrectomy.** Complications that may require a revisional procedure after sleeve gastrectomy are staple line leaks, sleeve stricture and sleeve dilatation, and gastroesophageal reflux. Obstruction due to stricture usually at the angularis incisura will require first-line treatment of endoscopic dilation and may require a revision to a RYGB. Staple line leaks are initially controlled with endoscopic stenting or drainage. Acute and chronic leaks may develop into fistulous disease and require conversion to RYGB. Persistent gastroesophageal reflux may also require conversion to RYGB.\textsuperscript{281,285}

**Roux-en-Y Gastric Bypass.** Several complications after primary RYGB have been demonstrated as indications for revisional surgery: gastric pouch dilation, gastrojejunostomy dilation/stricture, marginal ulcers, bowel loss due to internal hernia or volvulus, roux stasis syndrome, gastrogastric fistulas, anastomotic structures or ulcers, and metabolic/endocrine derangements. These require a revision focused on the mechanism of the complication. Reversal of an RYGB is reserved for severe instances of intractable nausea/vomiting, extreme weight loss and malnutrition, metabolic abnormalities, nonhealing ulceration or leaks, and patient choice. It has been performed rarely, and case reports indicate successful resolution of endocrine, metabolic, and nutritional abnormalities with improved
metabolic parameters. However, 50% to 88% of patients have been reported to regain significant weight.281,285

Biliopancreatic Diversion With Duodenal Switch. Acute complications are similar to RYGB and treated the same. The most severe chronic complication is protein-calorie malnutrition, and incidence ranges from 1% to 6%. Management is meticulous nutritional evaluation and nutritional and pancreatic enzyme support, with surgery as a fall back if weight and protein stores are not stabilized. Surgery would entail lengthening the common channel and is rarely necessary.281

SPECIAL ISSUES IN BARIATRIC SURGERY

Bariatric Procedures in Adolescents

The major controversy with regard to adolescents undergoing bariatric surgery includes the general aversion to subjecting an adolescent to surgery as well as the concern for the secondary side effects of bariatric surgery on remaining growth and development. Clearly the younger the patient, the more relevant the latter concern becomes. Bariatric surgery in adolescents has been performed more frequently, with only 800 cases a year in 2003 increasing to 1600 cases in 2009, but overall rates are much lower than in adults.310,311 Much of the clinical outcomes data have been extrapolated from the adult literature, and more evidence is needed to demonstrate whether weight loss is durable over time, the impact on obesity-related conditions such as T2DM, hypertension, and others, and how often patients experience short- and longer-term complications.

A meta-analysis involving 131 adolescents undergoing bariatric surgery demonstrated a 17.8 to 22.3 kg/m² decrease in BMI after RYGB.312 They also observed improvement of hypertension in more than half the patients and sleep apnea resolution in all 131 patients. There were four mortalities in this cohort, but only one of them was potentially associated with the procedure (Clostridium difficile colitis 9 months after operation). Morbidity in the adolescent literature ranges from 0% to 38%. The most common complication in the meta-analysis was nutrient deficiencies. The ASMSB pediatric guidelines suggest using BMI criteria similar to the adult population but with some modifications to comorbidity thresholds.313 They recommend considering surgery in patients with a BMI of 35 kg/m² or greater with major comorbidities (e.g., T2DM, severe nonalcoholic fatty liver disease, OSA) or a BMI of ≥40 kg/m² or greater with minor comorbidities (e.g., hypertension, dyslipidemia, insulin resistance). One more recent multicenter, prospective study of bariatric surgery in adolescents, the Teen-Longitudinal Assessment of Bariatric Surgery (Teen-LABS) study, is following 242 adolescent patients for at least 10 years, who underwent bariatric surgery at 5 academic centers.314 Fifty-one percent of adolescents had four or more major comorbid conditions before surgery. LRYGB, SG, and LAGB were performed in 66%, 28%, and 6% of patients, respectively. There were no deaths during the initial hospitalization or within 30 days of operation; major complications such as reoperation were reported in 19 patients (8%). Minor complications such as dehydration were reported in 15%. All reoperations and 85% of readmissions in the 30-day period were related to the bariatric surgery.315 Three-year follow-up data on the Teen-LABS cohort reported a mean weight loss of 28% for LRYGB and 26% for SG. Remission of T2DM occurred in 95% of participants who had had the condition before surgery, remission of abnormal kidney function occurred in 86%, remission of prediabetes in 76%, remission of elevated blood pressure in 74%, and remission of dyslipidemia in 66%. Rates of improvements in comorbid conditions, including T2DM, occurred at higher rates than in adults. Hypoglycemia was found in 57% of the participants, and 13% of the participants had undergone one or more additional intraabdominal procedures at 3 years.316 So despite the clinically significant improvements in weight, diabetes, cardiometabolic health, and weight-related quality of life that were observed at 3 years following surgery, the reoperation and micronutrient risks warrant longer observation and further study.

Cost Effectiveness

In a Canadian study, including five systematic reviews, two economic evaluations, two reviews of guidelines, and six primary evidence-based guidelines, the cost effectiveness for the use of bariatric surgery in adolescents was reviewed.316 The limited available evidence suggested superior weight loss, resolution of comorbidities compared to nonsurgical interventions, and potential superior weight loss with RYGB compared to other procedures. Cost-effectiveness data was lacking, but limited evidence suggested that bariatric surgery was cost effective several years after intervention, but not immediately.316 A U.S. cost-effectiveness analysis of bariatric surgery in adolescents has been published. In addition to the cost of the surgery, perioperative mortality, complications, and quality of life improvement were included in the modelled analysis. By the fifth year of follow-up, bariatric surgery was found to be cost effective in adolescent patients when compared to a cohort of patients with obesity who had not undergone surgery.317

For adults, the overall impact of bariatric surgery to reduce expenditures sufficiently to achieve cost savings continues to be debated. In a Canadian matched cohort study prior to the laparoscopic era, it was shown that bariatric surgery decreases long-term direct healthcare costs and the initial costs of surgery can be amortized over 3.5 years.318 In two observational studies, bariatric surgery was shown to be cost saving over a relatively short period of time.319,320 In more recent observational studies, including the large SOS study and another an analysis of 30,000 single payor enrollees in the United States, show no evidence of overall cost savings.206,321,322

In general, review of the evidence to date suggests that outpatient costs, including pharmacy costs, are significantly reduced after bariatric surgery. However, long-term inpatient hospital costs are increased or unchanged in those who have undergone bariatric surgery compared with matched nonsurgical patients, so no long-term net cost benefit is achieved. Other modeled cost effectiveness studies are consistent with these results as well.323,324 So it is likely that bariatric procedures are cost effective, but do not produce cost savings, compared with nonsurgical treatments.

Quality Assurance

Between 1998 and 2003, with emergence of the laparoscopic technique for bariatric surgery, there was rapid increase in the number bariatric surgical procedures performed. With bariatric surgeons becoming accustomed to laparoscopic techniques and laparoscopic surgeons learning bariatric procedures and patient care, there were realistic concerns regarding the safety of bariatric surgery.325 In 2004, ASMSB utilized the volume-outcome concept of centers of excellence and developed the first bariatric surgical accreditation program, ASMSB-Center of Excellence (COE). In 2005, the American College of Surgeons (ACS) initiated the ACS Bariatric Surgery Center Network (BSCN).
These accreditation programs verified that bariatric surgery centers had the infrastructure and equipment to care for the morbidly obese, experienced and qualified surgeons and staff, appropriate pre- and postoperative processes in place, and reported outcome data on all surgical cases.\textsuperscript{325}

In 2012, these two bariatric surgical accreditations merged into a single unified program, the Metabolic and Bariatric Surgery Accreditation and Quality Improvement Program (MBSAQIP). The MBSAQIP has continued to grant accreditation only after a rigorous review process during which a center proves that it can maintain certain physical resources, human resources, and standards of practice.\textsuperscript{326} Additionally, prospective outcome data is collected at the clinical center and validated based on standardized definitions and submitted to the quality improvement program. Centers receive back risk-stratified analysis of their data as a means to compare with the national statistics/standards and are then required to utilize the data in quality improvement projects at their center. This newer focus on quality improvement over centers of excellence will likely continue to improve the quality of bariatric surgery in the United States. Recently, in a systematic review, bariatric facility accreditation by this program has been associated with improved outcomes (length of stay, mortality, morbidity).\textsuperscript{327}

**Plastic Surgery After Weight Loss**

Patients who have undergone bariatric surgery are often left with skin and subcutaneous tissue deformities. Additional problems include skin rashes and maceration under folds in the pannus, thighs, and breasts; body odor; and poorly fitting clothes. Excess skin can also be a limiting factor in exercise and sexual activity. Plastic and reconstructive surgery is now a part of the continuum of care for bariatric surgery patients. Reconstructive surgery requires careful preoperative planning and is based on the patient’s deformities and priorities. Timing of plastic and reconstructive surgery is typically deferred until weight stability at approximately 1 to 2 years postoperatively to ensure improved healing. Excess tissue of the lower torso is the most common area for which patients undergo surgical intervention and a standard abdominoplasty is typically performed. More radical body contouring can include a circumferential abdominoplasty and lower body lift.\textsuperscript{328,329} This procedure involves excision of tissue from the buttocks and lateral thighs, with skin undermining down the thighs. Circumferential abdominoplasty removes redundant skin of the lower abdomen, flattens the abdomen, and incorporates the lower body lift. It requires central undermining to the xiphoid and minimal lateral undermining of the superior flap. The central abdominal fascia often requires imbrication. If simultaneous abdominal hernia repair is performed, this performs the function of fascial imbrication by creating a repair with some degree of fascial tension. The closure of the superior flap to the inferior skin edge incorporates lateral tension to narrow the waist and advance the anterolateral thighs. Medial thighplasty also may be needed for patients with significant excess medial thigh skin.

Mid-back and epigastric deformity, along with sagging breasts, are corrected with an upper body lift. The upper body lift is a reverse abdominoplasty, removal of mid-torso excessive skin, and reshaping of the breasts. For highly selected individuals, and with a well-organized team, a single-stage total body lift, which includes a circumferential abdominoplasty, lower body lift, medial thighplasty, an upper body lift, and breast reshaping, can be performed safely in under 8 hours (Figs. 27-30 and 27-31).\textsuperscript{330} Increasing numbers of patients are seeking these corrective procedures, and data about the results is evolving. There

![Figure 27-30](https://example.com/image.png)

*Preoperative frontal, right lateral, and left anterior oblique views of a 36-year-old, 150-lb (68-kg) 5'6" woman who lost 120 lb (54 kg) 2 years after laparoscopic Roux-en-Y bypass procedure. She desired a one-stage total body lift and bilateral brachioplasties, which were performed in the manner described in the text. (Used with permission from Dennis Hurwitz, MD, Clinical Professor of Plastic Surgery, University of Pittsburgh.)*
is a hypothesis that if body image is improved with corrective surgery, that weight maintenance in the longer-term may also be positively affected. Several matched controlled studies suggest that plastic surgery after bariatric surgery may improve long-term weight loss results.228,231

**FUTURE IMPORTANT QUESTIONS**

The volume and quality of literature in the field of bariatric surgery has grown tremendously in the last 10 years. High-quality evidence now shows that bariatric surgical procedures result in greater weight loss than nonsurgical treatments, improved survival, and are more effective at inducing remission of T2DM in people with obesity. More information is still needed about the long-term durability of comorbid health improvements and long-term complications after each of the different bariatric surgical procedures. In addition, the underlying specific mechanism(s) of action for both bariatric and metabolic surgery is still incompletely understood. Future knowledge will come from translational human studies, the ongoing longer-term studies and data registries, randomized studies comparing surgical to nonsurgical treatments, integrated health care systems data, and national “big data” networks. The following are some of the high-priority questions that future research will address.

- What are the specific mechanisms of action responsible for weight loss and the T2DM response to bariatric surgical procedures?
- What patient level factors can predict success with weight loss, health improvements, and cost savings after bariatric surgical procedures? Understanding pre- and postsurgery predictors will help to tailor an individual’s treatment.
- Is bariatric surgery more effective than nonsurgical care for the longer-term treatment of T2DM in people with less severe obesity (class I obesity, BMI <35)?
- With more standardized reporting of complications across bariatric studies, what are the long-term complication rates after different bariatric procedures?
- What is the effect of bariatric surgery on long-term microvascular and macrovascular event rates?
- What are the reproductive and mental health outcomes including risk for self-harm and suicide, alcohol use disorders, substance abuse, and other risk-taking behaviors?

**REFERENCES**

Entries highlighted in bright blue are key references.


16. Rubino F, Nathan DM, Eckel RH, et al. Metabolic surgery in the treatment algorithm for type 2 diabetes: a joint statement by international diabetes organizations. Diabet Care. 2016;39(6):861-877. This paper summarizes the most recent Diabetes Surgery Summit (DSS-II) conclusions for which three rounds of Delphi-like questionnaires were used to measure consensus for 32 data-based conclusions. Based on the evidence, bariatric/metabolic surgery was recommended to treat Type 2 diabetes in patients with Class III obesity (BMI>40 kg/m²) and in those with Class II obesity (BMI 35.0-39.9 kg/m²) when hyperglycemia is inadequately controlled by lifestyle and optimal medical therapy. Surgery should also be considered for patients with T2D and BMI 30.0-34.9 kg/m² (Class I obesity) if hyperglycemia is inadequately controlled despite optimal medical treatment.

17. Stefater MA, Wilson-Perez HE, Chambers AP, Sandoval DA, Seeley RJ. All bariatric surgeries are not created equal: insights from mechanistic comparisons. Endocr Rev. 2012;33(4):595-622. This paper is a review of what is known from published animal and human studies about the underlying basic mechanisms of effect of each of 3 most common surgical procedures. This review forms the basis of why research has shifted away from anatomic characterizations of the bariatric procedures as either “restrictive” or “malabsorptive” in nature, towards the physiology of common metabolic changes.


59. Flum DR, Belle SH, King WC, et al. Perioperative safety in the longitudinal assessment of bariatric surgery. *N Engl J Med*. 2009;361(5):445-454. This is a 30-day safety paper from the Longitudinal Assessment of Bariatric Surgery (LABS-1) study. It is a definitive paper on the morbidity and mortality of bariatric surgery from a multi-center and carefully studied cohort of patients undergoing bariatric surgery. The 30-day mortality was 0.3% for all procedures with a major adverse outcome rate [a pre-defined composite endpoint that included: death, venous thromboembolism, re-intervention (percutaneous, endoscopic, or operative), or failure to be discharged from the hospital in 30 days] of 4.1%.


82. Mingrone G, Panunzi S, De Gaetano A, et al. Bariatric-metabolic surgery versus conventional medical treatment in obese patients with type 2 diabetes: 5 year follow-up of an open-label, single-centre, randomised controlled trial. *Lancet.* 2015;386(9997):964-973. This is a randomized clinical trial comparing gastric bypass, biliopancreatic diversion, and medical treatment for the treatment of type 2 diabetes in people with obesity. It is only one of 2 randomized studies addressing this question with 5 year follow up, at this time.


1214
SPECIFIC CONSIDERATIONS

PART II


169. Courcoulas AP, Christian NJ, Belle SH, et al. Weight change and health outcomes at 3 years after bariatric surgery among individuals with severe obesity. JAMA. 2013;310(22):2416-2425. This is a 3-year outcomes paper from the Longitudinal Assessment of Bariatric Surgery (LABS-2) effectiveness study. This is a multi-center, longitudinal study of over 2500 people undergoing both gastric bypass and laparoscopic gastric banding. Results in this report are 3-year weight and health (including type 2 diabetes) outcomes. This paper highlights the variable weight trajectories following bariatric surgery.


183. Eid GM, McCloskey CA, Eagleton JK, Lee LB, Courcoulas AP. StomaphyX vs a sham procedure for revisional surgery to reduce regained weight in Roux-en-Y gastric bypass patients:


204. Adams TD, Davidson LE, Litwin SE, et al. Weight and metabolic outcomes 12 years after gastric bypass. *N Engl J Med.* 2017;377(12):1143-1155. This is the longest term paper to date from the Utah Obesity study which compares gastric bypass to 2 different control groups for weight and health outcomes, including type 2 diabetes.


215. Carlin AM, Zeni TM, English WJ, et al. The comparative effectiveness of sleeve gastrectomy, gastric bypass, and adjustable gastric banding procedures for the treatment of morbid obesity. *Ann Surg.* 2013;257(5):791-797. This is data from a statewide clinical registry, where outcomes of the 3 most common procedures; gastric bypass, sleeve gastrectomy, and laparoscopic adjustable banding are compared. It showed that rates for both short term weight loss and overall complications for the sleeve gastrectomy were positioned between gastric bypass and laparoscopic adjustable banding.


236. Schauer PR, Bhatt DL, Kirwan JP, et al. Bariatric surgery versus intensive medical therapy for diabetes—5-year outcomes. N Engl J Med. 2017;376(7):641-651. This is a randomized clinical trial comparing gastric bypass, gastric sleeve, and intensive medical management for the treatment of type 2 diabetes in people with obesity. It is only one of 2 randomized studies with 5 year follow up, at this time. It shows that surgical treatments are superior to intensive medical treatment for glycemic control.

CHAPTER 27
THE SURGICAL MANAGEMENT OF OBESITY

[261-OR]. Presented at the American Diabetes Association’s 76th Scientific Session; June 10-14, 2016; New Orleans, LA.


INTRODUCTORY COMMENTS

The small intestine is a remarkable and complex organ that is not only the principle site of nutrient digestion and absorption but also contains the body’s largest reservoir of immunologically active and hormone-producing cells. Hence, it can be conceptualized as the largest organ of the immune and endocrine systems.1 It achieves this diversity of action through unique anatomical features, which provide it with a massive surface area, a diversity of cell types, and a complex neural network to coordinate these functions.

Despite the size and importance of the small intestine, diseases of this organ are relatively infrequent and can present diagnostic and therapeutic challenges. Despite introduction of novel imaging techniques such as capsule endoscopy and double balloon endoscopy, diagnostic tests lack sufficient ability to reliably assess the small bowel. Furthermore, few high-quality, controlled data on the efficacy of surgical therapies for small bowel diseases are available.

Therefore, sound clinical judgment and a thorough understanding of anatomy, physiology, and pathophysiology remain essential to the care of patients with suspected small bowel disorders.

GROSS ANATOMY

The small intestine is a tubular structure that extends from the pylorus to the cecum. The estimated length varies depending on whether radiologic, surgical, or autopsy measurements are made. In the living, it is thought to measure 4 to 6 meters.2 The small intestine consists of three segments lying in series: the duodenum, the jejunum, and the ileum. The duodenum, the most proximal segment, lies in the retroperitoneum immediately adjacent to the head and inferior border of the body of the pancreas. The duodenum is demarcated from the stomach by the pylorus and from the jejunum by the ligament of Treitz. The jejunum and ileum lie within the peritoneal cavity and are tethered to the
Key Points

1. The small intestine performs a diverse set of functions.
2. Small bowel obstruction is one of the most common surgical diagnoses.
3. Most cases of small bowel obstruction are due to adhesions from previous surgery and resolve with conservative management.
4. Tumors and malignancies of the small bowel are rare and difficult to diagnose.
5. If following surgical resection less than 200 cm of small bowel remains, patients are at risk of developing short bowel syndrome.

The mucosa is the innermost layer and it consists of three layers: epithelium, lamina propria, and muscularis mucosae. The epithelium is exposed to the intestinal lumen and is the surface through which absorption from and secretion into the lumen occurs. The lamina propria is located immediately external to the epithelium and consists of connective tissue and a heterogeneous population of cells. It is demarcated from the more external submucosa by the muscularis mucosae, a thin sheet of smooth muscle cells.

The mucosa is organized into villi and crypts (crypts of Lieberkuhn). Villi are finger-like projections of epithelium and underlying lamina propria that contain blood and lymphatic (lacteals) vessels that extend into the intestinal lumen. Intestinal, epithelial cellular proliferation is confined to the crypts, each of which carries 250 to 300 cells. All epithelial cells in each crypt are derived from an unknown number of multipotent stem cells located at or near the crypt’s base. Our understanding of these crypt cells is rapidly expanding. It appears that there are two subgroups of intestinal stem cells, with specific cell markers. Bmi1-positive cells are usually quiescent, radiation-resistant cells that are induced by injury, while LGR5-positive cells facilitate homeostatic vs. injury-induced regeneration and are radiation sensitive.1

The stem cells can differentiate along one of four pathways that ultimately yield enterocytes and goblet, enteroendocrine, and Paneth cells. Except for Paneth cells, these lineages complete their terminal differentiation during an upward migration from each crypt to adjacent villi. The journey from the crypt to the villus tip is completed in 2 to 5 days and terminates with cells being removed by apoptosis and/or exfoliation. Thus, the small-intestinal epithelium undergoes continuous renewal, making it one of the body’s most dynamic tissues. The high cellular turnover rate contributes to mucosal resiliency but also makes the intestine uniquely susceptible to certain forms of injury such as that induced by radiation and chemotherapy.

The small intestine contains internal mucosal folds known as plicae circulares or valvulae conniventes that are visible upon gross inspection. These folds are also visible radiographically and help in the distinction between small intestine and colon, which does not contain them, on abdominal radiographs. These folds are more prominent in the proximal intestine than in the distal small intestine. Other features evident on gross inspection that are more characteristic of the proximal than distal small intestine include larger circumference, thicker wall, less fatty mesentery, and longer vasa recta (Fig. 28-1). Gross examination of the small-intestinal mucosa also reveals aggregates of lymphoid follicles. Those follicles, located in the ileum, are the most prominent and are designated Peyer’s patches.

Most of the duodenum derives its arterial blood from branches of both the celiac and the superior mesenteric arteries. The distal duodenum, the jejunum, and the ileum derive their arterial blood from the superior mesenteric artery. Their venous drainage occurs via the superior mesenteric vein. Lymph drainage occurs through lymphatic vessels coursing parallel to corresponding arteries. This lymph drains through mesenteric lymph nodes to the cisterna chyli, then through the thoracic duct, and ultimately into the left subclavian vein. The parasympathetic and sympathetic innervation of the small intestine is derived from the vagus and splanchnic nerves, respectively.

HISTOLOGY

The wall of the small intestine consists of four distinct layers: mucosa, submucosa, muscularis propria, and serosa (Fig. 28-2).

![Jejunum and Ileum](Image)

Figure 28-1. Gross features of jejunum contrasted with those of ileum. Relative to the ileum, the jejunum has a larger diameter, a thicker wall, more prominent plicae circulares, a less fatty mesentery, and longer vasa recta.
Enterocytes are the predominant absorptive cell of the intestinal epithelium. Their apical (lumen-facing) cell membrane contains specialized digestive enzymes, transporter mechanisms, and microvilli that are estimated to increase the absorptive surface area of the small intestine by up to 40-fold. Goblet cells produce mucin believed to play a role in mucosal defense against pathogens. Enteroendocrine cells are characterized by secretory granules containing regulatory agents and are discussed in greater detail in the “Endocrine Function” section. Paneth cells are located at the base of the crypt and contain secretory granules containing growth factors, digestive enzymes, and antimicrobial peptides, through which they control the host-microbe interaction and influence the intestinal microbiome. In addition, the intestinal epithelium contains M cells and intraepithelial lymphocytes. These two components of the immune system are discussed in this chapter.

The submucosa consists of dense connective tissue and a heterogeneous population of cells, including leukocytes and fibroblasts. The submucosa also contains an extensive network of vascular and lymphatic vessels, nerve fibers, and ganglion cells of the submucosal (Meissner’s) plexus.

The muscularis propria consists of an outer, longitudinally-oriented layer and an inner, circularly-oriented layer of smooth muscle fibers. Located at the interface between these two layers are ganglion cells of the myenteric (Auerbach’s) plexus.

The serosa consists of a single layer of mesothelial cells and is a component of the visceral peritoneum.

**DEVELOPMENT**

The first recognizable precursor of the small intestine is the embryonic gut tube, formed from the endoderm during the fourth week of gestation. The gut tube is divided into foregut, midgut, and hindgut. Other than the duodenum, which is a foregut structure, the rest of the small intestine is derived from the midgut. The gut tube initially communicates with the yolk sac; however, the communication between these two structures narrows by the sixth week to form the vitelline duct. The yolk sac and vitelline duct usually undergo obliteration by the end of gestation. Incomplete obliteration of the vitelline duct results in the spectrum of defects associated with Meckel’s diverticuli.

Also during the fourth week of gestation, the mesoderm of the embryo splits. The portion of mesoderm that adheres to the endoderm forms the visceral peritoneum, while the portion that adheres to the ectoderm forms the parietal peritoneum. This mesodermal division results in the formation of a coelomic cavity that is the precursor of the peritoneal cavity.

At approximately the fifth week of gestation, the bowel begins to lengthen to an extent greater than that which can be accommodated by the developing abdominal cavity, resulting in the extracoelomic herniation of the developing bowel. The bowel continues to lengthen during the subsequent weeks and is retracted back into the abdominal cavity during the tenth week of gestation. Subsequently, the duodenum becomes a retroperitoneal structure. Coincident with extrusion and retraction, the bowel undergoes a 270° counterclockwise rotation relative to the posterior abdominal wall. This rotation accounts for the usual locations of the cecum in the right lower quadrant and the duodenojejunal junction to the left of midline (Fig. 28-3).

The celiac and superior mesenteric arteries and veins are derived from the vitelline vascular system, which in turn is derived from blood vessels formed within the splanchnopleuric mesoderm during the third week of gestation. Neurons found in the small intestine are derived from neural crest cells that begin to migrate away from the neural tube during the third week of gestation. These neural crest cells enter the mesenchyme of the primitive foregut and subsequently migrate to the remainder of the bowel.

During the sixth week of gestation, the lumen of the developing bowel becomes obliterated as bowel epithelial proliferation accelerates. Vacuoles form within the bowel substance during the subsequent weeks and coalesce to form the intestinal lumen by the ninth week of gestation. Errors in this recanalization may account for defects such as intestinal webs and stenoses. Most intestinal atresias, however, are believed to be related to ischemic episodes occurring after organogenesis has been completed rather than to errors in recanalization.

During the ninth week of gestation, the intestinal epithelium develops intestine-specific features such as crypt-villus architecture. Organogenesis is complete by approximately the twelfth week of gestation.
Figure 28-3. Developmental rotation of the intestine. A. During the fifth week of gestation, the developing intestine herniates out of the coelomic cavity and begins to undergo a counterclockwise rotation about the axis of the superior mesenteric artery. B and C. Intestinal rotation continues, as the developing transverse colon passes anterior to the developing duodenum. D. Final positions of the small intestine and colon resulting from a 270° counterclockwise rotation of the developing intestine and its return into the abdominal cavity.

**PHYSIOLOGY**

**Digestion and Absorption**

The intestinal epithelium is the interface through which absorption and secretion occur. It has features characteristic of absorptive epithelia in general, including epithelial cells with cellular membranes possessing distinct apical (luminal) and basolateral (serosal) domains demarcated by intercellular tight junctions and an asymmetric distribution of transmembrane transporter mechanisms that promotes vectorial transport of solutes across the epithelium.

Solute can traverse the epithelium by active or passive transport. Passive transport of solutes occurs through diffusion or convection and is driven by existing electrochemical gradients. Active transport is the energy-dependent net transfer of solutes in the absence of or against an electrochemical gradient.

Active transport occurs through transcellular pathways (through the cell), whereas passive transport can occur through either transcellular or paracellular pathways (between cells through the tight junctions). Transcellular transport requires solutes to traverse the cell membranes through specialized membrane proteins, such as channels, carriers, and pumps. The molecular characterization of transporter proteins is evolving rapidly, with different transporter families, each containing many individual genes encoding specific transporters, now identified. Similarly, understanding of the paracellular pathway is evolving. In contrast to what was once believed, it is becoming apparent that paracellular permeability is substrate-specific, dynamic, and subject to regulation by specific tight junction proteins.

**Water and Electrolyte Absorption and Secretion.** Eight to 9 L of fluid enter the small intestine daily. Most of this volume consists of salivary, gastric, biliary, pancreatic, and intestinal secretions. Under normal conditions, the small intestine absorbs over 80 percent of this fluid, leaving approximately 1.5 L that enters the colon (Fig. 28-4). Small-intestinal absorption and secretion are tightly regulated; derangements in water and electrolyte homeostasis characteristic of many of the disorders discussed in this chapter play an important role in contributing to their associated clinical features.

Gut epithelia have two pathways for water transport: (a) the paracellular route, which involves transport through the spaces between cells, (b) the transcellular route, through apical and the basolateral cell membranes, with most occurring through
the transcellular pathway. The specific transport mechanisms mediating this transcellular transport are not completely characterized, and they may involve passive diffusion through the phospholipid bilayer, cotransport with other ions and nutrients, or diffusion through water channels called aquaporins. Many different types of aquaporins have been identified; however, their contribution to overall intestinal water absorption appears to be relatively minor.

The prevailing model for intestinal epithelial Na⁺ absorption is shown in Fig. 28-5. Activity of the Na⁺/K⁺ ATPase enzyme, which is located in the basolateral membrane and exchanges three intracellular Na⁺ for every two extracellular K⁺ in an energy-dependent process, generates the electrochemical gradient that drives the transport of Na⁺ from the intestinal lumen into the cytoplasm of enterocytes. Na⁺ ions traverse the apical membrane through several distinct transporter mechanisms, including nutrient-coupled sodium transport (e.g., sodium glucose cotransporter-1, SGLT1), sodium channels, and sodium-hydrogen exchangers (NHEs). Absorbed Na⁺ ions are then extruded from enterocytes through the Na⁺/K⁺ ATPase located in the basolateral membrane. Similar mechanistic models that account for the transport of other common ions such as K⁺ and HCO₃⁻ also exist.

Substantial heterogeneity, with respect to both crypt-villus and craniocaudal axes, exists for intestinal epithelial transport mechanisms. This spatial distribution pattern is consistent with a model in which absorptive function resides primarily in the villus and secretory function in the crypt.

Intestinal absorption and secretion are subject to modulation under physiologic and pathophysiologic conditions by a wide array of hormonal, neural, and immune regulatory mediators (Table 28-1).

### Table 28-1

<table>
<thead>
<tr>
<th>Regulation of intestinal absorption and secretion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agents that stimulate absorption or inhibit secretion of water</td>
</tr>
<tr>
<td>Aldosterone</td>
</tr>
<tr>
<td>Glucocorticoids</td>
</tr>
<tr>
<td>Angiotensin</td>
</tr>
<tr>
<td>Norepinephrine</td>
</tr>
<tr>
<td>Epinephrine</td>
</tr>
<tr>
<td>Dopamine</td>
</tr>
<tr>
<td>Somatostatin</td>
</tr>
<tr>
<td>Neuropeptide Y</td>
</tr>
<tr>
<td>Peptide YY</td>
</tr>
<tr>
<td>Enkephalin</td>
</tr>
<tr>
<td>Agents that simulate secretion or inhibit absorption of water</td>
</tr>
<tr>
<td>Secretin</td>
</tr>
<tr>
<td>Bradykinin</td>
</tr>
<tr>
<td>Prostaglandins</td>
</tr>
<tr>
<td>Acetylcholine</td>
</tr>
<tr>
<td>Atrial natriuretic factor</td>
</tr>
<tr>
<td>Vasopressin</td>
</tr>
<tr>
<td>Vasoactive intestinal peptide</td>
</tr>
<tr>
<td>Bombesin</td>
</tr>
<tr>
<td>Substance P</td>
</tr>
<tr>
<td>Serotonin</td>
</tr>
<tr>
<td>Neuropeptide Y</td>
</tr>
<tr>
<td>Somatostatin</td>
</tr>
<tr>
<td>Dopamine</td>
</tr>
<tr>
<td>Epinephrine</td>
</tr>
<tr>
<td>Norepinephrine</td>
</tr>
<tr>
<td>Angiotensin</td>
</tr>
<tr>
<td>Aldosterone</td>
</tr>
</tbody>
</table>

Carbohydrate Digestion and Absorption. Approximately 45% of energy consumption in the average Western diet consists of carbohydrates, approximately one-half of which is in the form of starch (linear or branched polymers of glucose) derived from cereals and plants. Other major sources of dietary carbohydrates include sugars derived from milk (lactose), fruits and vegetables (fructose, glucose, and sucrose), or purified from sugar cane or beets (sucrose). Processed foods contain a variety of sugars including fructose, oligosaccharides, and polysaccharides. Glycogen derived from meat contributes only a small fraction of dietary carbohydrate.

Pancreatic amylase is the major enzyme of starch digestion, although salivary amylase initiates the process. The terminal products of amylase-mediated starch digestion are oligosaccharides, maltotriose, maltose, and alpha-limit dextrins (Fig. 28-6). These products, as well as the major disaccharides in the diet (sucrose and lactose), are unable to undergo absorption in this form. They must first undergo hydrolytic cleavage into their constituent monosaccharides; these hydrolytic reactions are catalyzed by specific brush border membrane hydrolases that are expressed most abundantly in the villi of the duodenum and jejunum. The three major monosaccharides that represent the terminal products of carbohydrate digestion are glucose, galactose, and fructose.

Under physiologic conditions, most of these sugars are absorbed through the epithelium via the transcellular route. Glucose and galactose are transported through the enterocyte brush border membrane via intestinal Na⁺–glucose cotransporter, SGLT1 (Fig. 28-7). Fructose is transported through the brush border membrane by facilitated diffusion via GLUT5 (a member of the facilitative glucose transporter family). All three monosaccharides are extruded through the basolateral membrane by facilitated diffusion using GLUT2 and five transporters. Exerted monosaccharides diffuse into venules and ultimately enter the portal venous system.

There is evidence of overexpression of hexose transporters, specifically SGLT1, in disease states such as diabetes. Several approaches aimed at downregulation of small intestinal...
Carbohydrate digestion. Dietary carbohydrates, including starch and the disaccharides sucrose and lactose, must undergo hydrolysis into constituent monosaccharides glucose, galactose, and fructose before being absorbed by the intestinal epithelium. These hydrolytic reactions are catalyzed by salivary and pancreatic amylase and by enterocyte brush border hydrolases.

Protein digestion and absorption. Ten percent to 15% of energy consumption in the average Western diet consists of proteins. In addition to dietary proteins, approximately one-half of the protein load that enters the small intestine is derived from endogenous sources, including salivary and gastrointestinal secretions and desquamated intestinal epithelial cells. Protein digestion begins in the stomach with action of pepsins. This is not, however, an essential step because surgical patients who are acholorhydric, or have lost part or all their stomach, are still able to successfully digest proteins. Digestion continues in the duodenum with the actions of a variety of pancreatic peptidases. These enzymes are secreted as inactive proenzymes. This contrast with pancreatic amylase and lipase, which are secreted in their active forms. In response to the presence of bile acids, enteroerinase is liberated from the intestinal brush border membrane to catalyze the conversion of trypsinogen to active trypsin; trypsin in turn activates itself and other proteases. The final products of intraluminal protein digestion consist of neutral and basic amino acids and peptides two to six amino acids in length (Fig. 28-8). Additional digestion occurs through the actions of peptidases that exist in the enterocyte brush border and cytoplasm. Epithelial absorption occurs for both single amino acids and di- or tripeptides via specific membrane-bound transporters. Absorbed amino acids and peptides then enter the portal venous circulation.

Of all amino acids, glutamine appears to be a unique, major source of energy for enterocytes. Active glutamine uptake into enterocytes occurs through both apical and basolateral transport mechanisms.

Fat digestion and absorption. Approximately 40% of the average Western diet consists of fat. Over 95% of dietary fat is in the form of long-chain triglycerides; the remainder includes phospholipids such as lecithin, fatty acids, cholesterol, and...
Dietary long-chain triglycerides → Gastric lipase, Pancreatic lipase → Long-chain fatty acids and monoglycerides → Absorbed → Chyle (lymphatics)

Short- & medium-chain triglycerides → Absorbed → Portal venous blood

**Figure 28-9.** Fat digestion. Long-chain triglycerides, which constitute the majority of dietary fats, must undergo lipolysis into constituent long-chain fatty acids and monoglycerides before being absorbed by the intestinal epithelium. These reactions are catalyzed by gastric and pancreatic lipases. The products of lipolysis are transported in the form of mixed micelles to enterocytes, where they are resynthesized into triglycerides, which are then packaged in the form of chylomicrons that are secreted into the intestinal lymph (chyle). Triglycerides composed of short- and medium-chain fatty acids are absorbed by the intestinal epithelium directly, without undergoing lipolysis, and are secreted into the portal venous circulation.
bacteria. Other factors likely to play important roles in intestinal mucosal defense include antimicrobial peptides such as the defensins. The intestinal component of the immune system, known as the gut-associated lymphoid tissue (GALT), contains over 70% of the body’s immune cells.

The GALT is conceptually divided into inductive and effector sites. Inductive sites include Peyer’s patches, mesenteric lymph nodes, and smaller isolated lymphoid follicles scattered throughout the small intestine (Fig. 28-10). Peyer’s patches are macroscopic aggregates of B-cell follicles and intervening T-cell areas found in the lamina propria of the small intestine, primarily the distal ileum. Overlying Peyer’s patches is a specialized epithelium containing microfold (M) cells. These cells possess an apical membrane with microfolds rather than microvilli, which is characteristic of most intestinal epithelial cells. Using transepithelial vesicular transport, M cells transfer microbes to underlying professional antigen presenting cells (APCs), such as dendritic cells. Dendritic cells, in addition, may sample luminal antigens directly through their dendrite-like processes that extend through epithelial tight junctions. APCs interact with and prime naive lymphocytes, which then exit through the draining lymphatics to enter the mesenteric lymph nodes, where they undergo differentiation. These lymphocytes then migrate into the systemic circulation via the thoracic duct and ultimately accumulate in the intestinal mucosa at effector sites. Alternative induction mechanisms, such as antigen presentation within mesenteric lymph nodes, are also likely to exist.

Effector lymphocytes are distributed into distinct compartments. IgA-producing plasma cells are derived from B cells and are located in the lamina propria. CD4+ T cells are also located in the lamina propria. CD8+ T cells migrate preferentially to the epithelium, but they are also found in the lamina propria. These T cells are central to immune regulation; in addition, the CD8+ T cells have potent cytotoxic (CTL) activity. IgA is transported through the intestinal epithelial cells into the lumen, where it exists in the form of a dimer complexed with a secretory component. This configuration renders IgA resistant to proteolysis by digestive enzymes. IgA is believed to both help prevent the entry of microbes through the epithelium and to promote excretion of antigens or microbes that have already penetrated the laminal propria.

It has been increasingly recognized that the gastrointestinal tract is colonized with many bacteria that are essential for health. Communication between the microbiota and the host defense allows for protective immune responses against pathogens while preventing adverse inflammatory responses to harmless commensal microbes, which could lead to chronic inflammatory disorders such as celiac disease and Crohn’s disease.

**Motility**

Myocytes of the intestinal muscle layers are electrically and mechanically coordinated in the form of syncytia. Contractions of the muscularis propria are responsible for small-intestinal peristalsis. Contraction of the outer longitudinal muscle layer results in bowel shortening; contraction of the inner circular layer results in luminal narrowing. Contractions of the muscularis mucosa contribute to mucosal or villus motility, but not to peristalsis.

Several distinctive patterns of muscularis propria activity have been observed to occur in the small intestine. These patterns include ascending excitation and descending inhibition in which muscular contraction occurs proximal to a stimulus, such as the presence of a bolus of ingested food, and muscular relaxation occurs distal to the stimulus (Fig. 28-11). These two reflexes are present even in the absence of any extrinsic innervation to the small intestine and contribute to peristalsis when they are propagated in a coordinated fashion along the length of the intestine. The fed or postprandial pattern begins within 10 to 20 minutes of meal ingestion and abates 4 to 6 hours afterwards. Rhythmic segmentations or pressure waves traveling only short distances also are observed. This segmenting pattern is hypothesized to assist in mixing intraluminal contents and in facilitating their contact with the absorptive mucosal surface. The fasting pattern or interdigestive motor cycle (IDMC) consists of three phases. Phase 1 is characterized by motor quiescence, phase 2 by seemingly disorganized pressure waves occurring at submaximal rates, and phase 3 by sustained pressure waves occurring at maximal rates. This pattern is hypothesized to expel residual debris and bacteria from the small intestine. The median duration of the IDMC ranges from 90 to 120 minutes. At any given time, different portions of the small intestine can be in different phases of the IDMC.
“Gut hormones” were previously conceptualized as peptides produced by the enteroendocrine cells of the intestinal mucosa that are released into the systemic circulation to reach receptors in target sites in the gastrointestinal tract. Now it is clear that “gut hormone” genes are widely expressed throughout the body, not only in endocrine cells but also in central and peripheral neurons. The products of these genes are general intercellular messengers that can act as endocrine, paracrine, or neurocrine mediators. Thus, they may act as true blood-borne hormones as well as through local effects.

There are notable homology patterns among individual regulatory peptides found in the gastrointestinal tract. Based on these homologies, approximately one-half of the known regulatory peptides can be classified into families. For example, the secretin family includes secretin, glucagon, and glucagon-like peptides, glucose-dependent insulinotropic peptide, vasoactive intestinal polypeptide, peptide histidine isoleucine, growth hormone releasing hormone, and pituitary adenyl cyclase-activating peptide. Other peptide families include those named for insulin, epidermal growth factor, gastrin, pancreatic polypeptide, tachykinin, and somatostatin.

Receptor subtype multiplicity and cell-specific expression patterns for these receptor subtypes that are characteristic of these regulatory mediators makes definition of their actions complex. Detailed description of these actions is beyond the scope of this chapter; however, examples of regulatory peptides produced by enteroendocrine cells of the small-intestinal epithelium and their most commonly ascribed functions are summarized in Table 28-2. Some of these peptides, or their analogues, are used in routine clinical practice. For example,

### Table 28-2

<table>
<thead>
<tr>
<th>HORMONE</th>
<th>SOURCE*</th>
<th>ACTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Somatostatin</td>
<td>D cell</td>
<td>Inhibits gastrointestinal secretion, motility, and splanchnic perfusion</td>
</tr>
<tr>
<td>Secretin</td>
<td>S cell</td>
<td>Stimulates exocrine pancreatic secretion, stimulates intestinal secretion</td>
</tr>
<tr>
<td>Cholecystokinin</td>
<td>I cell</td>
<td>Simulates pancreatic exocrine secretion, simulates gallbladder emptying, inhibits sphincter of Oddi contraction</td>
</tr>
<tr>
<td>Motilin</td>
<td>M cell</td>
<td>Simulates intestinal motility</td>
</tr>
<tr>
<td>Peptide YY</td>
<td>L cell</td>
<td>Inhibits intestinal motility and secretion</td>
</tr>
<tr>
<td>Glucagon-like Peptide 2</td>
<td>L cell</td>
<td>Stimulates intestinal epithelial proliferation</td>
</tr>
<tr>
<td>Neurotensin</td>
<td>N cell</td>
<td>Stimulates pancreatic and biliary secretion, inhibits small bowel motility, stimulates intestinal mucoosal growth</td>
</tr>
</tbody>
</table>

*This table indicates which enteroendocrine cell types located in the intestinal epithelium produce these peptides. These peptides are also widely expressed in nonintestinal tissues.
therapeutic applications of octreotide, a long-acting analogue of somatostatin, include the amelioration of symptoms associated with neuroendocrine tumors (e.g., carcinoid syndrome), postgastrectomy dumping syndrome, enterocutaneous fistulas, and the initial treatment of acute hemorrhage due to esophageal varices. The gastrin secretory response to secretin administration forms the basis for the standard test used to establish the diagnosis of Zollinger-Ellison syndrome. Cholecystokinin is used in evaluations of gallbladder ejection fraction, a parameter that may have utility in patients who have symptoms of biliary colic but are not found to have gallstones. Of the peptides listed in Table 28-2, glucagon-like peptide 2 (GLP-2) has been identified as a specific and potent intestino-trophic hormone and is currently under clinical evaluation as an intestinotrophic agent in patients suffering from the short bowel syndrome, as discussed in the “Short Bowel Syndrome” section.

### Intestinal Adaptation

The small intestine has the capacity to adapt in response to varying demands imposed by physiologic and pathologic conditions. Of relevance to many of the diseases discussed in this chapter is the adaptation that occurs in the remnant intestine following surgical resection of a large portion of the small intestine (massive small bowel resection). Postresection intestinal adaptation has been studied extensively using animal models. Within a few hours after bowel resection, the remnant small intestine displays evidence of epithelial cellular hyperplasia. With additional time, villi lengthen, intestinal absorptive surface area increases, and digestive and absorptive functions improve. Postresection intestinal adaptation in human patients is less well studied, but it seems to follow similar steps as that seen in experimental models, and it takes 1 to 2 years to complete.1

The mechanisms responsible for inducing postresection intestinal adaptation are under active investigation. Several classes of effectors that stimulate intestinal growth include specific nutrients, peptide hormones and growth factors, pancreatic secretions, and some cytokines. Nutritional components with intestinal growth-stimulating effects include fiber, fatty acids, triglycerides, glutamine, polyamines, and lectins.

Postresection adaptation serves to compensate for the function of intestine that has been resected. Jejunal resection is generally better tolerated, as ileum shows better capacity to compensate. However, the magnitude of this response is limited. If enough small intestine is resected, a devastating condition known as the short bowel syndrome results. This condition is discussed in the “Short Bowel Syndrome” section of this chapter.

### SMALL BOWEL OBSTRUCTION

#### Epidemiology

Mechanical small bowel obstruction is the most frequently encountered surgical disorder of the small intestine.2 Although a wide range of etiologies for this condition exist, the obstructing lesion can be conceptualized according to its anatomical relationship to the intestinal wall as:

1. **intraluminal** (e.g., foreign bodies, gallstones, or meconium)
2. **intramural** (e.g., tumors, Crohn’s disease–associated inflammatory strictures)
3. **extrinsic** (e.g., adhesions, hernias, or carcinomatosis)

Intra-abdominal adhesions related to prior abdominal surgery account for up to 75% of cases of small bowel obstruction. Over 300,000 patients are estimated to undergo surgery to treat adhesion-induced small bowel obstruction in the United States annually. A 20-year trend analysis between 1988 and 2007 has documented no decrease in this rate during this period, highlighting the ongoing problem of this “old” disease.3 In fact, small bowel resection and lysis of adhesions account for two of the seven procedures that were responsible for 80% of the emergency surgeries in the United States between 2008 and 2011.4

Less prevalent etiologies for small bowel obstruction include hernias, malignant bowel obstruction, and Crohn’s disease. The frequency with which obstruction related to these conditions is encountered varies according to the patient population and practice setting. Cancer-related small bowel obstructions are commonly due to extrinsic compression or invasion by advanced malignancies arising in organs other than the small bowel; few are due to primary small bowel tumors. The most commonly encountered etiologies of small bowel obstruction are summarized in Table 28-3. Although congenital abnormalities capable of causing small bowel obstruction usually become evident during childhood, they sometimes elude detection and are diagnosed for the first time in adult patients presenting with abdominal symptoms. For example, intestinal malrotation and midgut volvulus should not be forgotten when considering the differential diagnosis of adult patients with acute or chronic symptoms of small bowel obstruction, especially those without a history of prior abdominal surgery. A rare etiology of obstruction is the superior mesenteric artery syndrome, characterized by compression of the third portion of the duodenum by the superior mesenteric artery as it crosses over this portion of the duodenum. This condition should be

#### Table 28-3

<table>
<thead>
<tr>
<th>Small bowel obstruction: common etiologies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Adhesions</strong></td>
</tr>
<tr>
<td><strong>Neoplasms</strong></td>
</tr>
<tr>
<td>Primary small bowel neoplasms</td>
</tr>
<tr>
<td>Secondary small bowel cancer (e.g., melanoma-derived metastasis)</td>
</tr>
<tr>
<td>Local invasion by intra-abdominal malignancy (e.g., Desmoid tumors)</td>
</tr>
<tr>
<td>Carcinomatosis</td>
</tr>
<tr>
<td><strong>Hernias</strong></td>
</tr>
<tr>
<td>External (e.g., inguinal and femoral)</td>
</tr>
<tr>
<td>Internal (e.g., following Roux-en-Y gastric bypass surgery)</td>
</tr>
<tr>
<td><strong>Crohn’s disease</strong></td>
</tr>
<tr>
<td><strong>Volvulus</strong></td>
</tr>
<tr>
<td><strong>Intussusception</strong></td>
</tr>
<tr>
<td>Radiation-induced stricture</td>
</tr>
<tr>
<td>Postischemic stricture</td>
</tr>
<tr>
<td><strong>Foreign body</strong></td>
</tr>
<tr>
<td>Gallstone ileus</td>
</tr>
<tr>
<td><strong>Diverticulitis</strong></td>
</tr>
<tr>
<td><strong>Meckel’s diverticulum</strong></td>
</tr>
<tr>
<td><strong>Hematoma</strong></td>
</tr>
<tr>
<td>Congenital abnormalities (e.g., webs, duplications, and malrotation)</td>
</tr>
</tbody>
</table>

1. **Table 28-3**
2. **introluminal** (e.g., foreign bodies, gallstones, or meconium)
3. **intramural** (e.g., tumors, Crohn’s disease–associated inflammatory strictures)
4. **extrinsic** (e.g., adhesions, hernias, or carcinomatosis)
Pathophysiology

With onset of obstruction, gas and fluid accumulate within the intestinal lumen proximal to the site of obstruction. The intestinal activity increases to overcome the obstruction, accounting for the colicky pain and the diarrhea that some experience even in the presence of complete bowel obstruction. Most of the gas that accumulates originates from swallowed air, although some is produced within the intestine. The fluid consists of swallowed liquids and gastrointestinal secretions (obstruction stimulates intestinal epithelial water secretion). With ongoing gas and fluid accumulation, the bowel distends and intraluminal and intramural pressures rise. The intestinal motility is eventually reduced with fewer contractions. With obstruction, the luminal flora of the small bowel, which is usually sterile, changes and a variety of organisms have been cultured from the contents. Translocation of these bacteria to regional lymph nodes has been demonstrated, although the significance of this process is not well understood. If the intramural pressure becomes high enough, intestinal microvascular perfusion is impaired leading to intestinal ischemia, and, ultimately, necrosis. This condition is termed strangulated bowel obstruction.

With partial small bowel obstruction, only a portion of the intestinal lumen is occluded, allowing passage of some gas and fluid. The progression of pathophysiologic events described previously tends to occur more slowly than with complete small bowel obstruction, and development of strangulation is less likely.

A particularly dangerous form of bowel obstruction is closed loop obstruction in which a segment of intestine is obstructed both proximally and distally (e.g., with volvulus). In such cases, the accumulating gas and fluid cannot escape either proximally or distally from the obstructed segment, leading to a rapid rise in luminal pressure and a rapid progression to strangulation.

Clinical Presentation

The symptoms of small bowel obstruction are colicky abdominal pain, nausea, vomiting, and obstipation. Vomiting is a more prominent symptom with proximal obstructions than distal. Character of vomitus is important as with bacterial overgrowth, the vomitus is more feculent, suggesting a more established obstruction. Continued passage of flatus and/or stool beyond 6 to 12 hours after onset of symptoms is characteristic of partial rather than complete obstruction. The signs of small bowel obstruction include abdominal distention, which is most pronounced if the site of obstruction is in the distal ileum and may be absent if the site of obstruction is in the proximal small intestine. Bowel sounds may be hyperactive initially, but in late stages of bowel obstruction, minimal bowel sounds may be heard. Laboratory findings reflect intravascular volume depletion and consist of hemoconcentration and electrolyte abnormalities. Mild leukocytosis is common.

Features of strangulated obstruction include abdominal pain often disproportionate to the degree of abdominal findings, suggestive of intestinal ischemia. Patients often have tachycardia, localized abdominal tenderness, fever, marked leukocytosis, and acidosis. Any of these findings should alert the clinician to the possibility of strangulation and the need for early surgical intervention.
SPECIFIC CONSIDERATIONS
PART II

Figure 28-12. Small bowel obstruction. Plain radiographs (A) supine, which show dilated loops of small bowel in the right upper quadrant; (B) erect, which confirm the presence of airfluid level in the loops of small bowel as well as the stomach, consistent with small bowel obstruction.

Figure 28-13. Small bowel obstruction. A CT scan of a patient presenting with signs and symptoms of bowel obstruction. Image shows grossly dilated loops of small bowel, with decompressed terminal ileum (I) and ascending colon (C), suggesting a complete distal small bowel obstruction. At laparotomy, adhesive bands from a previous surgery were identified and divided.

in the colon within 24 hours of administration is predictive of nonsurgical resolution of bowel obstruction with a sensitivity of 92% and a specificity of 93%.16

A limitation of CT scanning is its low sensitivity (<50%) in the detection of low-grade or partial small bowel obstruction. A subtle transition zone may be difficult to identify in the axial images obtained during CT scanning. In such cases, contrast examinations of the small bowel, either small bowel series (small bowel follow-through) or enteroclysis, can be helpful. For standard small bowel series, contrast is swallowed or instilled into the stomach through a nasogastric tube. Abdominal radiographs are then taken serially as the contrast travels distally in the intestine. Although barium can be used, water-soluble contrast agents, such as gastrograffin, should be used if the possibility of intestinal perforation exists. These examinations are more labor-intensive and less rapidly performed than CT scanning but may offer greater sensitivity in the detection of luminal and mural etiologies of obstruction, such as primary intestinal tumors. For enteroclysis, 200 to 250 mL of barium followed by 1 to 2 L of a solution of methylcellulose in water is instilled into the proximal jejunum via a long nasoenteric catheter. The double-contrast technique used in enterocolysis

Figure 28-14. Chronic partial small bowel obstruction. This patient presented with a several months history of chronic abdominal pain and intermittent vomiting. The coronal CT image shows grossly dilated loops of proximal small bowel on the left side (wide arrow), with decompressed loops of small bowel on the right side (narrow arrow). The dilated segment shows evidence of feculization of bowel contents, consistent with the chronic nature of the obstruction. Patient’s vomitus had characteristic feculent smell and quality. At exploratory laparotomy, adhesive bands were identified and divided.
permits a better assessment of mucosal surface and detection of relatively small lesions, even through overlapping small bowel loops. Enterocolysis is rarely performed in the acute setting but offers greater sensitivity than small bowel series in the detection of lesions that may be causing partial small bowel obstruction. Recently, CT enterocolysis has been used, and it was reported to be superior to plain X-ray small bowel contrast studies.

**Therapy**

Small bowel obstruction is usually associated with a marked depletion of intravascular volume due to decreased oral intake, vomiting, and sequestration of fluid in bowel lumen and wall. Therefore, fluid resuscitation is integral to treatment. Isotonic fluid should be given intravenously, and an indwelling bladder catheter may be placed to monitor urine output. Central-venous fluid should be given intravenously, and an indwelling bladder catheter may be placed to monitor urine output. Central-venous or pulmonary-artery catheter monitoring is not generally indicated unless the patient has underlying cardiac disease and severe dehydration. Broad-spectrum antibiotics are not indicated unless there is concern for bowel ischemia and surgery is planned.

The stomach should be continuously evacuated of air and fluid using a nasogastric (NG) tube. Effective gastric decompression decreases nausea, distention, and the risk of vomiting and aspiration. Longer nasoenteric tubes, with tips placed into the jejunum or ileum, were favored in the past but are rarely used today, as they are associated with higher complication rates than NG tubes, with no proven greater efficacy in several studies.

While a period of close observation and nonoperative management has been the mainstay of treatment for partial bowel obstruction, the standard therapy for complete small bowel obstruction has generally been expeditious surgery, with the dictum that “the sun should never rise and set on a complete bowel obstruction.” The rationale for favoring early surgical intervention is to minimize the risk for bowel strangulation, which is associated with an increased risk for morbidity and mortality. Clinical signs and currently available laboratory tests and imaging studies do not reliably permit the distinction between patients with simple obstruction and those with strangulated obstruction prior to the onset of irreversible ischemia. Therefore, the goal is to operate before the onset of irreversible ischemia. This treatment approach has, however, undergone significant reassessment in recent years, with many advocating for nonoperative approaches in management of these patients, providing closed-loop obstruction is ruled out and there is no evidence of intestinal ischemia. In a study of 145 patients with CT-diagnosed high-grade complete small bowel obstruction, 46% of the overall cohort were managed nonoperatively. More specifically, of the 104 patients who did not meet criteria for immediate surgery, 66 patients were successfully managed nonoperatively.

Thus, conservative therapy in the form of NG decompression and fluid resuscitation is now commonly recommended in the initial management of nonischemic bowel obstruction. Nonoperative management has been documented to be successful in 65% to 81% of patients with partial small bowel obstruction. Of those successfully treated nonoperatively, only 5% to 15% have been reported to have symptoms that were not substantially improved within 48 hours after initiation of therapy. Therefore, most patients with partial small obstruction whose symptoms do not improve within 48 hours after initiation of nonoperative therapy should be considered for surgery. In a study using the National Inpatient Sample, this principle was further highlighted. The authors concluded that a 2-day limit of watchful waiting before surgery is not associated with an increase in mortality or postoperative morbidity, although inpatient costs were higher.

The observation that administration of water-soluble oral contrast has not only diagnostic but also therapeutic and prognostic value has led to the creation of several protocols and pathways for management of patients presenting with small bowel obstruction. An example of such a pathway that is utilized at our institution is outlined in Fig. 28-16. Several studies and subsequent meta-analyses have shown that use of water-soluble contrast not only predicts likelihood of success of nonoperative management but also reduces the need for surgery (odds ratio 0.44), length of stay by about 2 days, and time to resolution by about 28 hours, without an increase in morbidity or mortality.

The operative procedure performed for small bowel obstruction varies according to the etiology of the obstruction. For example, adhesions are lysed, tumors are resected, and hernias are reduced and repaired. Regardless of the etiology, the affected intestine should be examined, and nonviable bowel should be resected. Criteria suggesting viability are normal color, peristalsis, and marginal arterial pulsations. Usually, visual inspection alone is adequate in judging viability. In borderline cases, a Doppler probe may be used to check for pulsatile flow to the bowel, and arterial perfusion can be verified by visualizing intravenously administered fluorescein dye in the bowel wall under ultraviolet illumination. Neither technique has, however, been found to be superior to clinical judgment. In general, if the patient is hemodynamically stable, short lengths of bowel of questionable viability should be resected, and primary anastomosis of the remaining intestine should be performed. However, if the viability of a large proportion of the intestine is in question, a concerted effort to preserve intestinal tissue should be made. In such situations, the bowel of uncertain viability should be left intact and the patient reexplored in 24 to 48 hours in a “second-look” operation. At that time, definitive resection of nonviable bowel is completed.

**Figure 28-15.** Intestinal pneumatosis. This CT scan shows intestinal pneumatosis (arrow). The cause of this radiological finding was intestinal ischemia. Patient was taken emergently to the operating room and underwent resection of an infarcted segment of small bowel.
Successful laparoscopic surgery for bowel obstruction is being reported with greater frequency. In a propensity score-matched study of patients who underwent adhesiolysis for small bowel obstruction, the laparoscopic approach was associated with significantly lower rates of overall complications, surgical site infections, and a shorter length of hospital stay (4 vs. 10 days). Since distended loops of bowel can interfere with adequate visualization, early cases of proximal small bowel obstruction that are likely due to a single adhesive band are best suited for this approach. Presence of bowel distention and multiple adhesions can cause these procedures to be difficult, with a reported conversion rate of 17% to 33%. One of the major concerns with the laparoscopic approach has been the risk of iatrogenic bowel injury. A pooled analysis of 11 nonrandomized comparative studies has, however, shown that the risk of bowel injury and reoperation were not different between the two procedures, although the laparoscopic approach was associated with greater surgical time.

**Outcomes**

The perioperative mortality rate associated with surgery for nonstrangulating small bowel obstruction is less than 5%, with most deaths occurring in elderly patients with significant comorbidities. Mortality rates associated with surgery for strangulated obstruction is higher, highlighting the need for prompt intervention in this group. Long-term prognosis is related to the etiology of obstruction. Many patients who are treated conservatively for adhesive small bowel obstruction do not require future readmissions; less than 20% of such patients will have a readmission over the subsequent 5 years with another episode of bowel obstruction.

In a study of 286 patients who had undergone surgical intervention for adhesive small bowel obstruction, the risk of recurrent obstruction was 5.5% at 1 year, 11.3% at 3 years, and 13.5% at 5 years. The risk of reoperation for recurrent obstruction was 3.7% at 1 year, 4.8% at 3 years, and 5.8% at 5 years. Considering the frequency of small bowel obstruction and the varied degree of clinical severity and presentation, there is often variation in the care of patients admitted with bowel obstruction. Studies have shown that a standard hospital-wide policy can help improve care of patients with bowel obstruction, reducing their time to surgery and shortening their length of hospital stay.
**Prevention**

With adhesive small bowel obstruction representing a large therapeutic burden, prevention of postoperative adhesions has become an area of great interest. Good surgical technique, careful handling of tissue, and minimal use and exposure of peritoneum to foreign bodies, forms the cornerstone of adhesion prevention. These measures alone are often inadequate. In patients undergoing colorectal or pelvic surgery, hospital readmission rates of greater than 30% over the subsequent 10 years have been reported for adhesive small bowel obstruction.²⁵

Use of laparoscopic surgery, when possible, has been strongly promoted. A recent study using the Swedish National Inpatient Register has shown that, compared to laparoscopy, open surgery is associated with a fourfold increase in risk of small bowel obstruction within 5 years of the index procedure, even after accounting for other risk factors such as age, comorbidity, and previous abdominal surgery.²⁴

In those undergoing open surgery, several strategies for adhesion prevention have been tried; however, the only therapy that has shown some success has been the use of hyaluronan-based agents, such as Sperafilm. The use of this barrier has been clearly shown to reduce the incidence of postoperative bowel adhesions; however, their effect in actually reducing the incidence of small bowel obstruction remains less well defined.²⁵ The use of these products is often left to the discretion of the surgeon and the clinical context. Wrapping of an intestinal anastomosis with the material may be associated with increased leak rates and is generally discouraged.²⁶

**Other Causes of Small Bowel Obstruction**

*Early postoperative bowel obstruction*, as defined by signs, symptoms, and radiographic signs of SBO occurring within 30 days following surgery, been reported to occur in 0.7% to 9% of patients, with a higher rate in patients undergoing pelvic surgery, especially colorectal procedures.²⁷ CT scanning or small bowel series is often required to make the diagnosis. Obstruction that occurs in the early postoperative period is usually partial and only rarely is associated with strangulation. Therefore, a period of extended nonoperative therapy (2–3 weeks) consisting of bowel rest, hydration, and TPN administration is usually warranted. However, if complete obstruction is demonstrated or if signs suggestive of peritonitis are detected, expeditious reoperation should be undertaken without delay. In a series of 180 patients undergoing anterior resection for rectal cancer, 12.8% developed early postoperative bowel obstruction on the median postoperative day 5, with 4 requiring surgical exploration at a median interval of 2 weeks from the index case.²⁷

*Crohn’s disease* as a cause of small bowel obstruction is discussed in more detail later in this chapter in the “Crohn’s Disease” section.

*Malignant small bowel obstruction* can be a challenging problem. Although it often indicates advanced disease with poor prognosis, 25% to 33% of patients with a history of cancer who present with small bowel obstruction have adhesions as the etiology of their obstruction and therefore should not be denied appropriate therapy. Even in cases in which the obstruction is related to recurrent malignancy, palliative resection or bypass can be performed, and in select cases these procedures lead to improved quality of life. In a series of 81 patients with small bowel obstruction, palliation was achieved in over 80% of patients, with over 70% able to reestablish oral intake. In this series, the surgical morbidty was high, with 7% developing an enterocutaneous fistula/aneastomotic leak and a 30-day mortality rate of 6%.²⁸ Patients with obvious carcinomatosis and multifocal obstruction pose a difficult challenge, given their limited prognosis. Thus, management must be tailored to an individual patient’s prognosis and desires. At the time of surgery, relief of the obstruction may be best achieved by a bypass procedure, avoiding a potentially difficult bowel resection, and even if that is not feasible, a palliative gastrostomy tube can be considered to help resolve nausea and vomiting.

**ILEUS AND OTHER DISORDERS OF INTESTINAL MOTILITY**

Ileus and intestinal pseudo-obstruction are clinical syndromes caused by impaired intestinal motility and are characterized by symptoms and signs of intestinal obstruction in the absence of a lesion-causing mechanical obstruction. Ileus is a temporary motility disorder that is reversed with time as the inciting factor is corrected. In contrast, chronic intestinal pseudo-obstruction comprises a spectrum of specific disorders associated with irreversible intestinal dysmotility.

Ileus is a major cause of morbidity in hospitalized patients. A degree of intestinal ileus is a normal physiological response to abdominal surgery, which often resolves quickly without any long-term sequelae. However, when postoperative ileus is prolonged, it can cause significant morbidity and cost. Prolonged postoperative ileus is the most frequently implicated cause of delayed discharge following abdominal operations, and its economic impact has been estimated to be between $750 million and $1 billion annually in the United States.²⁹

**Pathophysiology**

Numerous factors capable of impairing intestinal motility, and thus inciting ileus, have been described (Table 28–4). The most

<table>
<thead>
<tr>
<th>Table 28–4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ileus: common etiologies</strong></td>
</tr>
<tr>
<td>Abdominal surgery</td>
</tr>
<tr>
<td>Infection</td>
</tr>
<tr>
<td>Sepsis</td>
</tr>
<tr>
<td>Intra-abdominal abscess</td>
</tr>
<tr>
<td>Peritonitis</td>
</tr>
<tr>
<td>Pneumonia</td>
</tr>
<tr>
<td>Electrolyte abnormalities</td>
</tr>
<tr>
<td>Hypokalemia</td>
</tr>
<tr>
<td>Hypomagnesemia</td>
</tr>
<tr>
<td>Hypophosphatemia</td>
</tr>
<tr>
<td>Medications</td>
</tr>
<tr>
<td>Anticholinergics</td>
</tr>
<tr>
<td>Opiates</td>
</tr>
<tr>
<td>Phenothiazines</td>
</tr>
<tr>
<td>Calcium channel blockers</td>
</tr>
<tr>
<td>Tricyclic antidepressants</td>
</tr>
<tr>
<td>Hypothyroidism</td>
</tr>
<tr>
<td>Ureteral colic</td>
</tr>
<tr>
<td>Retropertioneal hemorrhage</td>
</tr>
<tr>
<td>Spinal cord injury</td>
</tr>
<tr>
<td>Myocardial infarction</td>
</tr>
<tr>
<td>Mesenteric ischemia</td>
</tr>
</tbody>
</table>
frequently encountered factors are abdominal operations, infection and inflammation, electrolyte abnormalities, and drugs. Following most abdominal operations or injuries, the motility of the gastrointestinal tract is transiently impaired. Among the proposed mechanisms responsible for this dysmotility are surgical stress-induced sympathetic reflexes, inflammatory response mediator release, and anesthetic/analgesic side effects; each of which can inhibit intestinal motility. The return of normal motility generally follows a characteristic temporal sequence, with small-intestinal motility returning to normal within the first 24 hours after laparotomy and gastric and colonic motility returning to normal by 48 hours and 2 to 5 days, respectively. Since small bowel motility is returned before colonic and gastric motility, listening for bowel sounds is not a reliable indicator that ileus has fully resolved. Functional evidence of coordinated gastrointestinal motility in the form of passing flatus or bowel movement is a more useful indicator. Resolution of ileus may be delayed in the presence of other factors capable of inciting ileus such as the presence of intra-abdominal abscesses or electrolyte abnormalities.

Chronic intestinal pseudo-obstruction can be caused by a large number of specific abnormalities affecting intestinal smooth muscle, the myenteric plexus, or the extraintestinal nervous system (Table 28-5). Visceral myopathies constitute a group of diseases characterized by degeneration and fibrosis of the intestinal muscularis propria. Visceral neuropathies encompass a variety of degenerative disorders of the myenteric and submucosal plexuses. Both sporadic and familial forms of visceral myopathies and neuropathies exist. Systemic disorders involving the smooth muscle such as progressive systemic sclerosis and progressive muscular dystrophy, and neurological diseases such as Parkinson’s disease, can also be complicated by chronic intestinal pseudo-obstruction. In addition, viral infections, such as those associated with cytomegalovirus and Epstein-Barr virus, can cause intestinal pseudo-obstruction.

### Table 28-5

<table>
<thead>
<tr>
<th>Chronic intestinal pseudo-obstruction: etiologies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary Causes</strong></td>
</tr>
<tr>
<td>Familial types</td>
</tr>
<tr>
<td><em>Familial visceral myopathies (types I, II, and III)</em></td>
</tr>
<tr>
<td><em>Familial visceral neuropathies (types I and II)</em></td>
</tr>
<tr>
<td><em>Childhood visceral myopathies (types I and II)</em></td>
</tr>
<tr>
<td>Sporadic types</td>
</tr>
<tr>
<td><em>Visceral myopathies</em></td>
</tr>
<tr>
<td><em>Visceral neuropathies</em></td>
</tr>
<tr>
<td><strong>Secondary Causes</strong></td>
</tr>
<tr>
<td>Smooth muscle disorders</td>
</tr>
<tr>
<td><em>Collagen vascular diseases (e.g., scleroderma)</em></td>
</tr>
<tr>
<td><em>Muscular dystrophies (e.g., myotonic dystrophy)</em></td>
</tr>
<tr>
<td><em>Amyloidosis</em></td>
</tr>
<tr>
<td>Neurological disorders</td>
</tr>
<tr>
<td><em>Chagas disease, Parkinson’s disease, spinal cord injury</em></td>
</tr>
<tr>
<td>Endocrine disorders</td>
</tr>
<tr>
<td><em>Diabetes, hypothyroidism, hypoparathyroidism</em></td>
</tr>
<tr>
<td>Miscellaneous disorders</td>
</tr>
<tr>
<td><em>Radiation enteritis</em></td>
</tr>
<tr>
<td>Pharmacological causes</td>
</tr>
<tr>
<td><em>E.g., phenothiazines and tricyclic antidepressants</em></td>
</tr>
<tr>
<td>Viral infections</td>
</tr>
</tbody>
</table>

### Clinical Presentation

The clinical presentation of ileus resembles that of small bowel obstruction. Inability to tolerate liquids and solids by mouth, nausea, and lack of flatus or bowel movements are the most common symptoms. Vomiting and abdominal distension may occur. Although bowel sound characteristics are not diagnostic, they are usually diminished or absent, in contrast to the hyperactive bowel sounds that usually accompany mechanical small bowel obstruction. The clinical manifestations of chronic intestinal pseudo-obstruction include variable degrees of nausea and vomiting and abdominal pain and distention.

### Diagnosis

Routine postoperative ileus should be expected and requires no diagnostic evaluation. Definition of prolonged postoperative ileus has been varied but generally diagnosed if ileus persists beyond 5 days postoperatively. A recent global survey synthesized the results of the data to define postoperative ileus as “interval from surgery until passage of flatus/stool AND tolerance of an oral diet,” with prolonged postoperative ileus being defined as “two or more of nausea/vomiting, inability to tolerate oral diet over 24 h, absence of flatus over 24 h, distension, radiologic confirmation occurring on or after day 4 postoperatively without prior resolution of postoperative ileus.”

Once suspected, diagnostic evaluation to detect specific underlying factors capable of inciting ileus and to rule out the presence of mechanical obstruction is warranted.

Patient medication lists should be reviewed for the presence of drugs, especially opiates, known to be associated with impaired intestinal motility. Measurement of serum electrolytes may demonstrate electrolyte abnormalities commonly associated with ileus. Abdominal radiographs are often obtained, but the distinction between ileus and mechanical obstruction may be difficult based on this test alone. In the postoperative setting, CT scanning is the test of choice as it can demonstrate the presence of an intra-abdominal abscess or other evidence of peritoneal sepsis that may be causing ileus and can exclude the presence of complete mechanical obstruction. Distinction of postoperative ileus from early postoperative obstruction can be difficult but is helpful in developing the appropriate management plan.

The diagnosis of chronic pseudo-obstruction is suggested by clinical features and confirmed by radiographic and manometric studies. Diagnostic laparotomy or laparoscopy with full-thickness biopsy of the small intestine may be required to establish the specific underlying cause in cases of suspected neural disorder.

### Therapy

The management of ileus consists of limiting oral intake and correcting the underlying inciting factor. If vomiting or abdominal distention are prominent, the stomach should be decompressed using a nasogastric tube. Fluid and electrolytes should be administered intravenously until ileus resolves. If the duration of ileus is prolonged, total parental nutrition (TPN) may be required.

Given the frequency of postoperative ileus and its financial impact, many strategies have been tested to reduce its duration. The administration of nonsteroidal anti-inflammatory drugs such as ketorolac and concomitant reductions in opioid dosing have been shown to reduce the duration of ileus in most studies. Similarly, the use of perioperative thoracic epidural anesthesia/analgesia with
regimens containing local anesthetics combined with limitation or elimination of systemically administered opioids has been shown to reduce duration of postoperative ileus, although they have not reduced the overall length of hospital stay. Many studies have also suggested that limiting intra- and postoperative fluid administration can also result in reduction of postoperative ileus and shortened hospital stay. Furthermore, studies have shown that early postoperative feeding after GI surgery is generally well tolerated and can lead to reduced postoperative ileus and a shorter hospital stay. Table 28-6 summarizes some of the measures used to minimize postoperative ileus. Such data have generated significant interest in Early Recovery After Surgery (ERAS) pathways, which are a collection of steps taken to expedite postoperative recovery in general. ERAS protocols typically involve 15 to 20 steps that involve the pre-, intra- and postoperative phases of care and form a multimodal pathway. Although the contribution of each element to the overall outcome has not been well studied, the bundle of steps leads to reduced length of stay and surgical complications. In cases of GI surgery, many of these steps are targeted towards reducing postoperative ileus, which is often the barrier to early discharge.

Although prokinetic agents have been tried to promote return of GI motility, they are associated with efficacy-toxicity profiles that are too unfavorable to warrant routine use. Recently, administration of alvimopan, a novel, peripherally active mu-opioid receptor antagonist with limited oral absorption, has been shown to reduce duration of postoperative ileus, hospital stay, and rate of readmissions in several prospective, randomized, placebo-controlled trials and the subsequent meta-analysis. Any cost savings associated with the use of this drug outside of a clinical trial has, however, been debated.

The therapy of patients with chronic intestinal pseudo-obstruction focuses on palliation of symptoms as well as fluid, electrolyte, and nutritional management. Surgery should be avoided if possible. No standard therapies are curative or delay the natural history of any of the specific disorders causing intestinal pseudo-obstruction. Prokinetic agents, such as metoclopramide and erythromycin, are associated with poor efficacy. Cisapride has been associated with palliation of symptoms; however, because of cardiac toxicity and reported deaths, this agent is restricted to compassionate use in the United States. Patients with refractory disease may require strict limitation of oral intake and long-term TPN administration. Despite these measures, some patients will continue to have severe abdominal pain or such copious intestinal secretions that vomiting and fluid and electrolyte losses remain substantial. These patients may require a decompressive gastrostomy or an extended small bowel resection to remove abnormal intestine. Small-intestinal transplantation has been applied in these patients with increasing frequency; the ultimate role of this modality remains to be defined.

### CROHN’S DISEASE

Crohn’s disease is a chronic, idiopathic transmural inflammatory disease with skip lesions that may affect any part of the alimentary tract, although there is propensity to affect the distal small bowel. Nearly 80% of patients with Crohn’s disease have small bowel involvement, with 30% having terminal ileitis exclusively. Recent studies suggest a prevalence of about 241 cases per 100,000 in the United States. The rates of Crohn’s and ulcerative colitis have been increasing globally over the past several decades with substantial regional variations in incidence. The highest incidences are reported in western nations and those in northern latitudes, with Canada having the highest reported rates. In countries such as China, the prevalence of Crohn’s disease is substantially below that seen in the West, but rates have been rapidly increasing recently. The incidence of Crohn’s disease varies among ethnic groups within the same geographic region. For example, members of Eastern European Ashkenazi Jewish population are at a two- to fourfold higher risk of developing Crohn’s disease than members of other populations living in the same location.

Most studies suggest that Crohn’s disease is slightly more prevalent in females than in males. The mean age at which patients are diagnosed with Crohn’s disease falls in the third decade of life years, with a second smaller peak in the sixth decade of life, giving it a bimodal distribution. The age at diagnosis can, however, range from early childhood through the entire lifespan.

Both genetic and environmental factors appear to influence the risk for developing Crohn’s disease. The relative risk among first-degree relatives of patients with Crohn’s disease is 14 to 15 times higher than that of the general population, with about 20% of patients reporting a family history. The concordance rate among monozygotic twins is as high as 67%; however, Crohn’s disease is not associated with simple Mendelian inheritance patterns. Although there is a tendency within families for either ulcerative colitis or Crohn’s disease to be present exclusively, mixed kindreds also occur, suggesting the presence of some shared genetic traits as a basis for both diseases.

Higher socioeconomic status is associated with an increased risk of Crohn’s disease. Most studies have found breastfeeding to be protective against the development of Crohn’s disease. Crohn’s disease is more prevalent among smokers. Furthermore, smoking is associated with the increased risk for both the need for surgery and the risk of relapse after surgery for Crohn’s disease.

#### Pathophysiology

Crohn’s disease is characterized by sustained inflammation. Whether this inflammation represents an appropriate response to a yet unrecognized pathogen or an inappropriate response to a normally innocuous stimulus is unknown. Various hypotheses on the roles of environmental and genetic factors in the pathogenesis of Crohn’s disease have been proposed. Many infectious agents have been suggested to be the causative organism of Crohn’s disease; however, there has been no conclusive evidence to confirm any. Studies using animal models suggest that in a genetically susceptible host, a nonpathogenic gut microbiome is sufficient to induce a chronic inflammatory response resembling
that associated with Crohn’s disease. In these models, the sustained intestinal inflammation is the result of either abnormal epithelial barrier function or immune dysregulation. A full discussion of the role of gut immune system and microbiome in the development of Crohn’s disease is beyond the scope of this work, but it is an area of great interest and under investigation. In general, poor barrier function is hypothesized to permit inappropriate exposure of lamina propria lymphocytes to antigenic stimuli derived from the intestinal lumen. In addition, a variety of defects in immune regulatory mechanisms, e.g., oversensitivity of mucosal T cells to enteric flora-derived antigens, can lead to defective immune tolerance and sustained inflammation.

Specific genetic defects associated with Crohn’s disease in human patients are beginning to be defined. For example, the presence of a locus on chromosome 16 (the so-called IBD1 locus) has been linked to Crohn’s disease. The IBD1 locus has been identified as the NOD2 gene. Persons with allelic variants on both chromosomes have a 40-fold relative risk of Crohn’s disease compared to those without variant NOD2 genes. The relevance of this gene to the pathogenesis of Crohn’s disease is biologically plausible, as the protein product of the NOD2 gene mediates the innate immune response to microbial pathogens. Other putative IBD loci have been identified on other chromosomes (IBD2 on chromosome 12q, and IBD3 on chromosome 6), and are under investigation.

Although appendectomy has been shown to lower the risk of subsequent development of ulcerative colitis, it was suspected that the surgery may increase the risk of developing Crohn’s disease. A meta-analysis has, however, suggested that the observed increased risk of Crohn’s disease in the first few years after an appendectomy may in fact reflect diagnostic difficulty in a group of patients with incipient Crohn’s.

Although the pathological hallmark of Crohn’s disease is focal, transmural inflammation of the intestine, a spectrum of pathological lesions can be present. The earliest lesion characteristic of Crohn’s disease is the aphthous ulcer. These superficial ulcers are up to 3 mm in diameter and are surrounded by a halo of erythema. In the small intestine, aphthous ulcers typically arise over lymphoid aggregates. Granulomas are highly characteristic of Crohn’s disease and are reported to be present in up to 70% of intestinal specimens obtained during surgical resection. These granulomas are noncaseating and can be found in both areas of active disease and apparently normal intestine, in any layer of the bowel wall, and in mesenteric lymph nodes.

As disease progresses, aphthae coalesce into larger, stellate-shaped ulcers. Linear or serpiginous ulcers may form when multiple ulcers fuse in a direction parallel to the longitudinal axis of the intestine. With transverse coalescence of ulcers, a cobblestoned appearance of the mucosa may arise.

With advanced disease, inflammation can be transmural. Serosal involvement results in adhesion of the inflamed bowel to other loops of bowel or other adjacent organs. Transmural inflammation can also result in fibrosis with stricture formation, intra-abdominal abscesses, fistulas, and, rarely, free perforation. Inflammation in Crohn’s disease can affect discontinuous portions of intestine, so-called skip lesions that are separated by intervening normal-appearing intestine.

A feature of Crohn’s disease that is grossly evident and helpful in identifying affected segments of intestine during surgery is the presence of fat wrapping, which represents encroachment of mesenteric fat onto the serosal surface of the bowel (Fig. 28-17). This finding is virtually pathognomonic of Crohn’s disease. The presence of fat wrapping correlates well with the presence of underlying acute and chronic inflammation.

Features that allow for differentiation between Crohn’s disease of the colon and ulcerative colitis include the layers of the bowel wall affected (inflammation in ulcerative colitis is limited to the mucosa and submucosa but may involve the full-thickness of the bowel wall in Crohn’s disease) and the longitudinal extent of inflammation (inflammation is continuous and characteristically affects the rectum in ulcerative colitis but may be discontinuous and spare the rectum in Crohn’s disease). In the absence of full expression of features of advanced disease, Crohn’s colitis can sometimes be difficult to distinguish from ulcerative colitis. It is also important to remember that although ulcerative colitis is a disease of the colon, it can be associated with inflammatory changes in the distal ileum (backwash ileitis).

Clinical Presentation
The most common symptoms of Crohn’s disease are abdominal pain, diarrhea, and weight loss. However, the clinical features are highly variable among individual patients and depend on which segment(s) of the gastrointestinal tract is (are) predominately affected, the intensity of inflammation, and the presence or absence of specific complications. In fact, some patients with Crohn’s disease may have been initially misdiagnosed as having irritable bowel syndrome or celiac disease.

Patients can be classified by their predominant clinical manifestation as having primarily (a) fibrostenotic disease, (b) fistulizing disease, and (c) aggressive inflammatory disease. There is substantial overlap among these disease patterns in individual patients, however. The onset of symptoms is insidious, and once present, their severity follows a waxing and waning course. Constitutional symptoms, particularly weight loss and fever, or growth retardation in children, may also be prominent and are occasionally the sole presenting features of Crohn’s disease.

The disease affects the small bowel in 80% of cases and colon alone in 20%. In those with small bowel disease, the majority have ileocecal disease. Isolated perineal and anorectal disease occurs in 5% to 10% of affected patients. Uncommon sites of involvement include the esophagus, stomach, and duodenum.

An estimated one-fourth of all patients with Crohn’s disease will have an extraintestinal manifestation of their disease. One
fourth of those affected will have more than one manifestation. Many of these complications can be seen with both Crohn’s disease and ulcerative colitis, although they are more prevalent among patients with Crohn’s disease. The most common extraintestinal manifestations are listed in Table 28-7. The clinical severity of some of these manifestations, such as erythema nodosum and peripheral arthritis, is correlated with the severity of intestinal inflammation. The severity of other manifestations, such as pyoderma gangrenosum and ankylosing spondylitis, bears no apparent relationship to the severity of intestinal inflammation.

### Diagnosis

The diagnosis is usually established with endoscopic findings in a patient with a compatible clinical history. The diagnosis should be considered in those presenting with acute or chronic abdominal pain, especially when localized to the right lower quadrant, chronic diarrhea, evidence of intestinal inflammation on radiography or endoscopy, the discovery of a bowel stricture or fistula arising from the bowel, and evidence of inflammation or granulomas on intestinal histology. Disorders associated with clinical presentations that resemble those of Crohn’s disease include ulcerative colitis, functional bowel disorders such as irritable bowel syndrome, mesenteric ischemia, collagen vascular diseases, carcinoma and lymphoma, diverticular disease, and infectious enteritides. Infectious enteritides are most frequently diagnosed in immunocompromised patients, but they can also occur in patients with normal immune function. Acute ileitis caused by *Campylobacter* and *Yersinia* species can be difficult to distinguish from that caused by an acute presentation of Crohn’s disease. Typhoid enteritis caused by *Salmonella typhosa* can lead to overt intestinal bleeding and perforation, most often affecting the terminal ileum. The distal ileum and cecum are the most common sites of intestinal involvement by infection due to *Mycobacterium tuberculosis*. This condition can result in intestinal inflammation, strictures, and fistula formation, like those seen in Crohn’s disease. Cytomegalovirus (CMV) can cause intestinal ulcers, bleeding, and perforation.

No single symptom, sign, or diagnostic test establishes the diagnosis of Crohn’s disease. Instead, the diagnosis is based on a complete assessment of the clinical presentation with confirmatory findings derived from radiographic, endoscopic, and in most cases, pathologic tests. Patients presenting with a history of Crohn’s disease should have their full blood count, electrolytes and renal function, liver function, iron, B12, ESR, and CRP levels checked. The results may be abnormal, showing anemia, but these results are nondiagnostic. Colonoscopy with intubation of terminal ileum is the main diagnostic tool and can reveal focal ulcerations adjacent to areas of normal appearing mucosa along with polyloid mucosal changes that give a “cobblestone appearance.” Skip areas of involvement are typical with segments of normal-appearing bowel interrupted by large areas of obvious disease; this pattern is different from the continuous involvement in ulcerative colitis. Pseudopolyps, as seen in ulcerative colitis, are also often present. Barium small bowel follow-through, CT enterography, or MR enterography may be used as contrast examinations of the small bowel to reveal strictures or networks of ulcers and fissures. CT scanning may reveal intra-abdominal abscesses and is useful in acute presentations to rule out the presence of other intra-abdominal disorders. Esophagogastroduodenoscopy (EGD) is done for disease of the proximal alimentary tract. Because Crohn’s disease often affects the small bowel, which is difficult to image, capsule endoscopy has been increasing used to make this diagnosis (Fig. 28-18).40

**Table 28-7** Extraintestinal manifestations of Crohn’s disease

<table>
<thead>
<tr>
<th>Dermatologic</th>
<th>Erythema nodosum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pyoderma gangrenosum</td>
</tr>
<tr>
<td>Rheumatologic</td>
<td>Peripheral arthritis</td>
</tr>
<tr>
<td></td>
<td>Ankylosing spondylitis</td>
</tr>
<tr>
<td></td>
<td>Sacroiliitis</td>
</tr>
<tr>
<td>Ocular</td>
<td>Conjunctivitis</td>
</tr>
<tr>
<td></td>
<td>Uveitis/iritis</td>
</tr>
<tr>
<td></td>
<td>Episcleritis</td>
</tr>
<tr>
<td>Hepatobiliary</td>
<td>Hepatic steatosis</td>
</tr>
<tr>
<td></td>
<td>Cholelithiasis</td>
</tr>
<tr>
<td></td>
<td>Primary sclerosing cholangitis</td>
</tr>
<tr>
<td></td>
<td>Pericholangitis</td>
</tr>
<tr>
<td>Urologic</td>
<td>Nephrolithiasis</td>
</tr>
<tr>
<td></td>
<td>Ureteral obstruction</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>Thromboembolic disease</td>
</tr>
<tr>
<td></td>
<td>Vasculitis</td>
</tr>
<tr>
<td></td>
<td>Osteoporosis</td>
</tr>
<tr>
<td></td>
<td>Endocarditis, myocarditis, pleuropericarditis</td>
</tr>
<tr>
<td></td>
<td>Interstitial lung disease</td>
</tr>
<tr>
<td></td>
<td>Amyloidosis</td>
</tr>
<tr>
<td></td>
<td>Pancreatitis</td>
</tr>
</tbody>
</table>

![Figure 28-18](image-url) Crohn’s disease. This image was captured by a wireless capsule endoscope as it was traveling through the small intestine. It demonstrates a superficial ulceration in the small bowel consistent with Crohn’s disease. *(Used with permission from Anne T. Wolf, M.D., Department of Medicine, Brigham and Women’s Hospital, Boston, MA.)*
Several antibodies have also been identified in patients with inflammatory bowel disease, which may have diagnostic value. The most commonly tested antibodies are antineutrophil cytoplasmic antibody (pANCA) and anti-saccharomyces cerevisiae antibody (ASCA). ASCA+/pANCA−, is associated with a diagnosis of Crohn’s disease, while ASCA−/pANCA+, correlates with ulcerative colitis. Although these antibody tests have high specificity, their use has been hampered by low test sensitivities. There is ongoing interest in developing other antibody tests to diagnose inflammatory bowel disease and help differentiate Crohn’s disease from ulcerative colitis. There have been attempts to develop stool tests to diagnose inflammatory bowel disease, and although fecal calprotectin or lactoferrin can identify patients with intestinal inflammation, they are not routinely done in clinical practice.

Because of the insidious, and often nonspecific, presentation of Crohn’s disease, a diagnosis of Crohn’s is typically made only after symptoms have been present for several years. However, in acute presentations, the diagnosis is sometimes made intraoperatively or during surgical evaluation. The initial manifestation of Crohn’s disease can consist of right lower quadrant abdominal mimic of the presentation of acute appendicitis. In patients with this presentation, Crohn’s disease can be discovered for the first time during laparotomy or laparoscopy performed for presumed appendicitis. In some patients, the initial manifestation of Crohn’s disease is an acute abdomen related to small bowel obstruction, intra-abdominal abscess, or free intestinal perforation. In other patients, perianal abscesses and fistulas requiring surgical therapy may be the first manifestation of Crohn’s disease.

**Therapy**

Because no curative therapies are available for Crohn’s disease, the goal of treatment is to palliate symptoms rather than to achieve cure. Medical therapy is used to induce and maintain disease remission. Surgery is reserved for specific indications described later in this chapter. In addition, nutritional support in the form of aggressive enteral regimens or, if necessary, parenteral nutrition, is used to manage the malnutrition that is common in patients with Crohn’s disease.

**Medical Therapy.** Pharmacologic agents used to treat Crohn’s disease include antibiotics, aminosalicylates, corticosteroids, immunomodulators, and biologic therapies. Antibiotics have an adjunctive role in the treatment of infectious complications associated with Crohn’s disease. They are also used to treat patients with perianal disease, enterocutaneous fistulas, and active colonic disease.

Crohn’s disease activity is assessed using the Crohn’s disease Activity Index or Harvey-Bradshaw Index, and depending on the scores, it can be categorized as asymptomatic, mild, moderate, or severe disease to guide therapy. While patients with mild and moderate disease can be managed on an outpatient basis, those with severe or fulminant disease often require hospitalization for treatment, bowel rest, and possible nutritional support. There are two general approaches to treating Crohn’s disease: top-down (which starts with the most potent agents to achieve remission with a subsequent decrease in medication) or step-up (starts with less potent and often safer drugs, and if symptoms fail to improve advances to the next group of medications).

The use of oral 5-aminosalicylic acid (5-ASA) drugs (e.g., mesalamine) is somewhat controversial with mixed results from several randomized studies and meta-analyses. Aminosalicylates are associated with minimal toxicity and are available in a variety of formulations that allow for their delivery to specific regions of the alimentary tract. Thus, many continue to recommend use of mesalamine as an initial step in management of mild symptoms in patients with small bowel Crohn’s disease.

Orally administered glucocorticoids are used to treat patients with mild disease that does not respond to aminosalicylates, or as initial treatment of patients with moderate disease. Patients with severe active disease usually require intravenous administration of glucocorticoids. Although glucocorticoids are effective in inducing remission, they are ineffective in preventing relapse, and their adverse side-effect profile makes long-term use hazardous. Therefore, they should be tapered once remission is achieved. Some patients are unable to undergo glucocorticoid tapering without suffering recurrence of symptoms. Such patients are said to have steroid dependence. These patients, along with those who do not respond to steroids at all (steroid resistant), should be considered for immune modulator therapies. Controlled ileal-released budesonide is an oral steroid with high first-pass hepatic metabolism and few systemic effects that can be tried in those with ileal and colonic Crohn’s disease.

For those with severe disease, the thiopurine antimetabolites azathioprine and its active metabolite, 6-mercaptopurine, have demonstrated efficacy in inducing remission, maintaining remission, and allowing for glucocorticoid tapering in glucocorticoid-dependent patients. A response to these medications is usually observed in 3 to 6 months, during which patients may need to continue with steroids. There is also some evidence that they decrease the risk of relapse after intestinal resection for Crohn’s disease. These agents are relatively safe but can induce bone marrow suppression and promote infectious complications. For patients who do not respond to the thiopurines, methotrexate is an alternative that is usually initially given intramuscularly before switching to oral form after achieving symptomatic control. There is little role for cyclosporine in Crohn’s disease; its efficacy/toxicity profile in this disease is poor.

The successful introduction of infliximab (Remicade), an anti-TNFα antibody, heralded the era of biological therapies for inflammatory bowel disease. Infliximab is a chimeric monoclonal antitumor necrosis-factor alpha (TNFα) antibody that has been shown to have efficacy in inducing remission and in promoting closure of enterocutaneous fistulae. There are two other anti-TNFα antibodies, with no randomized studies comparing efficacy of the drugs head to head. In general, it is thought that there is no significant difference in efficacy between them. While infliximab is a mouse-human chimeric antibody, adalimumab (Humira) is a fully human antibody. Certolizumab pegol (Cimzia) is a PEGylated Fab fragment of a humanized TNF inhibitor monoclonal antibody. These agents are generally used for patients who are resistant to standard therapy, to help taper steroid dosage. They are generally well tolerated, but they should not be used in patients with ongoing septic processes, such as undrained intra-abdominal abscesses. Antibodies against other targets in this inflammatory pathway have also been developed, including vedolizumab (Entyvio), a humanized anti-α4β7 integrin monoclonal antibody, with more specific anti-inflammatory effect in the intestine.

For patients with perianal disease, antibiotic therapy with metronidazole or ciprofloxacin is the primary step. Two to 4 weeks of therapy is needed before improvements are seen, and often long-term therapy is required to prevent relapse. In cases of relapse, azathioprine can be considered. In patients with fistulas, infliximab and azathioprine are drugs of choice.
**Surgical Therapy.** With introduction of new treatments, the need for surgery for Crohn’s disease has decreased steadily over the past few decades. Recent meta-analysis estimated the risk of surgery to be 16.3%, 33.3%, and 46.6% at 1, 5, and 10 years respectively. Surgery is generally reserved for patients whose disease is unresponsive to aggressive medical therapy or who develop complications of their disease (Table 28-8). Failure of medical management may be the indication for surgery if symptoms persist despite aggressive medical therapy for several months or if symptoms recur whenever aggressive therapy is tapered. Surgery should be considered if medication-induced complications arise, specifically corticosteroid-related complications, such as cushingoid features, cataracts, glaucoma, systemic hypertension, compression fractures, or aseptic necrosis of the femoral head. Growth retardation constitutes an indication for surgery in 30% of children with Crohn’s disease.

One of the most common indications for surgical intervention is intestinal obstruction. Abscesses and fistulas are frequently encountered during operations performed for intestinal obstruction in these patients, but they are rarely the only indication for surgery. Most abscesses are amenable to percutaneous drainage, and fistulas, unless associated with symptoms or metabolic derangements, do not require surgical intervention. Less common complications that require surgical intervention are acute gastrointestinal hemorrhage, perforations, and development of cancer.

Although surgery for Crohn’s disease is usually planned, an uncommon, but not rare, scenario is the intraoperative discovery of inflammation limited to the terminal ileum during operations performed for presumed appendicitis. This scenario can result from an acute presentation of Crohn’s disease or from acute ileitis caused by bacteria such as *Yersinia* or *Campylobacter*. Both conditions should be treated medically; ileal resection is not generally indicated. However, the appendix, even if normal appearing, should be removed (unless the cecum is inflamed, increasing the potential morbidity of this procedure) to eliminate appendicitis from the differential diagnosis of abdominal pain in these patients, particularly those with Crohn’s disease who may be destined to have recurring symptoms.

When the diagnosis of Crohn’s disease is known and surgery is planned, a thorough examination of the entire intestine should be performed. The presence of active disease is suggested by thickening of the bowel wall, narrowing of the lumen, serosal inflammation and coverage by creeping fat, and thickening of the mesentery. Skip lesions are present in approximately 20% of cases and should be sought. The length of uninvolved small intestine should be noted.

Segmental intestinal resection of grossly evident disease followed by primary anastomosis is the usual procedure of choice. Microscopic evidence of Crohn’s disease at the resection margins does not compromise a safe anastomosis, and frozen section analysis of resection margins is unnecessary. In a randomized prospective trial, the effects of achieving 2-cm resection margins beyond grossly evident disease were compared with achieving 12-cm resection margins. There were no evident differences with respect to clinical recurrence rates or anastomotic recurrences. Recurrence rates were similar whether margins were histologically free of or involved with Crohn’s disease. An area of controversy in surgical management of Crohn’s disease has been the ideal anastomotic technique for the bowel after intestinal resection. This issue was addressed in a randomized study of 139 patients undergoing an ileocolic resection for Crohn’s disease, with a mean follow-up of 11.9 months. There were no differences in endoscopic or symptomatic disease recurrence between the groups reconstructed using end-to-end sutured (2-0 PDS) anastomosis versus those with side-to-side staples anastomosis.

An alternative to segmental resection for obstructing lesions is stricturoplasty (Fig. 28-19). This technique allows for preservation of intestinal surface area and is especially well suited to patients with extensive disease and fibrotic strictures who may have undergone previous resection and are at risk for developing short bowel syndrome. In this technique, the bowel is opened longitudinally to expose the lumen. Any

<table>
<thead>
<tr>
<th>Table 28-8</th>
<th>Indications for surgical intervention in Crohn’s disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute onset of severe disease:</td>
<td>Crohn’s colitis +/- toxic megacolon (rare)</td>
</tr>
<tr>
<td>Failure of medical therapy:</td>
<td>Persistent symptoms despite long-term steroid use</td>
</tr>
<tr>
<td></td>
<td>Recurrence of symptoms when high-dose steroids are tapered</td>
</tr>
<tr>
<td></td>
<td>Drug-induced complications (Cushing’s disease, hypertension)</td>
</tr>
<tr>
<td>Development of disease complications:</td>
<td>Obstruction</td>
</tr>
<tr>
<td></td>
<td>Perforation</td>
</tr>
<tr>
<td></td>
<td>Complicated fistulas</td>
</tr>
<tr>
<td></td>
<td>Hemorrhage</td>
</tr>
<tr>
<td></td>
<td>Malignancy risk</td>
</tr>
</tbody>
</table>

Figure 28-19. Stricturoplasty. The wall of the strictured bowel is incised longitudinally. Reconstruction is performed by closing the defect transversely in a manner similar to the Heineke-Mikulicz pyloroplasty for short strictures (A), or the Finney pyloroplasty for longer strictures (B).
intraluminal ulcerations should be biopsied to rule out the presence of neoplasia. Depending on the length of the stricture, the reconstruction can be fashioned in a manner similar to the Heinecke-Mickulicz pyloroplasty (for strictures less than 12 cm in length) or the Finney pyloroplasty (for longer strictures as much as 25 cm in length). For longer strictures, variations on the standard stricturoplasty, namely the side-to-side isoperistaltic enterenterostomy, have been advocated and used for strictures with mean lengths of 50 cm.44 Stricturoplasty sites should be marked with metallic clips to facilitate their identification on radiographs and during subsequent operations. Stricturoplasty is associated with recurrence rates that are different from those associated with segmental resection. Because the affected bowel is left in situ rather than resected, there is the potential for cancer developing at the stricturoplasty site. However, as data on this complication are limited to anecdotes, this risk remains a theoretical one. Stricturoplasty is contraindicated in patients with intra-abdominal abscesses or intestinal fistulas. The presence of a solitary stricture relatively close to a segment for which resection is planned is a relative contraindication. In general, stricturoplasty is performed in cases where single or multiple strictures are identified in diffusely involved segments of bowel, or where previous resections have been performed and maintenance of intestinal length is of great importance.

Intestinal bypass procedures are sometimes required in the presence of intramesenteric abscesses or if the diseased bowel is coalesced in the form of a dense inflammatory mass, making its mobilization unsafe. Bypass procedures (gastrojejunostomy) are also used in the presence of duodenal strictures, for which stricturoplasty and segmental resection can be technically difficult.

Since the 1990s, laparoscopic surgical techniques have been applied to patients with Crohn’s disease. The inflammatory changes associated with Crohn’s disease such as thickened and for eshorted mesentery, obliterated tissue planes, and friable tissues with engorged vasculature can make laparoscopic approach challenging. Randomized studies and a meta-analysis have confirmed that laparoscopic surgery for Crohn’s disease is associated with less postoperative pain, shorter duration of ileus, and a shorter hospital stay. The rates of disease recurrence were similar between the two groups.45

Outcomes
Overall complication rates following surgery for Crohn’s disease range from 15% to 30%. Wound infections, postoperative intra-abdominal abscesses, and anastomotic leaks account for most of these complications.

Surgery is not a curative intervention in Crohn’s disease, and many patients develop recurrence. If recurrence is defined endoscopically, 70% recur within 1 year of a bowel resection and 85% by 3 years.46 Clinical recurrence, defined as the return of symptoms confirmed as being due to Crohn’s disease, affects 60% of patients by 5 years and 94% by 15 years after intestinal resection. Reoperation becomes necessary in approximately one-third of patients by 5 years after the initial operation, with a median time to reoperation of 7 to 10 years.47 Of patient-modifying factors, smoking is a strong risk factor for disease recurrence.

INTESTINAL FISTULAS
A fistula is defined as an abnormal communication between two epithelialized surfaces. The communication occurs between two parts of the gastrointestinal tract or adjacent organs in an internal fistula (e.g., enterocolonic fistula or colovesicular fistula). An external fistula (e.g., enterocutaneous fistula or rectovaginal fistula) involves the skin or another external surface epithelium. Enterocutaneous fistulas that drain less than 200 mL of fluid per day are known as low-output fistulas, whereas those that drain more than 500 mL of fluid per day are known as high-output fistulas.

Over 80% of enterocutaneous fistulas represent iatrogenic complications that occur as the result of enterotomies or intestinal anastomotic dehiscences. Fistulas that arise spontaneously without antecedent iatrogenic injury are usually manifestations of progression of underlying Crohn’s disease or cancer.

Pathophysiology
The manifestations of fistulas depend on which structures are involved. Low-resistance enterenteric fistulas, which allow luminal contents to bypass a significant proportion of the small intestine, may result in clinically-significant malabsorption. Enterovesicular fistulas often cause recurrent urinary tract infections. The drainage emanating from enterocutaneous fistulas are irritating to the skin and cause excoriation. The loss of enteric luminal contents, particularly from high-output fistulas originating from the proximal small intestine, results in dehydration, electrolyte abnormalities, and malnutrition.

Fistulas have the potential to close spontaneously. Factors inhibiting spontaneous closure, however, include malnutrition, sepsis, inflammatory bowel disease, cancer, radiation, obstruction of the intestine distal to the origin of the fistula, foreign bodies, high output, short fistulous tract (<2 cm) and epithelialization of the fistula tract (Table 28-9).

Clinical Presentation
Iatrogenic enterocutaneous fistulas usually become clinically evident between the fifth and tenth postoperative days. Fever, leukocytosis, prolonged ileus, abdominal tenderness, and wound infection are the initial signs. The diagnosis becomes obvious when drainage of enteric material through the abdominal wound or through existing drains occurs. These fistulas are often associated with intra-abdominal abscesses.

Diagnosis
CT scanning following the administration of enteral contrast is the most useful initial test. Leakage of contrast material from the intestinal lumen can be observed. Intra-abdominal abscesses

| Table 28-9 |
| Factors negatively impacting enteric fistula closure |
| Patient factors |
| Poor nutrition |
| Medications such as steroids |
| Etiological factors |
| Malignant fistula |
| Fistula related to Crohn’s disease |
| Fistula in radiated fields |
| Fistula site |
| Gastric |
| Duodenal |
| Local Factors |
| Persistence of local inflammation and sepsis |
| Presence of a foreign body (e.g., meshes or sutures) |
| Epithelialization of fistula tract |
| Fistula tract <2 cm |
| Distal obstruction to the fistula site |
should be sought and drained percutaneously. If the anatomy of the fistula is not clear on CT scanning, a small bowel series or enteroclysis examination can be obtained to demonstrate the fistula’s site of origin in the bowel. This study is also useful to rule out the presence of intestinal obstruction distal to the site of origin. Occasionally, contrast administered into the intestine does not demonstrate the fistula tract. A fistulogram, in which contrast is injected under pressure through a catheter placed percutaneously into the fistula tract, may offer greater sensitivity in localizing the fistula origin.

**Therapy**

The treatment of enterocutaneous fistulas should proceed through an orderly sequence of steps:

1. **Stabilization.** Fluid and electrolyte resuscitation is begun. Nutrition is provided, usually through the parenteral route initially. Sepsis is controlled with antibiotics and drainage of abscesses. The skin is protected from the fistula effluent with ostomy appliances or fistula drains.
2. **Investigation.** The anatomy of the fistula is defined using the aforementioned studies.
3. **Decision.** The available treatment options are considered, and a time line for conservative measures is determined.
4. **Definitive Management.** This entails the surgical procedure and requires appropriate preoperative planning and surgical experience.
5. **Rehabilitation.**

The overall objectives are to increase the probability of spontaneous closure. Nutrition and time are the key components of this approach. Most patients will require TPN; however, a trial of oral or enteral nutrition should be attempted in patients with low-output fistulas originating from the distal intestine. The somatostatin analogue octreotide is a useful adjunct, particularly in patients with high-output fistulas. A meta-analysis of several randomized studies confirmed that somatostatin treatment reduced length of hospital stay and time to closure of fistulas; however, its administration did not lead to a significant difference in fistula closure rates. Use of negative pressure wound therapy has increased in management of enterocutaneous fistulas. The system can allow better management of the fistula output. In a study of 91 patients with enterocutaneous fistulas, 40% of fistulas reached minimal output within a week, and with an average follow-up of 90 days, spontaneous closure rate was 46%.

**Timing of Surgical Intervention.** Most surgeons would pursue 2 to 3 months of conservative therapy before considering surgical intervention. This approach is based on evidence that 90% of fistulas that are going to close do so within 5 weeks and that surgical intervention after this period is associated with better outcomes and lower morbidity.

If the fistula fails to resolve during this period, surgery may be required, during which the fistula tract, together with the segment of intestine from which it originates, should be resected. Simple closure of the opening in the intestine from which the fistula originates is associated with high recurrence rates. Patients with intestinal fistulas typically have extensive and dense intra-abdominal adhesions. Thus, operations performed for nonhealing fistulas can present formidable challenges. Successful applications of alternative therapies to close intestinal fistulas such as the use of biologic sealants have been reported. The indications for their use remain to be defined.

**Outcomes**

Over 50% of intestinal fistulas close spontaneously. A useful mnemonic designates factors that inhibit spontaneous closure of intestinal fistulas: “FRIEND” (Foreign body within the fistula tract, Radiation enteritis, Infection/Inflammation at the fistula origin, Epithelialization of the fistula tract, Neoplasm at the fistula origin, Distal obstruction of the intestine).

In a 23-year old retrospective review of 153 cases of enterocutaneous fistulas that were treated surgically, most fistulas were found to originate from the small bowel and be iatrogenic in nature, with patients having undergone five or more previous abdominal surgeries. Operative repair was associated with a 30-day mortality of approximately 4% and a 1-year mortality of 15%. Morbidity was over 80%. First attempt at surgical repair was successful in 70% of cases, with an overall closure rate of 84% and some patients requiring up to 3 attempts at surgical repair. The authors identified closure of the abdominal fascia as an important factor in reducing rates of refistulization and postoperative mortality. In another similar study, fistula recurrence rates of 30% were documented and were independently associated with high output fistulas and the type of surgical treatment: operations not involving resection of the fistula had a much higher rate of recurrence.

**SMALL BOWEL NEOPLASMS**

Adenomas are the most common benign neoplasm of the small intestine. Other benign tumors include fibromas, lipomas, hemangiomas, lymphangiomas, and neurofibromas. The prevalence of small bowel tumors identified at autopsy is 0.2% to 0.3%, which is significantly higher than the rate of operation for small bowel tumors. This suggests that majority of small bowel tumors are asymptomatic. These lesions are most frequently encountered in the duodenum as incidental findings during esophagogastroduodenoscopic (EGD) examinations (Fig. 28-20). The reported prevalence of
duodenal polyps, as detected during EGD performed for other reasons, range from 0.3% to 4.6%.

Benign neoplasms account for 30% to 50% of small bowel tumors and include adenomas, lipomas, hematomas, and hemangiomas. Primary small bowel cancers are rare but have been increasing in incidence, with an estimated incidence of 10,190 cases in 2017 in the United States. Among small bowel cancers, adenocarcinomas comprise 35% to 50% of all cases, carcinoid tumors comprise 20% to 40%, and lymphomas comprise approximately 10% to 15%. In a retrospective review of a large U.S. database (SEER) between 1992 and 2006, of a total number of 10,945 small intestine cancers, 4315 were neuroendocrine in origin, 3412 were carcinomas, 2023 were lymphomas, and 1084 were sarcomas. Gastrointestinal stromal tumors (GISTs) are the most common mesenchymal tumors arising in the small intestine and comprise the vast majority of tumors that were formerly classified as leiomyomas, leiomyosarcomas, and smooth muscle tumors of the intestine. The small intestine is frequently affected by metastases from or local invasion by cancers originating at other sites. Melanoma, in particular, is associated with a propensity for metastasis to the small intestine.

Most patients with small-intestinal cancers are in their fifth or sixth decade of life. Reported risk factors for developing small-intestinal cancers include consumption of red meat, ingestion of smoked or cured foods, Crohn’s disease, celiac sprue, hereditary nonpolyposis colorectal cancer (HNPCC), familial adenomatous polyposis (FAP), and Peutz-Jeghers syndrome.

Pathophysiology
The small intestine contains over 90% of the mucosal surface area of the gastrointestinal tract but only 1.1% to 2.4% of all gastrointestinal malignancies. Proposed explanations for the low frequency of small-intestinal cancers include consumption of red meat, ingestion of smoked or cured foods, Crohn’s disease, celiac sprue, hereditary nonpolyposis colorectal cancer (HNPCC), familial adenomatous polyposis (FAP), and Peutz-Jeghers syndrome.

Table 28-10

<table>
<thead>
<tr>
<th>TUMOR TYPE</th>
<th>CELL OF ORIGIN</th>
<th>FREQUENCY(^a)</th>
<th>PREDOMINANT SITE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenocarcinoma</td>
<td>Epithelial cell</td>
<td>35–50%</td>
<td>Duodenum</td>
</tr>
<tr>
<td>Carcinoid</td>
<td>Enterochromaffin cell</td>
<td>20–40%</td>
<td>Ileum</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>Lymphocyte</td>
<td>10–15%</td>
<td>Ileum</td>
</tr>
<tr>
<td>GIST</td>
<td>Interstitial cell of Cajal</td>
<td>10–15%</td>
<td>–</td>
</tr>
</tbody>
</table>

\(^a\)Frequencies given as percentages of small intestinal malignancies comprised by each of the tumor types. Gastrointestinal stromal tumors (GISTs) display no regional variation in prevalence within the small intestine.

Recent advances have begun to clarify the molecular pathogenesis of small-intestinal adenocarcinomas and GISTs; there has been less progress with respect to the pathogenesis of the other small-intestinal malignancies (Table 28-10). Small-intestinal adenocarcinomas are believed to arise from preexisting adenomas through a sequential accumulation of genetic abnormalities in a model like that described for the pathogenesis of colorectal cancer. Adenomas are histologically classified as tubular, villous, and tubulovillous. Tubular adenomas have the least aggressive features. Villous adenomas have the most aggressive features and tend to be large, sessile, and located in the second portion of the duodenum. Malignant degeneration has been reported to be present in up to 45% of villous adenomas by the time of diagnosis. Patients with FAP have a nearly 100% cumulative lifetime risk of developing duodenal adenomas that have the potential to undergo malignant transformation. The risk of duodenal cancer in these patients is over 100-fold greater than in the general population. Indeed, duodenal cancer is the leading cause of cancer-related death among patients with FAP who have undergone colectomy. Patients with Peutz-Jeghers syndrome develop hamartomatous polyps; however, these polyps can contain adenomatous foci that can undergo malignant transformation (Fig. 28-21).

A defining feature of GISTs is their gain of function mutation of proto-oncogene KIT, a receptor tyrosine kinase.
Pathological KIT signal transduction is believed to be a central event in GIST pathogenesis. The majority of GISTs have activating mutations in the c-kit proto-oncogene, which cause KIT to become constitutively activated, presumably leading to persistence of cellular growth or survival signals. Because the interstitial cells of Cajal normally express KIT, these cells have been implicated as the cell of origin for GISTs. KIT expression is assessed by staining the tissues for CD117 antigen, which is part of the KIT receptor and is present in 95% of GISTs.

Clinical Presentation
Most small-intestinal neoplasms are asymptomatic until they become large. Partial small bowel obstruction, with associated symptoms of crampy abdominal pain and distention, nausea, and vomiting, is the most common mode of presentation. Obstruction can be the result of either luminal narrowing by the tumor itself or intussusception, with the tumor serving as the lead point. Hemorrhage, usually indolent, is the second most common mode of presentation.

Physical examination may be unrevealing, but it also may reveal a palpable abdominal mass in those with large tumors. Findings of intestinal obstruction may be present in some patients. A fecal occult blood test may be positive. Jaundice secondary to biliary obstruction or hepatic metastasis may be present. Cachexia, hepatomegaly, and ascites may be present with advanced disease.

Although the clinical presentation is usually not specific for tumor type, some general comments are appropriate. Adenocarcinomas, as well as adenomas (from which most are believed to arise), are most commonly found in the duodenum, except in patients with Crohn’s disease, in whom most are found in the ileum. Lesions in the periampullary location can cause obstructive jaundice or pancreatitis. Adenocarcinomas located in the duodenum tend to be diagnosed earlier in their progression than those located in the jejunum or ileum, which are rarely diagnosed prior to the onset of locally advanced or metastatic disease.

Carcinoid tumors of the small intestine are also usually diagnosed after the development of metastatic disease. These tumors are associated with a more aggressive behavior than the more common appendiceal carcinoid tumors. Approximately 25% to 50% of patients with carcinoid tumor-derived liver metastases will develop manifestations of the carcinoid syndrome. These manifestations include diarrhea, flushing, hypotension, tachycardia, and fibrosis of the endocardium and valves of the heart. Candidate tumor-derived mediators of the carcinoid syndrome such as serotonin, bradykinin, and substance P undergo nearly complete metabolism during the first passage through the liver. As a result, symptoms of carcinoid syndrome are rare in the absence of liver metastases.

Lymphoma may involve the small intestine primarily or as a manifestation of disseminated systemic disease. Primary small-intestinal lymphomas are most commonly located in the ileum, which contains the highest concentration of lymphoid tissue in the intestine. Although partial small bowel obstruction is the most common mode of presentation, 10% of patients with small-intestinal lymphoma present with bowel perforation.

The small intestine is the second most common site of GIST tumors after the stomach, containing 25% to 35% of GISTs. There appears to be no regional variation in the prevalence of GISTs within the small intestine. GISTs have a greater propensity to be associated with overt hemorrhage than the other small-intestinal malignancies (Fig. 28-22).

Metastatic tumors involving the small intestine can induce intestinal obstruction and bleeding.

Diagnosis
Because of the absent or nonspecific symptoms associated with most small-intestinal neoplasms, these lesions are rarely diagnosed preoperatively. Laboratory tests are nonspecific, except for elevated urinary 5-hydroxyindole acetic acid (5-HIAA) and serum chromogranin A (CgA) levels in patients with carcinoid syndrome, both of which have low sensitivity. Elevated carcinoembryonic antigen (CEA) levels are associated with small-intestinal adenocarcinomas, but only in the presence of liver metastases.

Contrast radiography of the small intestine may demonstrate benign and malignant lesions. Enteroclysis is reported to have a sensitivity of over 90% in the detection of small bowel tumors and is the test of choice, particularly for tumors located in the distal small bowel. Upper GI with small bowel follow-through examinations have reported sensitivities ranging from only 30% to 44% (Fig. 28-23). CT scanning can detect abnormalities in 70% to 80% of cases with small bowel tumor and assess for metastatic spread. Tumors associated with significant bleeding can be localized with angiography or radioisotope-tagged red blood cell (RBC) scans.

Tumors located in the duodenum can be visualized and biopsied on EGD. In addition, endoscopic ultrasonography (EUS) can offer additional information such as the layers of the intestinal wall involved by the lesion. Occasionally, the distal ileum can successfully be visualized during colonoscopy. Intraoperative enteroscopy can be used to directly visualize small-intestinal tumors beyond the reach of standard endoscopic techniques. For more distal small bowel lesions, capsule endoscopy and double-balloon

Figure 28-22. Jejunal gastrointestinal stromal tumor (GIST). This patient presented with overt obscure GI bleeding and was found to have a 7-cm jejunal GIST. The picture represents the laparoscopic view of the mass (black arrow), arising from the antimesenteric side of the small bowel (*). He underwent a successful laparoscopic resection.
Figure 28-23. Small bowel lesion identified during a small bowel follow-through (arrow). This patient had previously had a metastatic renal lesion to his duodenum requiring a Whipple procedure. During follow-up imaging 3 years later, he was found to have this new lesion in his jejunum. He underwent a laparoscopic small bowel resection. Pathology revealed a benign jejunal sessile polyp.

endoscopy have been used to evaluate small bowel. CT and MR enterography are also increasing utilized as noninvasive tests to look for small bowel masses. PET scans can also help assess metabolic activity of lesions and risk of malignancy.

**Therapy**

Benign neoplasms of the small intestine that are symptomatic should be surgically resected or removed endoscopically, if feasible. Tumors located in the duodenum, including asymptomatic lesions incidentally found during EGD, can pose the greatest therapeutic challenges. These lesions should be biopsied; symptomatic tumors and adenomas, because of their malignant potential, should be removed. In general, duodenal tumors less than 1 cm in diameter are amenable to endoscopic polypectomy. Lesions greater than 2 cm in diameter are technically difficult to remove endoscopically and may need to be removed surgically. Surgical options include transduodenal polypectomy and segmental duodenal resection. Tumors located in the second portion of the duodenum near the ampulla of Vater may require pancreaticoduodenectomy. EUS may offer utility for duodenal tumors ranging in size between 1 and 2 cm in diameter, with those limited to the mucosa being amenable to endoscopic polypectomy. Endoscopic resection of biopsy-proven benign duodenal periampullary adenomas leads to equivalent efficacy to surgery but with lower morbidity. Adenomas can recur; therefore, surveillance endoscopy is required after these procedures.55

Duodenal adenomas occurring in the setting of FAP require an especially aggressive approach to management. Patients with FAP should undergo screening EGD starting sometime during their second or third decade of life. Adenomas detected should be removed endoscopically, if possible, followed by surveillance endoscopy in 6 months and yearly thereafter, in the absence of recurrence. If surgery is required, pancreaticoduodenectomy is generally necessary because adenomas in patients with FAP tend to be multiple and sessile, with a predilection for the peri-ampullary region. Further, localized resections are complicated by high recurrence rates. Given the potential for recurrences in the duodenal remnant following pylorus-preserving pancreaticoduodenectomy, there is rationale for recommending the application of standard pancreaticoduodenectomy in these patients. However, recurrences have been reported even following this procedure; therefore, continuing surveillance is necessary. For most adenocarcinomas of the duodenum, except those in the third or fourth portion of the duodenum where a local resection could be considered, pancreaticoduodenectomy is required.

The surgical therapy of jejunal and ileal malignancies usually consists of wide-local resection of the intestine harboring the lesion. For adenocarcinomas, a wide excision of corresponding mesentery is done to achieve regional lymphadenectomy, as is done for adenocarcinomas of the colon. In the presence of locally-advanced or metastatic disease, palliative intestinal resection or bypass is performed. Chemotherapy has no proven efficacy in the adjuvant or palliative treatment of small-intestinal adenocarcinomas.

The goal of surgical therapy for carcinoids is resection of all visible disease. Localized small-intestinal carcinoid tumors should be treated with segmental intestinal resection and regional lymphadenectomy. Nodal metastases are unusual with tumors less than 1 cm in diameter, but they are present with 75% to 90% of tumors larger than 3 cm in diameter. In approximately 30% of cases, multiple small-intestinal carcinoid tumors are present (Fig. 28-24). Therefore, the entire small intestine should

Figure 28-24. Small bowel carcinoid tumor. This patient presented with history of abdominal pain and on CT was found to have a circumferentially thickened loop of distal small bowel with associated mesenteric stranding and lymphadenopathy. An octreotide scan demonstrated abnormal activity in the area, concerning for a carcinoid tumor. He underwent an open small bowel resection. Pathology revealed a multifocal carcinoid tumor with 50 distinct nodules and metastasis to mesenteric lymph nodes.
be examined before planning extent of resection. In the presence of metastatic disease, tumor debulking should be conducted as it can be associated with long-term survival and amelioration of symptoms of the carcinoid syndrome. Response rates of 30% to 50% have been reported to chemotherapy regimens based on agents such as doxorubicin, 5-fluorouracil, and streptozocin. However, none of these regimens is associated with a clearly demonstrable impact on the natural history of disease. Ocretide is the most effective pharmacologic agent for management of symptoms of carcinoid syndrome.

Localized small-intestinal lymphoma should be treated with segmental resection of the involved intestine and adjacent mesentery. If the small intestine is diffusely affected by lymphoma, chemotherapy rather than surgical resection should be the primary therapy. The value to adjuvant chemotherapy after resection of localized lymphoma is controversial.

Small-intestinal GISTs should be treated with segmental intestinal resection. If the diagnosis is known prior to resection, wide lymphadenectomy can be avoided as GISTs are rarely associated with lymph node metastases. GISTs are resistant to conventional chemotherapy agents. Imatinib (Gleevec) is a tyrosine kinase inhibitor with potent activity against tyrosine kinase KIT, and it is used in those with metastatic disease. Clinical trials have shown that 80% of patients with unresectable or metastatic GISTs derive clinical benefit from the administration of Imatinib, with 50% to 60% having objective evidence of reduction in tumor volume. Imatinib has shown great promise as a neoadjuvant and adjuvant therapy for GISTs. Studies have emphasized the potential for development of tumor resistance to this agent. In this setting, an alternative tyrosine kinase inhibitor, sunitinib, has been used with good results.

Metastatic cancers affecting the small intestine that are symptomatic should be treated with palliative resection or bypass except in the most advanced cases. Systemic therapy may be offered if effective chemotherapy exists for the primary cancer.

Outcomes
Complete resection of duodenal adenocarcinomas is associated with postoperative 5-year survival rates ranging from 50% to 60%. Complete resection of adenocarcinomas located in the jejunum or ileum is associated with 5-year survival rates of 20% to 30%. Five-year survival rates of 75% to 95% following resection of localized small-intestinal carcinoid tumors have been reported. In the presence of carcinoid tumor-derived liver metastases, 5-year survival rates of 19% to 54% have been reported. The overall 5-year survival rate for patients diagnosed with intestinal lymphoma ranges from 20% to 40%. For patients with localized lymphoma amenable to surgical resection, the 5-year survival rate is 60%.

The recurrence rate following resection of GISTs averages 35%. The 5-year survival rate following surgical resection has been reported to range from 35% to 60%. Both tumor size and mitotic index are independently correlated with prognosis. Low-grade tumors (mitotic index <10 per high-power field) measuring less than 5 cm in diameter are associated with excellent prognosis.

Radiation therapy is a component of multi-modality therapy for many intra-abdominal and pelvic cancers such as those of the cervix, endometrium, ovary, bladder, prostate, and rectum. An undesired side effect of radiation therapy is radiation-induced injury to the small intestine, which can present clinically as two distinct syndromes: acute and chronic radiation enteritis. Acute radiation enteritis is a transient condition that occurs in approximately 75% of patients undergoing radiation therapy for abdominal and pelvic cancers. Chronic radiation is enteritis is inexorable and develops in approximately 5% to 15% of these patients.

Pathophysiology
Radiation induces cellular injury directly and through the generation of free radicals. The principal mechanism of radiation-induced cell death is believed to be apoptosis resulting from free-radical–induced breaks in double-stranded DNA. Because radiation has its greatest impact on rapidly proliferating cells, the small-intestinal epithelium is acutely susceptible to radiation-induced injury. Pathological correlates of this acute injury include villus blunting and a dense infiltrate of leukocytes and plasma cells within the crypts. With severe cases, mucosal sloughing, ulceration, and hemorrhage are observed. The intensity of injury is related to the dose of radiation administered, with most cases occurring in patients who have received at least 4500 cGy. Risk factors for acute radiation enteritis include conditions that may limit splanchic perfusion such as hypertension, diabetes mellitus, coronary artery disease, and restricted mobility of the small intestine due to adhesions. Injury is potentiated by concomitant administration of chemotherapeutic agents, such as doxorubicin, 5-fluorouracil, actinomycin D, and methotrexate, that act as radiation-sensitizers. Because of the intestinal epithelium’s capacity for regeneration, the mucosal injury that is characteristic of acute radiation enteritis resolves after the cessation of radiation therapy.

In contrast, chronic radiation enteritis is characterized by a progressive occlusive vasculitis that leads to chronic ischemia and fibrosis that affects all layers of the intestinal wall, rather than the mucosa alone. These changes can lead to strictures, abscesses, and fistulas, which are responsible for the clinical manifestations of chronic radiation enteritis.

Clinical Presentation
The most common manifestations of acute radiation enteritis are nausea, vomiting, diarrhea, and crampy abdominal pain. Symptoms are generally transient and subside after the discontinuation of radiation therapy. Because the diagnosis is usually obvious, given the clinical context, no specific diagnostic tests are required. However, if patients develop signs suggestive of peritonitis, CT scanning should be performed to rule out the presence of other conditions capable of causing acute abdominal syndromes.

The clinical manifestations of chronic radiation enteritis usually become evident within 2 years of radiation administration, although they can begin as early as several months or as late as decades afterwards. The most common clinical presentations are diarrhea or one of partial small bowel obstruction with nausea, vomiting, intermittent abdominal distention, crampy abdominal pain, and weight loss. The terminal ileum is the most frequently affected segment. Other manifestations of chronic radiation enteritis include complete bowel obstruction, acute or chronic intestinal hemorrhage, and abscess or fistula formation.

Diagnosis
Evaluation of patients suspected of having chronic radiation enteritis should include review of the records of their
radiation treatments for information on total radiation dose administered, fractionation, and volume of treatment. Areas that received high doses should be noted, as lesions subsequently found in imaging studies usually localize to areas that had received high radiation doses. Enteroclysis is the most accurate imaging test for diagnosing chronic radiation enteritis, with reported sensitivities and specificities of over 90% (Fig. 28-25). CT scan findings are neither very sensitive nor specific for chronic radiation enteritis. However, CT scanning should be obtained to rule out the presence of recurrent cancer since its clinical manifestations may overlap with those of chronic radiation enteritis.

Therapy
Most cases of acute radiation enteritis are self-limited. Supportive therapy, including the administration of antiemetics, is usually sufficient. Patients with diarrhea-induced dehydration may require hospital admission and parenteral fluid administration. Rarely are symptoms severe enough to necessitate reduction in or cessation of radiation therapy.

In contrast, the treatment of chronic radiation enteritis represents a formidable challenge. Antidiarrheal agents may have a role in the management of diarrhea while, in those with obstructive symptoms, a low residue diet may be tried. Surgery for this condition is difficult, is associated with high morbidity rates, and should be avoided in the absence of specific indications such as high-grade obstruction, perforation, hemorrhage, intra-abdominal abscesses, and fistulas. The goal of surgery is limited resection of diseased intestine with primary anastomosis between healthy bowel segments. However, the characteristically diffuse nature of fibrosis and dense adhesions among bowel segments can make limited resection difficult to achieve. Further, it is difficult to distinguish between normal and irradiated intestine intraoperatively by either gross inspection or even frozen section analysis. This distinction is important as anastomoses between irradiated segments of intestine have been associated with leak rates as high as 50%. If limited resection is not achievable, an intestinal bypass procedure may be an option, except in cases for which hemorrhage is the surgical indication. There remain cases in which resections extensive enough to cause short bowel syndrome are unavoidable. This condition is discussed in detail below in the “Short Bowel Syndrome” section.

Outcomes
Acute radiation injury to the intestine is self-limited; its severity is not correlated with the probability of chronic radiation enteritis developing. Surgery for chronic radiation enteritis is associated with high morbidity rates and reported mortality rates averaging 10%.

Prevention
In view of significant morbidity associated with radiation enteritis, groups have studies possible measures to reduce or prevent such side effects. Keeping radiation exposure to below 5000 cGy is associated with minimal long-term side effects and is recommended where clinically possible.

Uses of multibeam radiation techniques to minimize the area of maximal radiation exposure, as well as tilt tables to move the bowel out of the pelvic during radiation, are increasingly utilized. Few small studies have suggested that oral sulphasalazine may help reduce the incidence of acute radiation-induced enteritis.

In patients undergoing pelvic surgery that are likely to require postoperative radiation therapy, surgical techniques that keep the small bowel out of the pelvic have been recommended. These measures include use of absorbable mesh sling to separate the pelvis from the true abdominal cavity and prevent the small bowel from being exposed to pelvic radiation.

MECKEL’S DIVERTICULA
Meckel’s diverticulum is the most prevalent congenital anomaly of the gastrointestinal tract, affecting approximately 2% of the general population. Meckel’s diverticuli are designated true diverticuli because their walls contain all the layers found in normal small intestine. Their location varies among individual patients, but they are usually found in the ileum within 100 cm of the ileocecal valve (Fig. 28-26). Approximately 60% of Meckel’s diverticuli contain heterotopic mucosa, of which over 60% consist of gastric mucosa. Pancreatic acini are the next most common; others include Brunner’s glands, pancreatic islets, colonic mucosa, endometriosis, and hepatobiliary tissues. A useful, although crude, mnemonic describing Meckel’s diverticuli is the “rule of twos”: 2% prevalence, 2:1 male predominance, location 2 feet proximal to the ileocecal valve in adults, and half of those who are symptomatic are under 2 years of age.

Pathophysiology
During the eighth week of gestation, the omphalomesenteric (vitelline) duct normally undergoes obliteration. Failure or...
incomplete vitelline duct obliteration results in a spectrum of abnormalities, the most common of which is Meckel’s diverticulum. Other abnormalities include omphalomesenteric fistula, enterocyst, and a fibrous band connecting the intestine to the umbilicus. A remnant of the left vitelline artery can persist to form a mesodiverticular band tethering a Meckel’s diverticulum to the ileal mesentery.

Bleeding associated with Meckel’s diverticulum is usually the result of ileal mucosal ulceration that occurs adjacent to acid-producing, heterotopic gastric mucosa located within the diverticulum. Intestinal obstruction associated with Meckel’s diverticulum can result from several mechanisms:

1. Volvulus of the intestine around the fibrous band attaching the diverticulum to the umbilicus
2. Entrapment of intestine by a mesodiverticular band (Fig. 28-27)
3. Intussusception with the diverticulum acting as a lead point
4. Stricture secondary to chronic diverticulitis

Meckel’s diverticuli can be found in inguinal or femoral hernia sacs (known as Littre’s hernia). These hernias, when incarcerated, can cause intestinal obstruction.

Clinical Presentation
Meckel’s diverticuli are asymptomatic unless associated complications arise. The lifetime incidence rate of complications arising in patients with Meckel’s diverticuli has been estimated to be approximately 4% to 6%. However, initial data has suggested that the risk of developing a complication related to Meckel’s diverticulum decreases with age, this has been questioned. In a population-based review at Olmsted County, Cullen and colleagues suggested that the risk of developing Meckel’s diverticulum–related complications does not change with age.

The most common presentations associated with symptomatic Meckel’s diverticuli are bleeding, intestinal obstruction, and diverticulitis. Bleeding is the most common presentation in children with Meckel’s diverticuli, representing over 50% of Meckel’s diverticulum-related complications among patients less than 18 years of age. Bleeding associated with Meckel’s diverticulum is rare among patients older than 30 years of age.

Intestinal obstruction is the most common presentation in adults with Meckel’s diverticuli. Diverticulitis, present in 20% of patients with symptomatic Meckel’s diverticuli, is associated with a clinical syndrome that is indistinguishable from acute appendicitis. Neoplasms, most commonly carcinoid tumors, are present in 0.5% to 3.2% of symptomatic Meckel’s diverticuli that are resected.

Diagnosis
Most Meckel’s diverticuli are discovered incidentally on radiographic imaging, during endoscopy, or at the time of surgery. In the absence of bleeding, Meckel’s diverticuli rarely are diagnosed prior to the time of surgical intervention. For those presenting with symptoms suggestive of a Meckel’s diverticulum, confirmatory imaging can be challenging. The sensitivity of CT scanning for the detection of Meckel’s diverticuli is too low to be clinically useful. Enterocolysis is associated with an accuracy of 75% but is usually not applicable during acute presentations of complications related to Meckel’s diverticuli. Radionuclide scans (99mTc-pertechnetate) can be helpful in the diagnosis of Meckel’s diverticulum; this test is, however, positive only when the diverticulum contains associated ectopic gastric mucosa that is capable of uptake of the tracer (Fig. 28-28). The accuracy of radionuclide scanning is reported to be 90% in pediatric patients but less than 50% in adults. Angiography can localize the site of bleeding during acute hemorrhage related to Meckel’s diverticuli.

Figure 28-27. Meckel’s diverticulum with mesodiverticular band (A). One mechanism by which Meckel’s diverticuli can cause small bowel obstruction is entrapment of the intestine by a mesodiverticular band (B).
**Therapy**

The surgical treatment of symptomatic Meckel’s diverticuli should consist of diverticulectomy with removal of associated bands connecting the diverticulum to the abdominal wall or intestinal mesentery. If the indication for diverticulectomy is bleeding, segmental resection of ileum that includes both the diverticulum and the adjacent ileal peptic ulcer should be performed. Segmental ileal resection may also be necessary if the diverticulum contains a tumor or if the base of the diverticulum is inflamed or perforated.

The management of incidentally found (asymptomatic) Meckel’s diverticuli is controversial. Until recently, most authors recommended against prophylactic removal of asymptomatic Meckel’s diverticuli, given the low lifetime incidence of complications. Supporting this approach, a meta-analysis has shown that 758 prophylactic diverticulectomies needed to be performed to prevent one Meckel’s-related death. Others have had greater enthusiasm for prophylactic diverticulectomy has appeared in the literature. Proponents of this approach cite the minimal morbidity associated with removing Meckel’s diverticuli and the possibility that previous estimates of the lifetime incidence of complications related to Meckel’s diverticuli may be erroneously low. Many have advocated a selective approach, with a recommendation to remove diverticuli in patients younger than 50 years of age, or those with band attachments, those with ectopic tissue, or those >2 cm in length on the assumption that these diverticuli are more likely to develop complications. No controlled data supporting or refuting these recommendations exist.

**ACQUIRED DIVERTICULA**

Acquired diverticuli are designated *false diverticuli* because their walls consist of mucosa and submucosa but lack a complete muscularis. Acquired diverticuli are more common in the duodenum, and tend to be located near the ampulla; such diverticuli are known as *periampullary, juxtapapillary, and perivaterian diverticuli*. Approximately 75% of juxtapapillary diverticuli arise on the medial wall of the duodenum. Acquired diverticuli in the jejunum or ileum are known as *jejunoileal diverticuli*. Eighty percent of jejunoileal diverticuli are localized to the jejunum, 15% to the ileum, and 5% to both jejunum and ileum. Diverticuli in the jejunum tend to be large and accompanied by multiple other diverticuli, whereas those in the ileum tend to be small and solitary.

The prevalence of duodenal diverticuli, as detected on upper GI examinations (Fig. 28-29), has been reported to range from 0.16% to 6%. Their prevalence, as detected during ERCP examinations, has been reported to range from 5% to 27%. A 23% prevalence rate has been reported in an autopsy series. The prevalence of duodenal diverticuli increases with age; they are
Duodenal diverticulum. This contrast radiograph demonstrates a duodenal diverticulum (arrows) that extends medially into the substance of the head of the pancreas.

Figure 28-30. Jejunoileal diverticuli. This picture demonstrates incidental jejunal diverticuli identified during a laparoscopic cholecystectomy. The diverticuli are typically located on the mesenteric aspect of the jejunum. Resection was not indicated as the diverticuli were asymptomatic.

Pathophysiology
The pathogenesis of acquired diverticuli is hypothesized to be related to acquired abnormalities of intestinal smooth muscle or dysregulated motility, leading to herniation of mucosa and submucosa through weakened areas of muscularis.

Acquired diverticuli can be associated with bacterial overgrowth, leading to vitamin B₁₂ deficiency, megaloblastic anemia, malabsorption, and steatorrhea. Periampullary duodenal diverticuli have been described to become distended with intraluminal debris and to compress the common bile duct or pancreatic duct, thus causing obstructive jaundice or pancreatitis, respectively. Jejunoileal diverticuli can also cause intestinal obstruction through intussusception or compression of adjacent bowel.
Clinical Presentation

Acquired diverticuli are asymptomatic unless associated complications arise. Such complications are estimated to occur in 6% to 10% of patients with acquired diverticuli and include intestinal obstruction, diverticulitis, hemorrhage, perforation, and malabsorption. Periampullary duodenal diverticuli may be associated with choledocholithiasis, cholangitis, recurrent pancreatitis, and sphincter of Oddi dysfunction. However, a clear link between the presence of the diverticuli and the development of these conditions has not been demonstrated. Symptoms such as intermittent abdominal pain, flatulence, diarrhea, and constipation are reported to be present in 10% to 30% of patients with jejunoileal diverticuli. The relationship between these symptoms and the presence of the diverticuli is similarly unclear.

Diagnosis

Most acquired diverticuli are discovered incidentally on radiographic imaging, during endoscopy, or at the time of surgery. On ultrasound and CT scanning, duodenal diverticuli may be mistaken for pancreatic pseudocysts and fluid collections, biliary cysts, and periampullary neoplasms. These lesions can be missed on endoscopy, particularly with forward-viewing endoscopes, and are best diagnosed on upper gastrointestinal radiographs. Enterocolysis is the most sensitive test for detecting jejunoileal diverticuli.

Therapy

Asymptomatic-acquired diverticuli should be left alone. Bacterial overgrowth associated with acquired diverticuli is treated with antibiotics. Other complications, such as bleeding and diverticulitis, are treated with segmental intestinal resection for diverticuli located in the jejunum or ileum.

Bleeding and obstruction related to lateral duodenal diverticuli are generally treated with diverticulectomy alone. Treatment of such complications in medial duodenal diverticuli that penetrate the substance of the pancreas can be very challenging. Complications related to these medial duodenal diverticuli should be managed nonoperatively if possible, using endoscopy. In emergent situations, bleeding related to medial duodenal diverticuli can be controlled using a lateral duodenotomy and oversewing of the bleeding vessel. Similarly, perforation can be managed with wide drainage rather than complex surgery. Whether diverticulectomy should be done in patients with biliary or pancreatic symptoms is controversial and is not routinely recommended.

Mesenteric Ischemia

Mesenteric ischemia can present as one of two distinct clinical syndromes: acute mesenteric ischemia and chronic mesenteric ischemia.

Four distinct pathophysiologic mechanisms can lead to acute mesenteric ischemia:

1. Arterial embolus
2. Arterial thrombosis
3. Vasospasm (also known as nonocclusive mesenteric ischemia or NOMI)
4. Venous thrombosis

Emboli is the most common cause of acute mesenteric ischemia and is responsible for over 50% of cases. The embolic source is usually in the heart, most often the left atrial or ventricular thrombi or valvular lesions. Indeed, up to 95% of patients with acute mesenteric ischemia due to emboli will have a documented history of cardiac disease. Embolism to the superior mesenteric artery accounts for 50% of cases; most of these emboli become wedged and cause occlusion at branch points in the mid- to distal superior mesenteric artery, usually distal to the origin of the middle colic artery. In contrast, acute occlusions due to thrombosis tend to occur in the proximal mesenteric arteries, near their origins. Acute thrombosis is usually superimposed on preexisting atherosclerotic lesions at these sites. NOMI is the result of vasospasm and is usually diagnosed in critically ill patients receiving vasopressor agents.

Mesenteric venous thrombosis accounts for 5% to 15% of cases of acute mesenteric ischemia and involves the superior mesenteric vein in 95% of cases. The inferior mesenteric vein is only rarely involved. Mesenteric venous thrombosis is classified as primary if no etiologic factor is identifiable, or as secondary if an etiologic factor, such as heritable or acquired coagulation disorders, is identified.

Regardless of the pathophysiologic mechanism, acute mesenteric ischemia can lead to intestinal mucosal sloughing within 3 hours of onset and full-thickness intestinal infarction by 6 hours.

In contrast, chronic mesenteric ischemia develops insidiously, allowing for development of collateral circulation, and, therefore, rarely leads to intestinal infarction. Chronic mesenteric arterial ischemia results from atherosclerotic lesions in the main splanchnic arteries (celiac, superior mesenteric, and inferior mesenteric arteries). In most patients with symptoms attributable to chronic mesenteric ischemia, at least two of these arteries are either occluded or severely stenosed. A chronic form of mesenteric venous thrombosis can involve the portal or splenic veins and may lead to portal hypertension, with resulting esophageal varices, splenomegaly, and hypersplenism.

Severe abdominal pain, out of proportion to the degree of tenderness on examination is the hallmark of acute mesenteric ischemia, regardless of the pathophysiologic mechanism. The pain is typically perceived to be colicky and most severe in the mid-abdomen. Associated symptoms include nausea, vomiting, and diarrhea. Physical findings are characteristically absent early in the course of ischemia. With the onset of bowel infarction, abdominal distension, peritonitis, and passage of bloody stools occur.

Chronic mesenteric ischemia presents insidiously. Postprandial abdominal pain is the most prevalent symptom, producing a characteristic aversion to food (“food fear”) and weight loss. These patients are often thought to have a malignancy and suffer a prolonged period of symptoms before the correct diagnosis is made.

Most patients with chronic mesenteric venous thrombosis are asymptomatic because of the presence of extensive collateral venous drainage routes; this condition is usually discovered as an incidental finding on imaging studies. However, some patients with chronic mesenteric venous thrombosis present with bleeding from esophagogastric varices.

The diagnosis and management of these disorders, which are of primary vascular origin, are discussed in the section on “Mesenteric Artery Occlusive Disease” in Chapter 23 “Arterial Disease.”

Miscellaneous Conditions

Obscure GI Bleeding

Up to 90% of lesions responsible for GI bleeding are within the reach of EGD and colonoscopy. Obscure GI bleeding refers to gastrointestinal bleeding for which no source has been identified by routine endoscopic studies (EGD and colonoscopy). Overt
GI bleeding refers to the presence of hematemesis, melena, or hematochezia. In contrast, occult GI bleeding occurs in the absence of overt bleeding and is identified on laboratory tests (e.g., iron-deficiency anemia) or examination of the stool (e.g., positive guaiac test). Occult GI bleeding is occult in 20% of cases.68

Obscure bleeding can be frustrating for both the patient and the clinician, and this is particularly true for obscure-overt bleeding, which cannot be localized despite aggressive diagnostic measures. Most of the small bowel is beyond the reach of these examinations and, hence, contains most lesions responsible for obscure GI bleeding. Small intestinal angiodysplasias account for approximately 75% of cases in adults, and neoplasms account for approximately 10%. Meckel’s diverticulum is the most common etiology of obscure GI bleeding in children. Other etiologies of obscure GI bleeding include Crohn’s disease, infectious enteritides, nonsteroidal anti-inflammatory drug (NSAID)-induced ulcers and erosions, vasculitis, ischemia, varices, diverticuli, and intussusception.

The diagnostic evaluation of patients with obscure GI bleeding should be tailored to the severity of bleeding and to the availability of technology and expertise. Enteroscopy is playing an increasingly important role. Several endoscopic techniques for visualizing the small intestine are available: push enteroscopy, sonde enteroscopy, intraoperative enteroscopy, double balloon endoscopy, and wireless capsule enteroscopy.

Push enteroscopy entails advancing a long endoscope (such as a pediatric or adult colonoscope or a specialized instrument) beyond the ligament of Treitz into the proximal jejunum. This procedure can allow for visualization of approximately 60 cm of the proximal jejunum. Diagnostic yield in patients with obscure GI bleeding ranges from 3% to 65%. In addition to diagnosis, push enteroscopy allows for cauterization of bleeding sites.

In Sonde enteroscopy, a long, thin fiberoptic instrument is propelled through the intestine by peristalsis following inflation of a balloon at the instrument’s tip. Visualization is done during instrument withdrawal; approximately 50% to 75% of the small intestinal mucosa can be examined. However, this instrument lacks biopsy or therapeutic capability. Further, it lacks tip deflection capability, limiting complete mucosal visualization, and has therefore been abandoned in favor of capsule endoscopy.

Wireless capsule endoscopy is an excellent tool in the patient who is hemodynamically stable but continues to bleed. This technique has reported success rates as high as 90% in identifying a small bowel pathology. In a randomized study of patients with obscure GI bleeding, evaluation with capsule endoscopy vs. small bowel contrast study had a much higher diagnostic yield (30% vs. 7%, respectively); however, this did not translate into an improvement in outcomes. The rates of rebleeding, hospitalization, need for blood transfusion, and therapeutic interventions were similar between the two arms.

The inability to perform biopsies or carry out any therapeutic interventions of capsule endoscope likely prevents the improved diagnostic yield of the test from translating into improved patient outcomes and highlights the continuing challenge with evaluation of the small bowel.69

For patients in whom bleeding from an obscure GI source has apparently stopped, push enteroscopy or capsule enteroscopy is a reasonable initial study. If these examinations do not reveal a potential source of bleeding, then enterocolysis should be performed. Standard small bowel follow-through examinations are associated with a low diagnostic yield in this setting and should be avoided. 99mTc-pertechnetate scintigraphy to diagnose Meckel’s diverticulum should be considered, although its yield in patients older than 40 years of age is extremely low. If still no diagnosis has been made, a “watch-and-wait” approach is reasonable, although angiography should be considered if the episode of bleeding was overt. Angiography can reveal angiodysplasia and vascular tumors in the small intestine even in the absence of ongoing bleeding. In many instances, however, angiography is done in patients with persistent bleeding from an obscure GI source, and it is also often performed after a 99mTc-labeled RBC scan, which, if positive, is followed by angiography to localize the source of bleeding. Patients who remain undiagnosed but continue to bleed and those with recurrent episodic bleeding significant enough to require blood transfusions should then undergo exploratory laparoscopy or laparotomy with intraoperative enteroscopy. An endoscope (usually a colonoscope) is inserted into the small bowel through peroral intubation or through an enterotomy made in the small bowel or cecum. The endoscope is advanced by successively telescoping short segments of intestine onto the end to the instrument. In addition to the endoscopic image, the transilluminated bowel should be examined externally with the operating room lights dimmed, as this maneuver may facilitate the identification of angiodysplasias. Identified lesions should be marked with a suture placed on the serosal surface of the bowel; these lesions can be resected after completion of endoscopy. Examination should be performed during instrument insertion rather than withdrawal because instrument-induced mucosal trauma can be confused with angiodysplasias.

Figure 28-31 provides a diagnostic and management algorithm for patients with obscure GI bleeding.

Small Bowel Perforation

Prior to the 1980s, duodenal perforation due to peptic ulcer disease was the most common form of small bowel perforation. Today, iatrogenic injury incurred during gastrointestinal endoscopy is the most common cause of small bowel perforation. Other etiologies of small bowel perforation include infections (especially tuberculosis, typhoid, and CMV), Crohn’s disease, ischemia, drugs (e.g., potassium- and NSAID-induced ulcers), radiation-induced injury, Meckel’s and acquired diverticuli, neoplasms (especially lymphoma, adenocarcinoma, and melanoma), and foreign bodies.

Among iatrogenic injuries, duodenal perforation during endoscopic retrograde cholangiography (ERCP) with endoscopic sphincterotomy (ES) is the most common. With improved technique and technology, the incidence of this is decreasing but remains at around 0.5%.70 The Stapfer classification is commonly used to categorize different types of ERCP-related perforations. These are:

- **Type I**: Free bowel wall perforation
- **Type II**: Retroperitoneal duodenal perforation secondary to periampullary injury
- **Type III**: Perforation of the pancreatic or bile duct
- **Type IV**: Retroperitoneal air alone

Type II (retroperitoneal duodenal injuries) are the most common and can often be managed nonsurgically. Manifestations of such contained duodenal perforation following ERCP can resemble those of ERCP-induced pancreatitis, including hyperamylasemia.
CT scanning is the most sensitive test for diagnosing duodenal perforations; positive findings include pneumoperitoneum for free perforations, but more commonly retroperitoneal air, contrast extravasation, and paraduodenal fluid collections. If all patients undergoing a therapeutic ERCP are imaged with a CT scan following the procedure, up to 30% will have evidence of air in the retroperitoneum, but the majority are asymptomatic. These patients do not require any specific therapy.71

True cases of retroperitoneal perforations of the duodenum can be managed nonoperatively in the absence of progression and sepsis. However, intraoperative duodenal perforations require surgical repair with pyloric exclusion and gastrojejunostomy or tube duodenostomy. Iatrogenic small bowel perforation incurred during endoscopy, if immediately recognized, can sometimes be repaired using endoscopic techniques.

Perforation of the jejunum and ileum occurs into the peritoneal cavity and usually causes overt symptoms and signs, such as abdominal pain, tenderness, and distention accompanied by fever and tachycardia. Plain abdominal radiographs may reveal free intraperitoneal air if intraperitoneal perforation has occurred. If perforation is suspected but not clinically obvious, CT scanning should be performed. Jejunal and ileal perforations require surgical repair or segmental resection.

**Chylous Ascites**

Chylous ascites refers to the accumulation of triglyceride-rich peritoneal fluid with a milky or creamy appearing, caused by the presence of intestinal lymph in the peritoneal cavity. Chylomicrons, produced by the intestine and secreted into lymph during the absorption of long-chain fatty acids, account for the characteristic appearance and triglyceride content of chyle.

The most common etiologies of chylous ascites in Western countries are abdominal malignancies and cirrhosis. In Eastern and developing countries, infectious etiologies, such as tuberculosis and filariasis, account for most cases. Chylous ascites can also develop as a complication of abdominal and thoracic operations and trauma. Operations particularly associated with this complication include abdominal aortic aneurysm repair, retroperitoneal lymph node dissection, inferior vena cava resection, and liver transplantation. Other etiologies of chylous ascites include congenital lymphatic abnormalities (e.g., primary lymphatic hypoplasia), radiation, pancreatitis, and right-sided heart failure.

Three mechanisms have been postulated to cause chylous ascites: (a) exudation of chyle from dilated lymphatics on the wall of the bowel and in the mesentery caused by obstruction of lymphatic vessels at the base of the mesentery or the cisterna chilii (e.g., by malignancies), (b) direct leakage of chyle through a lymphoperitoneal fistula (e.g., those that develop as a result of trauma or surgery), and (c) exudation of chyle through the wall of dilated retroperitoneal lymphatic vessels (e.g., in congenital lymphangiectasia or thoracic duct obstruction).

Patients with chylous ascites develop abdominal distention over a period of weeks to months. Postoperative chylous ascites can present acutely during the first postoperative week. Delayed presentations following surgery can occur if the mechanism of ascites formation is adhesion-induced lymphatic obstruction rather than lymphatic vessel disruption.

---

**Figure 28-31.** Diagnostic and management algorithm for obscure gastrointestinal bleeding.
Paracentesis is the most important diagnostic test. Chyle typically has a cloudy and turbid appearance; however, it may be clear in fasting patients (such as those in the immediate postoperative period). Fluid triglyceride concentrations above 110 mg/dL are diagnostic. CT scanning may be useful in identifying pathological intraabdominal lymph nodes and masses and in identifying extent and localization of fluid. Lymphangiography and lymphoscintigraphy may help localize lymph leaks and obstruction; this information is particularly useful for surgical planning.

There is little data on optimal management of patients with chylous ascites. The general approach is to focus on evaluating and treating the underlying causes, especially for patients with infectious, inflammatory, or hemodynamic etiologies for this condition.

Most patients respond to administration of a high-protein and low-fat diet supplemented with medium-chain triglycerides. This regimen is designed to minimize chyle production and flow. Medium-chain triglycerides are absorbed by the intestinal epithelium and are transported to liver through the portal vein; they do not contribute to chylomicron formation.

Patients who do not respond to this approach should be fasted and placed on TPN. Octreotide can further decrease lymph flow. Paracentesis is indicated for respiratory difficulties related to abdominal distention. Overall, two-thirds of patients will respond to conservative therapy. However, one-third of patients will require surgical therapy for chylous ascites. In general, postoperative and trauma-related cases that fail to respond to initial nonoperative therapy are best managed by surgical repair. Lymphatic leaks are localized and repaired with fine nonabsorbable sutures. If extravasation of chyle is localized to the periphery of the small bowel mesentery, then a limited small bowel resection can be performed instead. For patients who are poor surgical candidates and who do not respond to prolonged conservative therapy, peritoneovenous shunting may be an option. However, these shunts are associated with high rates of complications, including sepsis and disseminated intravascular coagulation. Because of the viscosity of chyle, these shunts are associated with a high occlusion rate.

**Intussusception**

Intussusception refers to a condition where one segment of the intestine becomes drawn in to the lumen of the distal segment of the bowel. It is usually seen in the pediatric population, where the ileum intussuscepts into the cecum (ileocolic intussusception). In children, it is often an idiopathic condition and treated nonsurgically by radiological reduction.

Intussusceptions are far less common and usually have a distinct pathologic lead point, which can be malignant in up to one-half of cases. They commonly present with a history of intermittent abdominal pain and signs and symptoms of bowel obstruction. CT scan is the investigation of choice, where a “target sign” may be seen (Fig. 28-32). Treatment is surgical resection of the involved segment and the lead point, which needs to undergo pathological evaluation to rule out an underlying malignancy.

With increasing use of CT imaging, target signs are sometimes seen on CT scans of patients who do not have a clinical presentation indicative of bowel obstruction. In such cases, the finding is of little clinical significance and is probably related to normal peristalsis.

In patients who have undergone a Roux-en-Y gastric bypass surgery, an atypical form of intussusception has been increasingly described. In these cases, the distal bowel is drawn into the lumen of the proximal bowel (retrograde intussusception). These intussusceptions are usually not associated with a lead point and may represent a motility disorder of the bowel following the Roux-en-Y reconstruction. Surgical reductions without resection have been successfully reported in these patients.

**Pneumatosis Intestinalis**

Pneumatosis intestinalis indicates the presence of gas within the bowel wall. It may affect any region of the GI tract, but it is most commonly seen in the jejunum. Pneumatosis intestinalis is not a disease but merely a sign that can be idiopathic or associated with many intestinal or nonintestinal disorders, such as obstructive pulmonary disease and asthma. Most cases of pneumatosis intestinalis are secondary to an identifiable cause, and 15% are idiopathic. The pathogenesis of pneumatosis intestinalis is not fully understood.

The surgical interest in this finding is the association of it with bowel ischemia and infarction, both of which necessitate emergent surgical intervention (Fig. 28-33). Thus, patients with this radiological finding need to be fully evaluated and monitored closely to rule out such intraabdominal catastrophes.
SHORT BOWEL SYNDROME

Intestinal resection is performed for many of the diseases discussed in this chapter and generally is associated with minimal morbidity. However, when extent of resection is great enough, a devastating condition known as short bowel syndrome may result. Although the best definition of short bowel syndrome is likely a functional one, reflecting a state of significant malabsorption of both macronutrients and micronutrients, some have used a more anatomical definition with it being arbitrarily defined as the presence of less than 200 cm of residual small intestine in adult patients.\(^5\)

In adults, the most common etiologies of short bowel syndrome are acute mesenteric ischemia, malignancy, and Crohn’s disease. Seventy-five percent of cases result from resection of a large amount of small bowel at a single operation. Twenty-five percent of cases result from the cumulative effects of multiple operations during which small intestine is resected. This latter pattern is typical of patients with Crohn’s disease who develop short bowel syndrome; the former is typical of patients with acute mesenteric ischemia who develop intestinal infarction. In pediatric patients, intestinal atresias, volvulus, and necrotizing enterocolitis are the most common etiologies of short bowel syndrome.

The prevalence of short bowel syndrome is hard to estimate due to its multifactorial nature and wide spectrum of presentation and treatment, which may include home total parenteral nutrition (TPN).

Pathophysiology

Resection of less than 50% of the small intestine is generally well tolerated. However, clinically significant malabsorption occurs when greater than 50% to 80% of the small intestine has been resected. Among adult patients who lack a functional colon, lifelong TPN dependence is likely to persist if there is less than 100 cm of residual small intestine. Among adult patients who have an intact and functional colon, lifelong TPN dependence is likely to persist if there is less than 60 cm of residual small intestine. Among infants with short bowel syndrome, weaning from TPN-dependence has been achieved with as little as 10 cm of residual small intestine.

Residual bowel length is not the only factor predictive of achieving independence from TPN (enteral autonomy), however. Other determinants of the severity of malabsorption include the presence or absence of an intact colon, as indicated previously. The colon has the capacity to absorb large fluid and electrolyte loads. In addition, the colon can play an important, albeit small, role in nutrient assimilation by absorbing short chain fatty acids. Second, an intact ileocecal valve is believed to be associated with decreased malabsorption. The ileocecal valve delays transit of chyme from the small intestine into the colon, thereby prolonging the contact time between nutrients and the small-intestinal absorptive mucosa. Third, healthy, rather than diseased, residual small intestine is associated with decreased severity of malabsorption. Fourth, resection of jejunum is better tolerated than resection of ileum, as the capacity for bile salt and vitamin B\(_12\) absorption is specific to the ileum (Table 28-11).

During the first 1 to 2 years following massive small bowel resection, the remaining intestine undergoes compensatory adaptation, as discussed previously. Clinically, the period of adaptation is associated with reductions in volume and frequency of bowel movements, increases in the capacity for enteral nutrient assimilation, and reductions in TPN requirements. As this process completes, some patients are successfully weaned off TPN. Understanding the mechanisms mediating intestinal adaptation may suggest strategies for enhancing adaptation in patients with short bowel syndrome who are unable to achieve independence from TPN. To date, the phenomenon of intestinal adaptation in patients remains poorly understood.\(^13\)

Malabsorption in patients who have undergone massive small bowel resection is exacerbated by a characteristic hypergastrinemia-associated gastric acid hypersecretion that persists for 1 to 2 years postoperatively. The increased acid

---

**Table 28-11**

**Risk factors for development of short bowel syndrome after massive small bowel resection**

<table>
<thead>
<tr>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small bowel length &lt;200 cm</td>
</tr>
<tr>
<td>Absence of ileocecal valve</td>
</tr>
<tr>
<td>Absence of colon</td>
</tr>
<tr>
<td>Diseased remaining bowel (e.g., Crohn’s disease)</td>
</tr>
<tr>
<td>Ileal resection</td>
</tr>
</tbody>
</table>
load delivered to the duodenum inhibits absorption by a variety of mechanisms, including the inhibition of digestive enzymes, most of which function optimally under alkaline conditions.

**Therapy**

**Medical Therapy.** For patients after massive small bowel resection, the initial treatment priorities include management of the primary condition precipitating the intestinal resection and the repletion of fluid and electrolytes lost in the severe diarrhea that characteristically occurs. Most patients will require TPN, at least initially. Enteral nutrition should be gradually introduced, once ileus has resolved. High-dose histamine-2 receptor antagonists or proton pump inhibitors should be administered to reduce gastric acid secretion. Antimotility agents, such as loperamide hydrochloride or diphenoxylate, may be administered to delay small-intestinal transit. Octreotide can be administered to reduce the volume of gastrointestinal secretions, although, in animal models, its use is associated with an inhibition of intestinal adaptation.

During the period of adaptation, generally lasting 1 to 2 years postoperatively, TPN and enteral nutrition are titrated to allow for independence from TPN. Patients who remain dependent on TPN face substantial TPN-associated morbidities including catheter sepsis, venous thrombosis, liver and kidney failure, and osteoporosis. Liver failure is a significant source of morbidity and often leads to liver transplantation (always in combination with small bowel transplantation). Due to these complications, patients with short bowel syndrome on TPN have a reduced life expectancy, with 5-year survival rates of 50% to 75%.

**Nontransplant Surgical Therapy.** Among patients with stomas, restoration of intestinal continuity should be performed whenever possible, to capitalize on the absorptive capacity of all residual intestine. Other forms of nontransplant surgery designed to improve intestinal absorption are associated with unclear efficacy and/or substantial morbidities and therefore should not be applied routinely.

The goal of these operations is to increase nutrient and fluid absorption by either slowing intestinal transit or increasing intestinal length. Operations designed to slow intestinal transit include segmental reversal of the small bowel, interposition of a segment of colon between segments of small bowel, construction of small-intestinal valves, and electrical pacing of the small intestine. Reported experience with these procedures is limited to case reports or series of a few cases. Objective evidence of increased absorption is lacking; further, these procedures are frequently associated with intestinal obstruction.

The intestinal lengthening operation for which has the longest history is the longitudinal intestinal lengthening and tailoring (LILT) procedure, first described by Bianchi in 1980. The procedure entails separation of the dual vasculature of the small intestine, followed by longitudinal division of the bowel with subsequent isoperistaltic end-to-end anastomosis. This procedure has the potential to double the length of small intestine to which it is applied. This procedure has generally been used for pediatric patients with dilated residual small bowel.

An alternative surgical approach to lengthening the small bowel is the serial transverse enteroplasty procedure (STEP) has been described. This procedure is designed to accomplish lengthening of dilated small intestine without the need for separating its dual vasculature (Fig. 28-33). A report from an international registry of 111 patients showed that 47% of patients achieved enteral autonomy at a median follow-up of 21 months.

**Intestinal Transplantation.** This complex procedure is being increasingly performed to treat patients with short bowel syndrome. The currently accepted indication for intestinal transplantation is the presence of life-threatening complications attributable to intestinal failure and/or long-term TPN therapy. Specific complications for which intestinal transplantation is indicated include (a) impending or overt liver failure, (b) thrombosis of major central veins, (c) frequent episodes of catheter-related sepsis, and (d) frequent episodes of severe dehydration.

Of the transplants involving the intestine, 37% were intestine-alone transplants, 30% included intestine, liver, and pancreas, and 24% were intestine and liver.

Isolated intestinal transplantation is used for patients with intestinal failure who have no significant liver disease or failure of other organs. Combined intestine/liver transplantation is used for patients with both intestinal and liver failure. Multivisceral transplantation has been used for patients with giant desmoid tumors involving the vascular supply of the liver and pancreas as well as that of the intestine, for diffuse gastrointestinal motility disturbances, and for diffuse splanchic thrombosis. Nearly 80% of survivors have full intestinal graft function with no need for TPN. However, morbidities associated with intestinal transplantation are substantial and include acute and chronic rejection, CMV infection, and posttransplant lymphoproliferative disease.

**Alternative Therapies.** Pharmacologic and biologic therapies designed to expand intestinal mucosal surface area or to enhance the efficiency of intestinal absorption are beginning to undergo clinical evaluation. Promising regimens include GLP-2 and the combination of glutamine and growth hormone with a modified, high-carbohydrate diet.

**Outcomes**

Approximately 50% to 70% of patients with short bowel syndrome who initially require TPN are ultimately able to achieve independence from TPN. Prognosis for achieving enteral autonomy is better among pediatric patients than among adults.

Information on survival among patients with short bowel syndrome is limited. In a recently reported study of 124 adults with short bowel syndrome due to nonmalignant etiologies, the survival rates at 2 and 5 years of follow up were 86% and 45%, respectively. Patients with end-enterostomies and those having less than 50 cm of residual small intestine had significantly worse survivals than those without these features.

No randomized trials comparing intestinal transplantation to chronic TPN administration among patients with short bowel syndrome have been reported. One-, 5-, and 10-year graft survival rates of intestine-alone recipients were 80%, 44%, and 26%; while those for intestine and liver and intestine, liver, and pancreas were 62%, 45%, 36% and 69%, 48%, 33%, respectively.

**REFERENCES**

Entries highlighted in bright blue are key references.


EMBRYOLOGY AND ANATOMY

Embryology

The embryonic gastrointestinal tract begins developing during the fourth week of gestation. The primitive gut is derived from the endoderm and divided into three segments: foregut, midgut, and hindgut. Both midgut and hindgut contribute to the colon, rectum, and anus.

The midgut develops into the small intestine, ascending colon, and proximal transverse colon, and receives blood supply from the superior mesenteric artery. During the sixth week of gestation, the midgut herniates out of the abdominal cavity and then rotates 270° counterclockwise around the superior mesenteric artery to return to its final position inside the abdominal cavity during the tenth week of gestation. The hindgut develops into the distal transverse colon, descending colon, rectum, and anal canal.
proximal anus, all of which receive their blood supply from the inferior mesenteric artery. During the sixth week of gestation, the distal-most end of the hindgut, the cloaca, is divided by the urorectal septum into the urogenital sinus and the rectum.

The distal anal canal is derived from ectoderm and receives its blood supply from the internal pudendal artery. The dentate line divides the endodermal hindgut from the ectodermal distal anal canal.

**Anatomy**

The large intestine extends from the ileocecal valve to the anus. It is divided anatomically and functionally into the colon, rectum, and anal canal. The wall of the colon and rectum comprise four distinct layers: mucosa, submucosa, muscularis propria (inner circular muscle, outer longitudinal muscle), and serosa. In the colon, the outer longitudinal muscle is separated into three teniae coli, which converge proximally at the appendix and distally at the rectum, where the outer longitudinal muscle layer is circumferential. In the distal rectum, the inner smooth muscle layer coalesces to form the internal anal sphincter. The intraperitoneal colon and proximal one-third of the rectum are covered by serosa; the mid and lower rectum lack serosa.

**Colon Landmarks.** The colon begins at the junction of the terminal ileum and cecum and extends approximately 150 cm (3 to 5 feet) to the rectum. The rectosigmoid junction is found at approximately the level of the sacral promontory and is arbitrarily described as the point at which the three teniae coli coalesce to form the outer longitudinal smooth muscle layer of the rectum. The cecum is the widest diameter portion of the colon (normally 7.5–8.5 cm) and has the thinnest muscular wall. As a result, the cecum is most vulnerable to perforation and least vulnerable to obstruction. The ascending colon is usually fixed to the retroperitoneum. The hepatic flexure marks the transition to the transverse colon. The transverse colon is relatively mobile, but it is tethered by the gastrocolic ligament and colonic mesentery. The greater omentum is attached to the anterior/superior edge of the transverse colon. These attachments explain the characteristic triangular appearance of the transverse colon observed during colonoscopy. The splenic flexure marks the transition from the transverse colon to the descending colon. The attachments between the splenic flexure and the spleen (the lienocolic ligament) can be short and dense, making mobilization of this flexure during colectomy challenging. The descending colon is relatively fixed to the retroperitoneum. The sigmoid colon is the narrowest part of the large intestine and is extremely mobile. Although the sigmoid colon is usually located in the left lower quadrant, redundancy and mobility can result in a portion of the sigmoid colon residing in the right lower quadrant. This mobility explains why volvulus is most common in the sigmoid colon and why diseases affecting the sigmoid colon, such as diverticulitis, may occasionally present as right-sided abdominal pain. The narrow caliber of the sigmoid colon makes this segment of the large intestine the most vulnerable to obstruction.

**Colon Vascular Supply.** The arterial supply to the colon is highly variable (Fig. 29-1). In general, the superior mesenteric artery branches into the ileocolic artery (absent in up to 20% of people), which supplies blood flow to the terminal ileum and proximal ascending colon; the right colic artery, which supplies the ascending colon; and the middle colic artery, which supplies the transverse colon. The inferior mesenteric artery branches into the left colic artery, which supplies the descending colon; several sigmoidal branches, which supply the sigmoid colon; and the superior rectal artery, which supplies the proximal rectum. The terminal branches of each artery form anastomoses with the terminal branches of the adjacent artery and communicate via the marginal artery of Drummond. This arcade is complete in only 15% to 20% of people.
Except for the inferior mesenteric vein, the veins of the colon parallel their corresponding arteries and bear the same terminology (Fig. 29-2). The inferior mesenteric vein ascends in the retroperitoneal plane over the psoas muscle and continues posterior to the pancreas to join the splenic vein. During a colectomy, this vein is often mobilized independently and ligated at the inferior edge of the pancreas.

Colon Lymphatic Drainage. The lymphatic drainage of the colon originates in a network of lymphatics in the muscularis mucosa. Lymphatic vessels and lymph nodes follow the regional arteries. Lymph nodes are found on the bowel wall (epicolic), along the inner margin of the bowel adjacent to the arterial arcades (paracolic), around the named mesenteric vessels (intermediate), and at the origin of the superior and inferior mesenteric arteries (main).

Colon Nerve Supply. The colon is innervated by both sympathetic (inhibitory) and parasympathetic (stimulatory) nerves, which parallel the course of the arteries. Sympathetic nerves arise from T6–T12 and L1–L3. The parasympathetic innervation to the right and transverse colon is from the vagus nerve; the parasympathetic nerves to the left colon arise from sacral nerves S2–S4 to form the nervi erigentes.

Anorectal Landmarks. The rectum is approximately 12 to 15 cm in length. Three distinct submucosal folds, the valves of Houston, extend into the rectal lumen. Posteriorly, the presecral fascia separates the rectum from the presacral venous plexus and the pelvic nerves. At S4, the rectosacral fascia (Waldeyer’s fascia) extends anteriorly and caudally and attaches to the fascia propria at the anorectal junction. Anteriorly, Denovilliers’ fascia separates the rectum from the prostate and seminal vesicles in men and from the vagina in women. The lateral ligaments support the lower rectum.

The anatomic anal canal extends from the dentate or pectinate line to the anal verge. The dentate or pectinate line marks the transition point between columnar rectal mucosa and squamous anoderm. The anal transition zone includes mucosa proximal to the dentate line that shares histologic characteristics of columnar, cuboidal, and squamous epithelium. Although the anal transition zone was long thought to extend only 1 to 2 cm proximal to the dentate line, it is known that the proximal extent of this zone is highly variable and can be as far as 15 cm proximal to the dentate line. The dentate line is surrounded by longitudinal mucosal folds, known as the columns of Morgagni, into which the anal crypts empty. These crypts are the source of cryptoglandular abscesses (Fig. 29-3). In contrast to the anatomic anal canal, the surgical anal canal begins at the

**Figure 29-1.** Arterial blood supply to the colon. a. = artery.

**Figure 29-2.** Venous drainage of the colon. v. = vein. (Reproduced with permission from Bell RH, Rikkers LF, Mulholland M: Digestive Tract Surgery: A Text and Atlas. Philadelphia, PA: Lippincott Williams & Wilkins; 1996.)
anorectal junction and terminates at the anal verge. The surgical anal canal measures 2 to 4 cm in length and is generally longer in men than in women. It begins at the anorectal junction and terminates at the anal verge.

In the distal rectum, the inner smooth muscle is thickened and comprises the internal anal sphincter that is surrounded by the subcutaneous, superficial, and deep external sphincter. The deep external anal sphincter is an extension of the puborectalis muscle. The puborectalis, iliooccygeus, and pubococcygeus muscles form the levator ani muscle of the pelvic floor (Fig. 29-4).

Anorectal Vascular Supply. The superior rectal artery arises from the terminal branch of the inferior mesenteric artery and supplies the upper rectum. The middle rectal artery arises from the internal iliac; the presence and size of these arteries are highly variable. The inferior rectal artery arises from the internal pudendal artery, which is a branch of the internal iliac artery. A rich network of collaterals connects the terminal arterioles of each of these arteries, thus making the rectum relatively resistant to ischemia (Fig. 29-5).

The venous drainage of the rectum parallels the arterial supply. The superior rectal vein drains into the portal system via the inferior mesenteric vein. The middle rectal vein drains into the internal iliac vein. The inferior rectal vein drains into the internal pudendal vein, and subsequently into the internal iliac vein. A submucosal plexus deep to the columns of Morgagni forms the hemorrhoidal plexus and drains into all three veins.

Anorectal Lymphatic Drainage. Lymphatic drainage of the rectum parallels the vascular supply. Lymphatic channels in the upper and middle rectum drain superiorly into the inferior mesenteric lymph nodes. Lymphatic channels in the lower rectum drain both superiorly into the inferior mesenteric lymph nodes and laterally into the internal iliac lymph nodes. The anal canal has a more complex pattern of lymphatic drainage. Proximal to the dentate line, lymph drains into both the inferior mesenteric lymph nodes and the internal iliac lymph nodes. Distal to the dentate line, lymph primarily drains into the inguinal lymph nodes, but can also drain into the inferior mesenteric lymph nodes and internal iliac lymph nodes.

Figure 29-3. The lining of the anal canal. (Reproduced with permission from Goldberg SM, Gordon PH, Nivatvongs S: Essentials of Anorectal Surgery. Philadelphia, PA: JB Lippincott Company; 1980.)
Anorectal Nerve Supply. Both sympathetic and parasympathetic nerves innervate the anorectum. Sympathetic nerve fibers are derived from L1–L3 and join the preaortic plexus. The preaortic nerve fibers then extend below the aorta to form the hypogastric plexus, which subsequently joins the parasympathetic fibers to form the pelvic plexus. Parasympathetic nerve fibers are known as the nervi erigentes and originate from S2–S4. These fibers join the sympathetic fibers to form the pelvic plexus. Sympathetic and parasympathetic fibers then supply the anorectum and adjacent urogenital organs.

The internal anal sphincter is innervated by sympathetic and parasympathetic nerve fibers; both types of fibers inhibit sphincter contraction. The external anal sphincter and puborectalis muscles are innervated by the inferior rectal branch of the internal pudendal nerve. The levator ani receives innervation from both the internal pudendal nerve and direct branches of S3 to S5. Sensory innervation to the anal canal is provided by the inferior rectal branch of the pudendal nerve. While the rectum is relatively insensitive, the anal canal below the dentate line receives somatic innervation.

Congenital Anomalies
Perturbation of the embryologic development of the midgut and hindgut may result in anatomic abnormalities of the colon, rectum, and anus. Failure of the midgut to rotate and return to the abdominal cavity during the tenth week of gestation results in varying degrees of intestinal malrotation and colonic nonfixation. Failure of canalization of the primitive gut can result in colonic duplication. Incomplete descent of the urogenital septum may result in imperforate anus and associated fistulas to the genitourinary tract. Many infants with congenital anomalies of the hindgut have associated abnormalities in the genitourinary tract.

NORMAL PHYSIOLOGY

Fluid and Electrolyte Exchanges
Water, Sodium, Potassium, Chloride, Bicarbonate, and Ammonia. The colon is a major site for water absorption and electrolyte exchange. Under normal circumstances, approximately 90% of the water contained in ileal fluid is absorbed in the colon (1000–2000 mL/d), but up to 5000 mL of fluid can be absorbed daily. Sodium is absorbed actively via sodium-potassium (Na+/K+) ATPase. The colon can absorb up to 400 mEq of sodium per day. Water accompanies the transported sodium and is absorbed passively along an osmotic gradient. Potassium is actively secreted into the colonic lumen and absorbed by passive diffusion. Chloride is absorbed actively via a chloride-bicarbonate exchange.

Bacterial degradation of protein and urea produces ammonia. Ammonia is subsequently absorbed and transported to the liver. Absorption of ammonia depends in part on intraluminal pH. A decrease in colonic bacteria (e.g., due to broad-spectrum antibiotic use) and/or a decrease in intraluminal pH (e.g., due to lactulose administration) will decrease ammonia absorption.

Colonic Microflora and Intestinal Gas
Approximately 30% of fecal dry weight is composed of bacteria (10^{11}–10^{12} bacteria/g of feces). Anaerobes are the predominant class of microorganism, and Bacteroides species are the most common (10^{11}–10^{12} organisms/mL). Escherichia coli are the most numerous aerobes (10^{8}–10^{10} organisms/mL). Endogenous microflora are crucial for the breakdown of carbohydrates and proteins in the colon and participate in the metabolism of bilirubin, bile acids, estrogen, and cholesterol. Colonic bacteria also are necessary for production of vitamin K. Endogenous bacteria also are thought to suppress the emergence of pathogenic microorganisms, such as Clostridium difficile, a phenomenon termed “colonization resistance.” However, the high bacterial load of the large intestine may contribute to sepsis in critically ill patients and may contribute to intra-abdominal sepsis, abscess, and wound infection following colectomy. Burgeoning research on the human gut microbiome offers new concepts of bacterial community structure and its impact on gastrointestinal disease. Use of 16S ribosomal RNA, universally present in bacteria, allows for a more complete understanding of colonic samples, without the use of stool cultures.

Intestinal gas arises from swallowed air, diffusion from the blood, and intraluminal production. Nitrogen, oxygen, carbon dioxide, hydrogen, and methane are the major components of intestinal gas. Nitrogen and oxygen are largely derived from swallowed air. Carbon dioxide is produced by the reaction of bicarbonate and hydrogen ions and by the digestion of triglycerides to fatty acids. Hydrogen and methane are produced by colonic bacteria. The production of methane is highly variable. The gastrointestinal tract usually contains between 100 and 200 mL of gas, and 400 to 1200 mL/d are released as flatus, depending on the type of food ingested.

Motility, Defecation, and Continence
Motility. Unlike the small intestine, the large intestine does not demonstrate cyclic motor activity characteristic of the migratory motor complex. Instead, the colon displays intermittent
contractions of either low or high amplitude. Low-amplitude, short-duration contractions occur in bursts and appear to move the colonic contents both antegrade and retrograde. It is thought that these bursts of motor activity delay colonic transit and thus increase the time available for absorption of water and exchange of electrolytes. High-amplitude, prolonged duration, propagated contractions (HAPCs) occur in a more coordinated fashion and create “mass movements,” four to ten times per day, mostly after meals and awakening. Bursts of “rectal motor complexes” also have been described. In general, cholinergic activation increases colonic motility.

Defecation. Defecation is a complex, coordinated mechanism involving colonic mass movement, increased intra-abdominal and rectal pressure, and relaxation of the pelvic floor. Distention of the rectum causes a reflex relaxation of the internal anal sphincter (the rectoanal inhibitory reflex) that allows the contents to make contact with the anal canal. This “sampling reflex” allows the sensory epithelium to distinguish solid stool from liquid stool and gas. If defecation does not occur, the rectum relaxes and the urge to defecate passes (accommodation response). Defecation proceeds by coordinating increasing intra-abdominal pressure via aValsalva maneuver with rectal contraction, relaxation of the puborectalis muscle, and opening of the anal canal.

Continence. The maintenance of fecal continence is at least as complex as the mechanism of defecation. Continence requires adequate rectal wall compliance to accommodate the fecal bolus, appropriate neurogenic control of the pelvic floor and sphincter mechanism, and functional internal and external sphincter muscles. At rest, the puborectalis muscle creates a “sling” around the distal rectum, forming a relatively acute angle that distributes intra-abdominal forces onto the pelvic floor. With defecation, this angle straightens, allowing downward force to be applied along the axis of the rectum and anal canal. The internal and external sphincters are tonically active at rest. The internal sphincter is responsible for most of the resting, involuntary sphincter tone (resting pressure). The external sphincter is responsible for most of the voluntary sphincter tone (squeeze pressure). Branches of the pudendal nerve innervate both the internal and external sphincter. The hemorrhoidal cushions may contribute to continence by mechanically blocking the anal canal. Finally, liquid stools exacerbate abnormalities with these anatomic and physiologic mechanisms, so a formed stool contributes to maintaining continence. Thus, impaired continence may result from poor rectal compliance, injury to the internal and/or external sphincter or puborectalis, or neuropathy.

CLINICAL EVALUATION

Clinical Assessment

Obtaining a complete history and performing a physical examination are the starting points for evaluating any patient with suspected disease of the colon, rectum, or anus. Special attention should be paid to the patient’s past medical and surgical history to detect underlying conditions that might contribute to a gastrointestinal problem. If patients have had prior intestinal surgery, it is essential that one understand the resultant gastrointestinal anatomy. A history of anorectal surgery may be critical for patients with either abdominal or anorectal complaints. The obstetrical history in women is essential to detect occult pelvic floor and/or anal sphincter damage. Identifying a family history of colorectal disease, especially inflammatory bowel disease, polyps, and colorectal cancer, is crucial. In addition to a family history of colorectal disease, a history of other malignancies may suggest the presence of a genetic syndrome. Medication use must be detailed as many drugs cause gastrointestinal symptoms. Before recommending operative intervention, the adequacy of medical treatment must be ascertained. In addition to examining the abdomen, visual inspection of the anus and perineum and careful digital rectal exam are essential.

Endoscopy

Anoscopy. The anoscope is a useful instrument for examination of the anal canal. Anoscopes are made in a variety of sizes and measure approximately 8 cm in length. A larger anoscope provides better exposure for anal procedures such as rubber band ligation or sclerotherapy of hemorrhoids. The anoscope, with obturator in place, should be adequately lubricated and gently inserted into the anal canal. The obturator is withdrawn, inspection of the visualized anal canal is done, and the anoscope should then be withdrawn. It is rotated 90° and reinserted to allow visualization of all four quadrants of the canal. If the patient complains of severe perianal pain and cannot tolerate a digital rectal examination, anoscopy should not be attempted without anesthesia.

Proctoscopy. The rigid proctoscope is useful for examination of the rectum and distal sigmoid colon and is occasionally used therapeutically. The standard proctoscope is 25 cm in length and available in various diameters. Most often, a 15- or 19-mm diameter proctoscope is used for diagnostic examinations. A smaller “pediatric” proctoscope (11-mm diameter) is better tolerated by patients with anal stricture. Suction is necessary for an adequate proctoscopic examination. An operating platform for transanal surgery known as transanal endoscopic microsurgery (TEM) has a much wider diameter and can be used for excisions of large polyps and tumors. Transanal minimally invasive surgery (TAMIS) can achieve similar resections to TEM, but it does not utilize a proctoscope and instead depends on insufflation to create a working space in the rectum while utilizing a circular wound protector to open the anus.

Flexible Sigmoidoscopy and Colonoscopy. Video or fiber-optic flexible sigmoidoscopy and colonoscopy provide excellent visualization of the colon and rectum. Sigmoidoscopes measure 60 cm in length. Full depth of insertion may allow visualization as high as the splenic flexure, although the mobility and redundancy of the sigmoid colon often limit the extent of the examination. Partial preparation with enemas is usually adequate for sigmoidoscopy, and most patients can tolerate this procedure without sedation. Colonoscopes measure 100 to 160 cm in length and are capable of examining the entire colon and terminal ileum. A complete oral bowel preparation is usually necessary for colonoscopy, and the duration and discomfort of the procedure usually require conscious sedation. Both sigmoidoscopy and colonoscopy can be used diagnostically and therapeutically. Electrocautery should generally not be used in the absence of a complete bowel preparation because of the risk of explosion of intestinal methane or hydrogen gases. Diagnostic colonoscopes possess a single channel through which instruments such as snares, biopsy forceps, or electrocautery can be passed; this channel also provides suction and irrigation capability. Therapeutic colonoscopes possess two channels to allow simultaneous suction/irrigation and the use of snares, biopsy forceps, or electrocautery.
Capsule Endoscopy. Capsule endoscopy uses a small ingestible camera. After swallowing the camera, images of the mucosa of the gastrointestinal tract are captured, transmitted by radiofrequency to a belt-held receiver, and then downloaded to a computer for viewing and analysis. Capsule endoscopy largely has been used to detect small bowel lesions. Recent advances in the development of maneuverable capsules may improve the sensitivity of this procedure. Finally, concern over the possibility of an acute obstruction has led to the development of a dissolvable capsule that can detect obstruction lesions. Although this technology is promising, the ultimate utility of capsule endoscopy remains unknown.

Imaging

Plain X-Rays and Contrast Studies. Despite advanced radiologic techniques, plain X-rays and contrast studies continue to play an important role in the evaluation of patients with suspected colon and rectal diseases. Plain X-rays of the abdomen (supine, upright, and diaphragmatic views) are useful for detecting free intra-abdominal air, bowel gas patterns suggestive of small or large bowel obstruction, and volvulus. Contrast studies are useful for evaluating obstructive symptoms, delineating fistulous tracts, and diagnosing small perforations or anastomotic leaks. Although Gastrografin cannot provide the mucosal detail provided by barium, this water-soluble contrast agent is recommended if perforation or leak is suspected. Double-contrast barium enema (use of barium followed by the insufflation of air into the colon) has been reported to be 70% to 90% sensitive for the detection of mass lesions greater than 1 cm in diameter. Detection of small lesions can be extremely difficult, especially in a patient with extensive diverticulosis. For this reason, a colonoscopy is preferred for evaluating nonobstructing mass lesions in the colon. Double-contrast barium enema has been used as a back-up examination if colonoscopy is incomplete.

Computed Tomography. Computed tomography (CT) commonly is employed in the evaluation of patients with abdominal complaints. Its utility is primarily in the detection of extraluminal disease, such as intra-abdominal abscesses and pericolic inflammation, and in staging colorectal carcinoma because of its sensitivity in detection of hepatic metastases. Extravasation of oral or rectal contrast may also confirm the diagnosis of perforation or anastomotic leak. Nonspecific findings such as bowel wall thickening or mesenteric stranding may suggest inflammatory bowel disease, enteritis, colitis, or ischemia. A standard CT scan is relatively insensitive for the detection of intraluminal lesions.

Computed Tomography Colonography. CT colonography (virtual colonoscopy) is a radiologic technique that is designed to overcome some of the limitations of traditional CT scanning. This technology uses helical CT and three-dimensional reconstruction to detect intraluminal colonic lesions. Oral bowel preparation, oral and rectal contrast, and colon insufflation have been used to maximize sensitivity. Experience with this technology has shown a sensitivity and specificity for detecting 1 cm or larger polyps of 85% to 90% in most studies, making it comparable to traditional colonoscopy. Although this technology has yet to be widely adopted, it remains an alternative to traditional colonoscopy for select patients.

Magnetic Resonance Imaging. The main use of magnetic resonance imaging (MRI) in colorectal disorders is in evaluation of pelvic lesions. MRI is more sensitive than CT for detecting bony involvement or pelvic sidewall extension of rectal tumors. MRI accurately determines the extent of spread of rectal cancer into the adjacent mesorectum and can reliably predict difficulty achieving radial margin clearance of a rectal cancer by surgery alone. When the radial margin is threatened, neoadjuvant chemoradiation is indicated. MRI also can be helpful in the detection and delineation of complex fistulas in ano. The use of an endorectal coil may increase sensitivity.

Positron Emission Tomography. Positron emission tomography (PET) is used for imaging tissues with high levels of anaerobic glycolysis, such as malignant tumors. Fluorodeoxyglucose (FDG) is injected as a tracer; metabolism of this molecule then results in positron emission. PET has been used as an adjunct to CT in the staging of colorectal cancer and may prove useful in discriminating recurrent cancer from fibrosis. By combining PET and CT technology (PET/CT), anatomic correlation between regions of high isotope accumulation (“hot spots”) on PET and abnormalities on CT can be determined. PET/CT increasingly is used to diagnose recurrent and/or metastatic colorectal cancer. However, the efficacy and utility of this technology remains unproven.

Scintigraphy to Assess Gastrointestinal Bleeding. Scintigraphy to assess gastrointestinal bleeding (technetium-99m-tagged red blood cell [RBC] scan; “tagged RBC scan”) is a nuclear medicine test that uses 99mTc-erythrocytes and dynamic images to localize a bleeding source. Patients must be actively bleeding at the time of imaging, and a normal distribution of 99mTc-erythrocytes in vasculature, liver, spleen, penile circulation with mild uptake in kidneys and bladder can interfere with localization in bowel segments near those structures. Patients must be stable enough to tolerate imaging intervals of up to 4 hours, but slow bleeding at a rate of 0.05 to 0.2 mL/minute can be detected.

Single Photon Emission Computed Tomography (SPECT/CT). Radiolabeled erythrocytes are also used for SPECT/CT, but cross-sectional images provide a more specific localization of the bleeding source, which can be very helpful for surgical planning, especially if direct visualization via endoscopy has not been successful.

Angiography. Angiography is occasionally used for the detection of bleeding within the colon or small bowel. To visualize hemorrhage angiographically, bleeding must be relatively brisk (approximately 0.5 to 1.0 mL per minute). If extravasation of contrast is identified, infusion of vasopressin or angiographic embolization can be therapeutic. If surgical resection is required, the angiographic catheter can be left in place to assist with identification of the bleeding site intraoperatively.

CT and magnetic resonance angiography are also useful for assessing patency of visceral vessels. This technique uses three-dimensional reconstruction to detect vascular lesions. If an abnormality is found, more traditional techniques (angiography, surgery) may then be used to further define and/or correct the problem.

Endorectal and Endoanal Ultrasound. Endorectal ultrasound is primarily used to evaluate the depth of invasion of neoplastic lesions in the rectum. The normal rectal wall appears as a five-layer structure (Fig. 29-6). Ultrasound can reliably differentiate most benign polyps from invasive tumors based on the integrity of the submucosal layer. Ultrasound can also differentiate superficial T1-T2 from deeper T3-T4 tumors. Overall,
Physiologic and Pelvic Floor Investigations

Anorectal physiologic testing uses a variety of techniques to investigate the function of the pelvic floor. These techniques are useful in the evaluation of patients with incontinence, constipation, rectal prolapse, obstructed defecation, and other functional disorders of the pelvic floor.

**Manometry.** Anorectal manometry is performed by placing a pressure-sensitive catheter in the lower rectum. The catheter is then withdrawn through the anal canal and pressures recorded. A balloon attached to the tip of the catheter also can be used to test anorectal sensation. The resting pressure in the anal canal reflects the function of the internal anal sphincter (normal 40–80 mmHg), whereas the squeeze pressure, defined as the maximum voluntary contraction pressure minus the resting pressure, reflects function of the external anal sphincter (normal 40–80 mmHg above resting pressure). The high-pressure zone estimates the length of the anal canal (normal 2.0–4.0 cm). The rectoanal inhibitory reflex can be detected by inflating a balloon in the distal rectum; absence of this reflex is characteristic of Hirschsprung’s disease.

**Neurophysiology.** Neurophysiologic testing assesses function of the pudendal nerves and recruitment of puborectalis muscle fibers. Pudendal nerve terminal motor latency measures the speed of transmission of a nerve impulse through the distal pudendal nerve fibers (normal 1.8–2.2 ms); prolonged latency suggests the presence of neuropathy. Electromyographic (EMG) recruitment assesses the contraction and relaxation of the puborectalis muscle during attempted defecation. Normally, recruitment increases when a patient is instructed to “squeeze” and decreases when a patient is instructed to “push.” Inappropriate recruitment is an indication of paradoxical contraction (nonrelaxation of the puborectalis). Needle EMG has been used to map both the pudendal nerves and the anatomy of the internal and external sphincters. However, this examination is painful and poorly tolerated by most patients. Needle EMG has largely been replaced by pudendal nerve motor latency testing to assess pudendal nerve function and endoanal ultrasound to map the sphincters.

**Rectal Evacuation Studies.** Rectal evacuation studies include the balloon expulsion test and video defecography. Balloon expulsion assesses a patient’s ability to expel an intrarectal balloon. Video defecography provides a more detailed assessment of defecation. In this test, barium paste is placed in the rectum, and defecation is then recorded fluoroscopically. Defecography is used to help diagnose obstructed defecation from nonrelaxation of the puborectalis muscle or anal sphincter dyssynergy, increased perineal descent, rectal prolapse and intussusception, rectocele, and enterocoele. The addition of vaginal contrast and intraperitoneal contrast is useful in delineating complex disorders of the pelvic floor.

**Laboratory Studies**

**Fecal Occult Blood Testing and Fecal Immunohistochemical Testing.** Fecal occult blood testing (FOBT) has been used as a screening test for colonic neoplasms in asymptomatic, average-risk individuals. The efficacy of this test is based on serial testing because the majority of colorectal malignancies will bleed intermittently. FOBT has been a nonspecific test for peroxidase contained in hemoglobin; consequently, occult bleeding from any gastrointestinal source will produce a positive result. Similarly, many foods (red meat, some fruits and vegetables, and vitamin C) will produce a false-positive result. Increased specificity for cancer detection is possible by using fecal immunohistochemical test (FIT). Reported sensitivity of 79% and specificity
of 94% has led to widespread use of FIT in current population-based screening approaches. These tests rely on monoclonal or polyclonal antibodies to react with the intact globin portion of human hemoglobin and are more specific for identifying occult bleeding from the colon or rectum. Any positive FOBT or FIT mandates further investigation, usually by colonoscopy. More recently, stool DNA testing has been proposed for early detection of colorectal cancer.

Stool Studies. Stool studies are often helpful in evaluating the etiology of diarrhea. Wet-mount examination reveals the presence of fecal leukocytes, which may suggest colonic inflammation or the presence of an invasive organism such as invasive E coli or Shigella species. Stool cultures can detect pathogenic bacteria, ova, and/or parasites. C difficile colitis is diagnosed by detecting bacterial toxin in the stool. Steatorrhea may be diagnosed by adding Sudan red stain to a stool sample.

Tumor Markers. Carcinoembryonic antigen (CEA) may be elevated in 60% to 90% of patients with colorectal cancer. Preoperative CEA level has recently been suggested to be a prognostic indicator. Despite this, CEA is not an effective screening tool for this malignancy. Many practitioners follow serial CEA levels after curative-intent surgery in order to detect early recurrence of colorectal cancer. However, this tumor marker is nonspecific, and no survival benefit associated with its serial measurements has yet been proven. It is also important to note that CEA may be mildly elevated in patients who smoke tobacco. Other biochemical markers (ornithine decarboxylase, urokinase) have been proposed, but none has yet proven sensitive or specific for detection, staging, or predicting prognosis of colorectal carcinoma.

Genetic Testing. Although familial colorectal cancer syndromes, such as familial adenomatous polyposis (FAP) and hereditary nonpolyposis colon cancer (HNPCC) are rare, information about the specific genetic abnormalities underlying these disorders has led to significant interest in the role of genetic testing for colorectal cancer.

Tests for mutations in the adenomatous polyposis coli (APC) gene responsible for FAP and in mismatch repair genes responsible for HNPCC are commercially available and extremely accurate in families with known mutations. However, in the absence of an identified mutation, a negative result is uninformative. For individuals from high-risk families without an identified mutation, increased surveillance is recommended. Although many of these mutations are also present in sporadic colorectal cancer, the accuracy of genetic testing in average-risk individuals is considerably lower, and these tests are not recommended for screening. Due to the potential psychosocial implications of genetic testing, it is strongly recommended that professional genetic counselors be involved in the care of any patient considering these tests.

Evaluation of Common Symptoms

Pain

Abdominal Pain Abdominal pain is a nonspecific symptom with myriad causes. Abdominal pain related to the colon and rectum can result from obstruction (either inflammatory or neoplastic), inflammation, perforation, or ischemia. Plain X-rays and judicious use of contrast studies and/or a CT scan can often confirm the diagnosis. Gentle retrograde contrast studies (Gastrografin enema) may be useful in delineating the degree of colonic obstruction. Sigmoidoscopy and/or colonoscopy performed by an experienced endoscopist can assist in the diagnosis of ischemic colitis, infectious colitis, and inflammatory bowel disease. However, if perforation or near complete obstruction is suspected, colonoscopy and/or sigmoidoscopy are generally contraindicated. Evaluation and treatment of abdominal pain from a colorectal source should follow the usual surgical principles of a thorough history and physical examination, appropriate diagnostic tests, resuscitation, and appropriately timed surgical intervention.

Pelvic Pain Pelvic pain can originate from the distal colon and rectum or from adjacent urogenital structures. Tenesmus may result from proctitis or from a rectal or retrorectal mass, or fecal impaction in a constipated patient. Cyclical pain associated with menses, especially when accompanied by rectal bleeding, suggests a diagnosis of endometriosis. Pelvic inflammatory disease also can produce significant abdominal and pelvic pain. The extension of a peridiverticular abscess or periappendiceal abscess into the pelvis may also cause pain. CT scan and/or MRI may be useful in differentiating these diseases. Proctoscopy (if tolerated) also can be helpful. Occasionally, laparoscopy will yield a diagnosis, although with access to high-quality imaging, indications for diagnostic surgery should be rare.

Anorectal Pain. Anorectal pain is most often secondary to an anal fissure, perirectal abscess and/or fistula, or a thrombosed hemorrhoid. Physical examination can usually differentiate these conditions. Other, less common causes of anorectal pain include anal canal neoplasms, perianal skin infection, and dermatologic conditions. Proctalgia fugax results from levator spasm and may present without any other anorectal findings. Physical exam is critical in evaluating patients with anorectal pain. If a patient is too tender to examine in the office, an examination under anesthesia is necessary. MRI or other imaging studies may be helpful in select cases where the etiology of pain is elusive.

Lower Gastrointestinal Bleeding. The first goal in evaluating and treating a patient with gastrointestinal hemorrhage is adequate resuscitation. The second goal is to identify the source of hemorrhage. Because the most common source of gastrointestinal hemorrhage is esophageal, gastric, or duodenal, nasogastric aspiration should always be performed; return of bile suggests that the source of bleeding is distal to the ligament of Treitz. If aspiration reveals blood or nonbile secretions, or if symptoms suggest an upper intestinal source, esophagogastroduodenoscopy should be performed. Anoscopy and/or limited proctoscopy can identify hemorrhoidal bleeding. A technetium-99 (99mTc)-tagged red blood cell (tagged RBC scan) scan is extremely sensitive and is able to detect as little as 0.1 mL/h of bleeding; however, localization is imprecise. If the 99mTc-tagged RBC scan is positive, angiography can then be both diagnostic and potentially therapeutic. If the patient is hemodynamically stable, a rapid bowel preparation (over 4–6 hours) can be performed to allow colonoscopy. Colonoscopy may identify the cause of the bleeding, and cautery or injection/application of epinephrine into the bleeding site may be used to control hemorrhage. A SPECT/CT may be helpful if other modalities have not achieved localization, particularly in distinguishing between small intestinal and colon sources. Colectomy may be required if bleeding persists despite interventions. Intraoperative colonoscopy and/or enteroscopy may assist in localizing bleeding. If colectomy is required, a segmental resection is preferred if the bleeding source can be localized. “Blind” subtotal colectomy
very rarely may be required in a patient who is hemodynamically unstable with ongoing colonic hemorrhage of an unknown source. In this setting, just prior to proceeding with a “blind” subtotal colectomy, it is crucial to irrigate the rectosigmoid and reexamine the mucosa of the anal canal and rectum by anoscopy and proctoscopy to ensure the source of ongoing bleeding is not distal to the planned resection margin (Fig. 29-7).

Occult blood loss from the gastrointestinal tract may manifest as iron-deficiency anemia or may be detected with FOBT or FIT. Because colon neoplasms bleed intermittently and rarely present with rapid hemorrhage, the presence of occult fecal blood should always prompt a colonoscopy. Unexplained iron-deficiency anemia is also an indication for colonoscopy.

Hematochezia is commonly caused by hemorrhoids or a fissure. Sharp, knife-like pain and bright red rectal bleeding with bowel movements suggest the diagnosis of fissure. Painless, bright red rectal bleeding with bowel movements is often secondary to a friable internal hemorrhoid that is easily detected by anoscopy. In the absence of a painful, obvious fissure, any patient with rectal bleeding should undergo a careful digital rectal examination, anoscopy, and proctosigmoidoscopy. Failure to diagnose a source in the distal anorectum should prompt colonoscopy.

**Constipation and Obstructed Defecation.** Constipation is an extremely common complaint, affecting more than 4 million people in the United States. Despite the prevalence of this problem, there is lack of agreement about an appropriate definition of constipation. Patients may describe infrequent bowel movements, hard stools, or excessive straining. A careful history of these symptoms often clarifies the nature of the problem.

Constipation has many causes. Underlying metabolic, pharmacologic, endocrine, psychological, and neurologic causes often contribute to the problem. A stricture or mass lesion should be excluded by colonoscopy, barium enema, or CT colonography. After these causes have been excluded, evaluation focuses on differentiating slow-transit constipation from outlet obstruction. Transit studies, in which radiopaque markers are swallowed and followed radiographically, are useful for diagnosing slow-transit constipation. In this study, patients ingest radiopaque markers and are followed radiographically for 5 days. Retention of 20% or greater of these markers in the colon demonstrated slow transit. If these markers are congregated in the rectosigmoid colon and rectum, obstructed defecation/outlet obstruction is suggested. Anorectal manometry and EMG can detect nonrelaxation of the puborectalis, which contributes to outlet obstruction. The absence of an anorectal inhibitory reflex
suggests Hirschsprung’s disease and may prompt a rectal mucosal biopsy. Defecography can identify rectal prolapse, intussusception, rectocele, or enterocele.

Medical management is the mainstay of therapy for constipation and includes fiber, increased fluid intake, and laxatives. Outlet obstruction from nonrelaxation of the puborectalis or anal sphincter dyssynergy often responds to biofeedback. Surgery to correct rectocele and rectal prolapse (with or without sigmoid resection) has a variable effect on symptoms of constipation but can be successful in selected patients. Subtotal colectomy is considered only for patients with severe slow-transit constipation (colonic inertia) refractory to maximal medical interventions. While this operation almost always increases bowel movement frequency, complaints of diarrhea, incontinence, and abdominal pain are not infrequent, and patients should be carefully selected and counseled.11

Diarrhea and Irritable Bowel Syndrome. Diarrhea is also a common complaint and is usually a self-limited symptom of infectious gastroenteritis. If diarrhea is chronic or is accompanied by bleeding or abdominal pain, further investigation is warranted. Bloody diarrhea and pain are characteristic of colitis; etiology can be an infection (invasive E coli, Shigella, Salmonella, Campylobacter, Entamoeba histolytica, or C difficile), inflammatory bowel disease (ulcerative colitis or Crohn’s colitis), or ischemia. Stool wet-mount and culture can often diagnose infection. Sigmoidoscopy or colonoscopy can be helpful in diagnosing inflammatory bowel disease or ischemia. However, if the patient has abdominal tenderness, particularly with peritoneal signs, or any other evidence of perforation, endoscopy is contraindicated.

Chronic diarrhea may present a more difficult diagnostic dilemma. Chronic ulcerative colitis, Crohn’s colitis, infection, malabsorption, and short gut syndrome can cause chronic diarrhea. Rarely, carcinoid syndrome and islet cell tumors (vasoactive intestinal peptide–secreting tumor [VIPoma], somatostatinoma, gastrinoma) present with this symptom. Large villous lesions may cause secretory diarrhea. Collagenous colitis can cause diarrhea without any obvious mucosal abnormality. Along with stool cultures, tests for malabsorption, and metabolic investigations, colonoscopy can be invaluable in differentiating these causes. Biopsies should be taken even if the colonic mucosa appears grossly normal.

Irritable bowel syndrome is a particularly troubling constellation of symptoms consisting of crampy abdominal pain, bloating, constipation, and urgent diarrhea. Workup reveals no underlying anatomic or physiologic abnormality. Once other disorders have been excluded, dietary restrictions and avoidance of caffeine, alcohol, and tobacco may help to alleviate symptoms. Antispasmodics and bulking agents may be helpful.

Incontinence. The true incidence of fecal incontinence is unknown, but has been estimated to occur in 10 to 13 individuals per 1000 people older than age 65 years. Incontinence ranges in severity from occasional leakage of gas and/or liquid stool to daily loss of solid stool. The underlying cause of incontinence is often multifactorial, and diarrhea is often contributory. In general, causes of incontinence can be classified as neurogenic or anatomic. Neurogenic causes include diseases of the central nervous system and spinal cord along with pudendal nerve injury. Anatomic causes include congenital abnormalities, procidentia (rectal prolapse), overflow incontinence secondary to impaction or an obstructing neoplasm, and trauma. The most common traumatic cause of incontinence is injury to the anal sphincter during vaginal birth. Other causes include anorectal surgery, impalement, and pelvic fracture.

After a thorough medical evaluation to detect underlying conditions that might contribute to incontinence, evaluation focuses on assessment of the anal sphincter and pudendal nerves. Pudendal nerve terminal motor latency testing may detect neuropathy. Anal manometry can detect low resting and squeeze pressures. Physical examination and defecography can detect rectal prolapse. Endoanal ultrasound is invaluable in diagnosing sphincter defects (Fig. 29-8).

Therapy depends on the underlying abnormality. Diarrhea should be treated medically (fiber, antidiarrheal agents). Even in the absence of frank diarrhea, the addition of dietary fiber may improve continence. Some patients may respond to biofeedback and this approach may be considered in patients who fail dietary modification. Many patients with a sphincter defect are candidates for an overlapping sphincteroplasty. Sacral nerve
stimulation been shown to decrease episodes of fecal incontinence and has proven durability in the long term (5 years). The artificial bowel sphincter may be useful in patients who fail other interventions. Other options include radiofrequency energy to the anal canal (SECCA procedure), magnetic anal sphincter, and injectable submucosal bulking agents, but long-term efficacy has not yet been proven. Finally, a stoma can provide relief for severely incontinent patients who have failed or are not candidates for other interventions.

GENERAL SURGICAL CONSIDERATIONS

Colorectal resections are performed for a wide variety of conditions, including neoplasms (benign and malignant), inflammatory bowel diseases, and other benign conditions. Although the indication and urgency for surgery will alter some of the technical details, the operative principles of colorectal resections, anastomoses, and use of ostomies are well established.

Resections

The mesenteric clearance technique dictates the extent of colonic resection and is determined by the nature of the primary pathology (malignant or benign), the intent of the resection (curative or palliative), the precise location(s) of the primary pathology, and the condition of the mesentery (thin and soft or thick and indurated). In general, a proximal mesenteric ligation will eliminate the blood supply to a greater length of colon and require a more extensive “colectomy.” Curative resection of a colorectal cancer is usually best accomplished by performing a proximal mesenteric vessel ligation and radical mesenteric clearance of the lymphatic drainage basin of the tumor site (Fig. 29-9). Resection of a benign process does not require wide mesenteric clearance.

Emergency Resection. Emergency resection may be required because of obstruction, perforation, or hemorrhage. In this setting, the bowel is almost always unprepared and the patient may be unstable. The surgical principles described earlier apply, and an attempt should be made to resect the involved segment along with its lymphovascular supply. If the resection involves the right colon or proximal transverse colon (right or extended right colectomy), a primary ileocolic anastomosis can usually be performed safely as long as the remaining bowel appears healthy and the patient is stable. For left-sided tumors, the traditional approach has involved resection of the involved bowel and end colostomy, with or without a mucus fistula. However, there is an increasing body of data to suggest that a primary anastomosis without a bowel preparation or with an on-table lavage, with or without a diverting ileostomy, may be equally safe in this setting. If the proximal colon appears unhealthy (vascular compromise, serosal tears, perforation), a subtotal colectomy can be performed with a small bowel to rectosigmoid anastomosis. Resection and diversion (ileostomy or colostomy) remain safe and appropriate if the bowel appears compromised or if the patient is unstable, malnourished, or immunosuppressed.

Minimally Invasive Techniques of Resection. With advances in minimally invasive technology, many procedures that previously have required laparotomy can now be performed laparoscopically, with hand-assisted laparoscopy (HAL), or robotically. Potential advantages of minimally invasive surgery include improved cosmetic result, decreased postoperative pain, and earlier return of bowel function. Moreover, some experimental data suggest that minimally invasive operations have less immunosuppressive impact on the patient and thus might improve postoperative outcome and even long-term survival. To date, most studies have demonstrated equivalence between laparoscopic, HAL, and open resection in terms of extent of resection. Return of bowel function and length of hospital stay are highly variable. Long-term outcome has yet to be determined; however, short-term quality of life appears to be improved by laparoscopy. Laparoscopic total mesorectal excision for rectal cancer, however, may not be appropriate. The most recent advances in minimally invasive surgery involve use of robotics and telesmanipulation in which the surgeon operates from a console remote from the patient. These procedures have been rapidly gaining in popularity, especially for pelvic and rectal resections. Early studies suggest equivalence between robotic resections and laparoscopic/HAL resections.

In addition, some proponents have suggested that robotic procedures may be easier to learn (a shorter “learning curve”) and that robotic surgery may be ergonomically better for the operating surgeon. Nevertheless, long-term superiority, or even equivalence, has yet to be demonstrated, and these advanced technologies are likely to be associated with significant cost.

Colectomy. A variety of terms are used to describe different types of colectomy (Fig. 29-10).

Ileocolic Resection An ileocolic resection describes a limited resection of the terminal ileum, cecum, and appendix. It is used to remove disease involving these segments of the intestine (e.g., ileocecal Crohn’s disease) and benign lesions or incurable cancers arising in the terminal ileum, cecum, and, occasionally, the appendix. If curable malignancy is suspected, more radical resections, such as a right hemicolec-tomy, are generally indicated. The ileocolic vessels are ligated and divided. A variable length of small intestine may be resected depending on the disease process. A primary anastomosis is created between the distal small bowel and the ascending colon. It is technically difficult to perform an anastomosis at or just proximal to the ileocecal valve; therefore, if the most distal ileum needs to be resected, the cecum is generally also removed.

Right Colectomy A right colectomy is used to remove lesions or disease in the right colon and is oncologically the most appropriate operation for curative intent resection of proximal colon carcinoma. The ileocolic vessels, right colic vessels, and right branches of the middle colic vessels are ligated and divided. Approximately 10 cm of terminal ileum are usually included in the resection. A primary ileal-transverse colon anastomosis is almost always possible.

Extended Right Colectomy An extended right colectomy may be used for curative intent resection of lesions located at the hepatic flexure or proximal transverse colon. A standard right colectomy is extended to include ligation of the middle colic vessels at their base. The right colon and proximal transverse colon are resected, and a primary anastomosis is created between the distal ileum and distal transverse colon. Such an anastomosis relies on the marginal artery of Drummond. If the blood supply to the distal transverse colon is questionable, the resection is extended distally beyond the splenic flexure to well-perfused descending colon where the ileocolic anastomosis can be performed safely.

Transverse Colectomy Lesions in the mid and distal transverse colon may be resected by ligating the middle colic vessels and resecting the transverse colon, followed by a colocolonic anastomosis. However, an extended right colectomy with an
anastomosis between the terminal ileum and descending colon may be a safer anastomosis with an equivalent functional result.

**Left Colectomy** For lesions or disease states confined to the distal transverse colon, splenic flexure, or descending colon, a left colectomy is performed. The left branches of the middle colic vessels, the left colic vessels, and the first branches of the sigmoid vessels are ligated. A colocolonic anastomosis can usually be performed.

**Extended Left Colectomy** An extended left colectomy is an option for removing lesions in the distal transverse colon. In this operation, the left colectomy is extended proximally to include the right branches of the middle colic vessels.

**Sigmoid Colectomy** Lesions in the sigmoid colon require ligation and division of the sigmoid branches of the inferior mesenteric artery. In general, the entire sigmoid colon should be resected to the level of the peritoneal reflection and an

---

anastomosis created between the descending colon and upper rectum. Full mobilization of the splenic flexure is often required to create a tension-free anastomosis.

**Total and Subtotal Colectomy**  
Total or subtotal colectomy is occasionally required for patients with fulminant colitis, attenuated FAP, or synchronous colon carcinomas. In this procedure, the ileocolic vessels, right colic vessels, middle colic vessels, and left colic vessels are ligated and divided. The superior rectal vessels are preserved. If it is desired to preserve the sigmoid, the distal sigmoid vessels are left intact, and an anastomosis is created between the ileum and distal sigmoid colon (subtotal colectomy with ileosigmoid anastomosis). If the sigmoid is to be resected, the sigmoidal vessels are ligated and divided, and the ileum is anastomosed to the upper rectum (total abdominal colectomy with ileorectal anastomosis). If an anastomosis is contraindicated, an end ileostomy is created, and the remaining sigmoid or rectum is managed either as a mucus fistula or a Hartmann’s pouch.

**Proctocolectomy**

**Total Proctocolectomy**  
In this procedure, the entire colon, rectum, and anus are removed, and the ileum is brought to the skin as a Brooke ileostomy.

**Restorative Proctocolectomy (Ileal Pouch–Anal Anastomosis)**  
The entire colon and rectum are resected, but the anal sphincter muscles and a variable portion of the distal anal canal are preserved. Bowel continuity is restored by anastomosis of an ileal reservoir to the anal canal. The original technique included a transanal mucosectomy and hand-sewn ileoanal anastomosis. Proponents of this technique argue that mucosectomy guarantees removal of all of the diseased mucosa, including the anal transition zone, and therefore decreases the risk of ongoing disease, dysplasia, and carcinoma. Opponents cite the increased risk of incontinence after mucosectomy and argue that even meticulous technique invariably leaves behind mucosal “islands” that are subsequently hidden under the anastomosis. However, persistent or recurrent dysplasia in the anal transition zone is uncommon (4.5%), and cancers occur even more rarely. Moreover, the “double-staple” technique using the circular stapling devices is considerably simpler than mucosectomy and a hand-sewn anastomosis and may be associated with a better functional outcome (Fig. 29-11). Regardless of the anastomotic technique, many surgeons recommend that patients undergo annual surveillance of the anastomosis and/or anal transition zone by digital rectal exam and anoscopy or proctoscopy.

The neorectum is made by anastomosis of the terminal ileum aligned in a “J,” “S,” or “W” configuration. Because functional outcomes are similar and because the J-pouch is the simplest to construct, it has become the most used configuration.

![Figure 29-10. Terminology of types of colorectal resections: A→C Ileocecectomy; + A + B→D Ascending colectomy; + A + B→F Right hemicolecotomy; + A + B→G Extended right hemicolecotomy; + E + F→G + H Transverse colectomy; G→I Left hemicolecotomy; F→J Extended left hemicolecotomy; J + K Sigmoid colectomy; + A + B→J Subtotal colectomy; + A + B→K Total colectomy; + A + B→L Total proctocolectomy. (Reproduced with permission from Fielding LP, Goldberg SM: Rob & Smith’s Operative Surgery of the Colon, Rectum, and Anus. London: Elsevier; 1993.)](image1)

![Figure 29-11. After a total colectomy and resection of the rectum (A), the anal canal with a short cuff of transitional mucosa and sphincter muscles is preserved (B). An ileal J-pouch has been constructed and is anastomosed to the anal canal using a double-staple technique (C). (Reproduced with permission from Bell RH, Rikers LF, Mulholland M: Digestive Tract Surgery: A Text and Atlas. Philadelphia, PA: Lippincott Williams & Wilkins; 1996.)](image2)
Colonic Resection

Anterior Resection. Anterior resection is the general term used to describe resection of the rectum from an abdominal approach to the pelvis with no need for a perineal, sacral, or other incision. Three types of anterior resection have been described.

High Anterior Resection A high anterior resection is the term used to describe resection of the distal sigmoid colon and upper rectum and is the appropriate operation for benign lesions and disease at the rectosigmoid junction such as diverticulitis. The upper rectum is mobilized, but the pelvic peritoneum is not divided and the rectum is not mobilized fully from the concavity of the sacrum. The inferior mesenteric artery is ligated at its base, and the inferior mesenteric vein, which follows a different course than the artery, is ligated separately. A primary anastomosis (usually end-to-end) between the colon and rectal stump with a short cuff of peritoneum surrounding its anterior two thirds generally can be performed.

Low Anterior Resection. A low anterior resection is used to remove lesions in the upper and mid rectum. The rectosigmoid is mobilized, the pelvic peritoneum is opened, and the inferior mesenteric artery is ligated and divided either at its origin from the aorta or just distal to the takeoff of the left colic artery. The rectum is mobilized from the sacrum by sharp dissection under direct view within the endopelvic fascial plane. The dissection may be performed distally to the anorectal ring, extending posteriorly through the rectosacral fascia to the coccyx and anteriorly through Denonvilliers’ fascia to the vagina in women or the seminal vesicles and prostate in men. The rectum and accompanying mesorectum are divided at the appropriate level, depending on the nature of the lesion. A low rectal anastomosis usually requires mobilization of the splenic flexure and ligation and division of the inferior mesenteric vein just inferior to the pancreas. Circular stapling devices have greatly facilitated the conduct and improved the safety of the colo-anal transitional zone.

Extended Low Anterior Resection An extended low anterior resection is necessary to remove lesions located in the distal rectum, but several centimeters above the sphincter. The rectum is fully mobilized to the level of the levator ani muscle just as for a low anterior resection, but the anterior dissection is extended along the rectovaginal septum in women and distal to the seminal vesicles and prostate in men. After resection at this level, a coloanal anastomosis can be created using one of a variety of techniques. An end-to-end stapled or hand-sewn anastomosis has traditionally been the procedure of choice. However, the functional consequences of a “straight” anastomosis have led to consideration for creation of a colon J-pouch to increase the capacity of the neorectal reservoir. A side-to-end anastomosis can be constructed by placing the anvil of an EEA stapler 3 to 4 cm away from the stapled end of the proximal colon, with similar functional outcomes to the colon J-pouch reconstruction.

Hartmann’s Procedure and Mucus Fistula. Hartmann’s procedure refers to a colon or rectal resection without an...
Abdominoperineal Resection. An abdominoperineal resection (APR) involves removal of the entire rectum, anal canal, and anus with construction of a permanent colostomy from the descending or sigmoid colon. The abdominal-pelvic portion of this operation proceeds in the same fashion as described for an extended low anterior resection. The perineal dissection can be performed with the patient in lithotomy position (often by a second surgeon) or in the prone position after closure of the abdomen and creation of the colostomy. For cancer, the perineal dissection is designed to excise the anal canal with a wide circumferential margin including a cylindrical cuff of the levator muscle. A deliberate resection of the levator muscles near their bony attachments, in order to avoid opening the space between the tumor and the levator ani, is known as the extralevator abdominoperineal resection (ELAPE). ELAPE is useful for low, locally advanced rectal cancers, but routine use for all rectal cancer has not been shown to improve cancer outcomes. Primary wound closure is usually successful, but a large perineal defect, especially if preoperative radiation has been used, may require a vascularized flap closure in some patients. For benign disease, proctectomy may be performed using an intersphincteric dissection between the internal and external sphincters. This approach minimizes the perineal wound, making it easier to close because the levator muscle remains intact.

Anastomoses
Anastomoses may be created between two segments of bowel in a multitude of ways. The geometry of the anastomosis may be end-to-end, end-to-side, side-to-end, or side-to-side. The anastomatic technique may be hand-sutured or stapled. The submucosal layer of the intestine provides the strength of the bowel wall and must be incorporated in the anastomosis to assure healing. The choice of anastomosis depends on the operative anatomy and surgeon preference. Although many surgeons advocate one method over another, none has been proven to be superior. Accurate approximation of two well-vascularized, healthy limbs of bowel without tension in a normotensive, well-nourished patient almost always results in a good outcome. Anastomoses at highest risk of leak or stricture are those that are in the distal rectal or anal canal, involve irradiated or diseased intestine including perforation with peritoneal soilage, are inadvertently fashioned above a partial distal obstruction, or are performed in malnourished, immunosuppressed, or ill patient.

Anastomotic Configuration
End-to-End An end-to-end anastomosis can be performed when two segments of bowel are roughly the same caliber. This technique is most often employed in rectal resections, but may be used for colocolostomy or small bowel anastomoses.

End-to-Side An end-to-side configuration is useful when one limb of bowel is larger than the other. This most commonly occurs in the setting of chronic obstruction.

Side-to-End A side-to-end anastomosis is used when the proximal bowel is of smaller caliber than the distal bowel. Ileorectal anastomoses commonly make use of this configuration. A side-to-end anastomosis may have a less tenuous blood supply than an end-to-end anastomosis.

Side-to-Side A side-to-side anastomosis allows a large, well-vascularized connection to be created on the antimesenteric side of two segments of intestine. This technique is commonly used in ileocolic and small bowel anastomoses.

Anastomotic Technique
Hand-Sutured Technique Any of the configurations described earlier may be created using a hand-sutured or stapled technique. Hand-sutured anastomoses may be single layer, using either running or interrupted stitches, or double layer. A double-layer anastomosis usually consists of a continuous inner layer and an interrupted outer layer. Suture material may be either permanent or absorbable. After distal rectal or anal canal resection, a transanal, hand-sewn coloanal anastomosis may be necessary to restore bowel continuity. This can be done in conjunction with an anal canal mucosectomy to allow the anastomosis to be created at the dentate line.

Stapled Techniques Linear cutting/stapling devices are used to divide the bowel and to create side-to-side anastomoses. The anastomosis may be reinforced with interrupted sutures if desired. Circular cutting/stapling devices can create end-to-end, end-to-side, or side-to-end anastomoses. These instruments are particularly useful for creating low rectal or anal canal anastomoses where the anatomy of the pelvis makes a hand-sewn anastomosis technically difficult or impossible.

Following resection of the colorectum, a stapled end-to-end colorectal, coloanal, or ileoanal anastomosis may be created by one of two techniques. With the open purse-string technique, the distal rectal stump purse-string is placed by hand, and the assembled circular stapler is inserted into the anus and guided up to the rectal purse-string. The stapler is opened, and the distal purse-string is tied. A purse-string is placed in the distal end of the proximal colon; the proximal colon is placed over the anvil and the purse-string tightened. The stapler is closed and fired (Fig. 29-13). With the alternative double-staple technique, the distal rectum or anal canal is closed with a transverse staple line. The circular stapler is inserted through the anus without its anvil until the cartridge effaces the transverse staple line. The stapler is opened, causing the trocar to perforate through the rectal stump adjacent to the transverse staple line. The anastomosis is then completed as described earlier (see Fig. 29-11). If the stapler cannot advance to the end of the rectal stump, further dissection of the stump may be necessary to optimize tissue apposition. After firing and removing the stapler, the resulting anastomotic rings should be inspected to ensure that they are full-thickness and concentric, and in cases of rectal cancer, the distal anastomotic ring should be sent to pathology as a specimen (true distal margin). A gap in an anastomotic ring suggests that the circular staple line is incomplete and the anastomosis should be reinforced with suture circumferentially, if technically feasible. A temporary proximal ileostomy may be indicated as well. Most surgeons will also leak test an anastomosis by instilling water or saline into the pelvis and insufflating the rectum with air via a proctoscope to look or alternatively instilling methylene blue or betadine into the rectum to look for extravasation. A leak test strongly suggests a defect and/or disruption of the anastomosis. As such, the suture should line reinforced and sometimes may require reanastomosis.
Ostomies and Preoperative Stoma Planning

Depending on the clinical situation, a stoma may be temporary or permanent. It may be end-on or a loop. However, regardless of the indication for a stoma, placement and construction are crucial for function.

The preoperative preparation of a patient who is expected to require a stoma should include a consultation with an enterostomal therapy (ET) nurse. ET nurses are specially trained and credentialed by the Wound, Ostomy, and Continence Nurses Society. Preoperative planning includes counseling, education, and stoma siting. Postoperatively, the ET nurse assists with local skin care and pouching. Other considerations in stoma planning include evaluation of other medical conditions that may impact on a patient’s ability to manage a stoma (e.g., eyesight, manual dexterity). Family or other caregivers can prove invaluable in caring for these patients.

Preoperative stoma siting is crucial for a patient’s postoperative function and quality of life. A poorly placed stoma can result in leakage and skin breakdown. Ideally, a stoma should be placed in a location that the patient can easily see and manipulate, within the rectus muscle, and below the belt line (Fig. 29-15). Because the abdominal landmarks in a supine, anesthetized patient may be dramatically different from those in an awake, standing, or sitting patient, the stoma site should always be marked with a tattoo, skin scratch, or permanent marker preoperatively, if possible. In an emergency operation where the stoma site has not been marked, an attempt should be made to place a stoma within the rectus muscle and away from

both the costal margin and iliac crest. In emergencies, placement high on the abdominal wall is preferred to a more caudal site.

For all stomas, a circular skin incision is created, and the subcutaneous tissue is dissected to the level of the anterior rectus sheath. The anterior rectus sheath is incised in a cruciate fashion, the muscle fibers separated bluntly, and the posterior sheath identified and incised. Care should be taken to avoid injuring and causing bleeding from the inferior epigastric artery and vein. The size of the defect depends on the size of the bowel used to create the stoma, but it should be as small as possible without compromising the intestinal blood supply (usually the width of two to three fingers). The bowel is then brought through the defect and secured with sutures. The abdominal incision is usually closed and dressed prior to maturing the stoma to avoid contaminating the wound. In order to make appliance use easier, a protruding nipple is fashioned by everting the bowel. Three or four interrupted absorbable sutures are placed through the edge of the bowel, then through the serosa, approximately 2 cm proximal to the edge, and then through the dermis (Brooke technique). After the stoma is everted, the mucocutaneous junction is sutured circumferentially with interrupted absorbable suture (Fig. 29-16).

**Ileostomy**

**Temporary Ileostomy** A temporary ileostomy is often used to “protect” an anastomosis that is at risk for leakage (low in the rectum, in an irradiated field, in an immunocompromised or malnourished patient, and during some emergency operations). In this setting, the stoma is often constructed as a loop ileostomy (see Fig. 29-12). A segment of distal ileum is brought through the defect in the abdominal wall as a loop. An enterotomy is created and the stoma matured as described earlier. The loop may be secured with or without an underlying rod. A divided loop may also be created by firing a linear cutting/stapler across the distal limb of the loop flush with the skin followed by maturation of the proximal limb of the loop. This technique prevents incomplete diversion that occasionally occurs with a loop ileostomy.

The advantage of a loop or divided loop ileostomy is that subsequent closure can often be accomplished without a formal laparotomy. An elliptical incision is created around the stoma and the bowel gently dissected free of the subcutaneous tissues and fascia. A hand-sewn or stapled anastomosis can then be created and the intestine returned to the peritoneal cavity. This

---

**Figure 29-14.** Technique of end-to-end colorectal anastomosis using a circular stapler. **A.** The patient is in modified lithotomy position. **B.** After resection of the rectosigmoid and placement of purse-string sutures proximally and distally, the stapler is inserted into the anal canal and opened. **C.** Rectal purse-string suture is tied to secure the rectal stump to the rod of the stapler, and the colonic purse-string is tied to secure the colon to the anvil of the stapler. **D.** The stapler is closed and fired. **E.** The stapler is removed, leaving a circular stapled end-to-end anastomosis.
avoids a long laparotomy incision and generally is well tolerated. The timing of ileostomy closure should take into account anastomotic healing as well as the patient’s overall condition. A flexible endoscopy exam and a contrast enema (Gastrografin) are recommended prior to closure to ensure that the anastomosis has not leaked and is patent. A patient’s nutritional status should be optimized. Because the timing of adjuvant chemotherapy affects survival, cancer patients receiving adjuvant chemotherapy usually should defer ileostomy closure until completion of treatment.

Permanent Ileostomy

A permanent ileostomy is sometimes required after total proctocolectomy or in patients with obstruction. An end ileostomy is the preferred configuration for a permanent ileostomy because a symmetric protruding nipple can be fashioned more easily than with a loop ileostomy (see Fig. 29-16). The end of the small intestine is brought through the abdominal wall defect and matured. Stitches are often used to secure the bowel to the posterior fascia.

Complications of Ileostomy

Stoma necrosis may occur in the early postoperative period and is usually caused by skeletonizing the distal small bowel and/or creating an overly tight fascial defect. Limited mucosal necrosis above the fascia may be treated expectantly, but necrosis below the level of the fascia requires surgical revision. Stoma retraction may occur early or late and may be exacerbated by obesity. Local revision may be necessary. The creation of an ileostomy bypasses the fluid-absorbing capability of the colon, and dehydration with fluid and electrolyte abnormalities is not uncommon. Ideally, ileostomy output should be maintained at less than 1500 mL/d to avoid this problem. Bulk agents and opioids (Lomotil, Imodium, tincture of opium) are useful. The somatostatin analogue, octreotide, has been used with varying success in this setting. Skin irritation can also occur, especially if the stoma appliance fits poorly. Skin-protecting agents and custom pouches can help to solve this problem. Obstruction may occur intra-abdominally or at the site where the stoma exits the fascia. Parastomal hernia is less common after an ileostomy than after a colostomy but can cause poor appliance fitting, pain, obstruction, or strangulation. In general, symptomatic parastomal hernias should be repaired. A variety of techniques to repair these hernias have been described, including local repair (either with or without mesh), laparoscopic repair, and stoma resiting. Prolapse is a rare, late complication and is often associated with a parastomal hernia.

Colostomy. Most colostomies are created as end colostomies rather than loop colostomies (Fig. 29-17). The bulkiness of the colon makes a loop colostomy awkward for use of an appliance, and prolapse is more likely with this configuration. Most colostomies are created on the left side of the colon. An abdominal wall defect is created and the end of the colon mobilized through it. Because a protruding stoma is considerably easier to pouch, colostomies should also be matured in a Brooke fashion. The distal bowel may be brought through the abdominal wall as a mucus fistula or left intra-abdominally as a Hartmann’s pouch. Tacking the distal end of the colon to the abdominal wall or tagging it with permanent suture can make identification of the stump easier if the colostomy is closed at a later date. Closure of an end colostomy has traditionally required a laparotomy, but increasingly minimally invasive techniques have been adopted. The stoma is dissected free of the abdominal wall and the distal bowel identified. An end-to-end anastomosis is then created.
Complications of Colostomy  

Colostomy necrosis may occur in the early postoperative period and results from an impaired vascular supply (skeletonization of the distal colon or a tight fascial defect). Like ileostomy necrosis, limited suprafascial necrosis may be followed expectantly, but necrosis below the fascia requires surgery. Retraction may also occur but is less problematic with a colostomy than with an ileostomy because the stool is less irritating to the skin than succus entericus. Obstruction is unusual, but may also occur. Parastomal hernia is the most common late complication of a colostomy and requires repair if it is symptomatic. Prolapse occurs rarely, but is more common with a loop colostomy. Interestingly, it is almost always the effenter limb of the loop that prolapses. Dehydration is rare after colostomy, and skin irritation is less common than with ileostomy.

Functional Results

Function following segmental colonic resection and primary anastomosis is generally excellent. A small percentage of patients following subtotal or total colectomy and ileosigmoid or ileorectal anastomosis may experience diarrhea and bowel frequency. This is especially true if the patient is elderly, if significant length of small bowel has been resected, and if residual proctocolitis is poorly controlled. In general, the more distal the anastomosis, the greater is the risk of troublesome diarrhea and frequency. However, some patients develop significant diarrhea after right colectomy due to malabsorption of bile acids; in these cases, bile acid binding resins (e.g., cholestyramine) sometimes can be helpful.

Function following anterior resection is highly dependent on the level of anastomosis, the use of pre- or postoperative pelvic radiation, and underlying sphincter function. Following low anterior or extended low anterior resection, some surgeons prefer to construct a short (5-cm) colon J-pouch to anastomose to the distal rectum or anal canal in order to increase the capacity of the neorectum. The reservoirs are thought to lessen urgency, frequency, and incontinence, but some patients have difficulty initiating defecation, and long-term superiority over a “straight” anastomosis has yet to be proven. In addition, these reservoirs can be technically difficult, especially in an obese male with a narrow pelvis. 

Figure 29-17. Intraperitoneal end colostomy.

Anesthesia Considerations

Local Anesthesia. Many anorectal procedures can be performed with local anesthetic alone. Intravenous sedation is often provided to calm the patient. Injection of 0.5% lidocaine (short acting) and 0.25% bupivacaine (long acting) into the perianal skin, sphincter, and area around the pudendal nerves usually provides an adequate block. The addition of dilute epinephrine decreases bleeding and prolongs the anesthetic effect.

Regional Anesthesia. Epidural, spinal, and caudal anesthetics can be used for anorectal procedures and transanal resections. In patients with severe medical comorbidity, regional anesthesia may occasionally be used for laparotomy and colectomy. Postoperative epidural anesthesia provides excellent pain relief and improves pulmonary function especially after an open operation.

General Anesthesia. General anesthesia is required for the vast majority of intra-abdominal procedures. Patients should
undergo a thorough preoperative cardiovascular evaluation. In patients with significant comorbid disease, an anesthesia consultation may be appropriate.

**Positioning.** Most abdominal colectomies can be performed in the supine position. Anterior resection and APR require lithotomy positioning to facilitate the pelvic dissection and mobilization of the splenic flexure. Adequate padding should be provided for the patient’s sacrum, and care should be taken to avoid stirrup pressure on the peroneal nerves.

Anorectal procedures may be performed in lithotomy or in the prone jackknife position. Some surgeons prefer the prone jackknife position because exposure may be better, especially for anterior lesions. Distal posterior lesions can usually be accessed from either position, but more proximal posterior lesions are better accessed in the prone position.

**Operative Preliminaries**

**Bowel Preparation.** The rationale for bowel preparation is that decreasing the bacterial load in the colon and rectum will decrease the incidence of postoperative infection. *Mechanical bowel preparation* uses cathartics to rid the colon of solid stool the night before surgery. The most commonly used regimens include polyethylene glycol (PEG) solutions or magnesium citrate. PEG solutions require patients to drink a large volume of fluid and may cause bloating and nausea. Magnesium citrate solutions are generally better tolerated but are more likely to cause fluid and electrolyte abnormalities. Both are equally efficacious in bowel cleansing. Preparatory formulations have been recently introduced in tablet form in an attempt to improve tolerance. However, these methods of bowel cleansing require ingestion of 40 or more tablets with water over several hours. To date, these formulations have not been proven to be superior to the more traditional products. *Antibiotic prophylaxis* also is recommended. The addition of oral antibiotics to the preoperative mechanical bowel preparation has been thought to decrease postoperative infection by further decreasing the bacterial load of the colon. A recent analysis of the Surgical Care Improvement Project-1 (SCIP-1) suggests that oral antibiotics reduce postoperative wound infection, especially if a mechanical bowel preparation is not used.

Longstanding, convincing data support the efficacy of parenteral antibiotic prophylaxis at the time of surgery. Broad-spectrum parenteral antibiotic(s) with activity against aerobic and anaerobic enteric pathogens should be administered just prior to the skin incision and redosed as needed depending on the length of the operation. There is no proven benefit to using antibiotics postoperatively after an uncomplicated colectomy.

Despite widespread use of mechanical bowel preparation, the necessity of bowel cleansing prior to colectomy has been questioned. European surgeons in particular have advocated abandoning this practice. Arguments against mechanical bowel preparation include dehydration and electrolyte abnormalities that often result from bowel cleansing, as well as the risk of spillage of liquid stool left over from the “prep.” Arguments in favor of mechanical bowel preparation included easier manipulation of an “empty” colon (especially in minimally invasive procedures) and avoidance of a “stool column” above an anastomosis, especially in the pelvis. Interestingly, a recent meta-analysis of 14 randomized controlled trials suggested that mechanical bowel preparation does not prevent surgical site infection and should be abandoned in clinical practice.

However, these studies did not include the use of oral antibiotics in the mechanical preparation groups, and further studies will be needed to determine the optimal preparation regimen.

**Ureteral Stents.** Ureteral stents may be useful for identifying the ureters intraoperatively and are placed via cystoscopy after the induction of general anesthesia and removed at the end of the operation. Stents can be invaluable during reoperative pelvic surgery or when there is significant retroperitoneal inflammation (such as complicated diverticulitis), as well as in obese patients. Lighted stents may be helpful in laparoscopic and robotic resections. Patients often have transient hematuria postoperatively, but major complications are rare.

**Multidisciplinary Teams.** Patients with complex colorectal disease often benefit from a multidisciplinary approach to their care. Patients with pelvic floor disorders (especially incontinence) often require evaluation by both a colorectal surgeon and a urologist or urogynecologist. Preoperative evaluation of cancer patients by a medical oncologist and/or radiation oncologist is crucial for planning either neoadjuvant or adjuvant therapy. Intraoperatively, complex pelvic resections often require the involvement of not only a colorectal surgeon but also a urologist, gynecologic oncologist, neurosurgeon, and/or plastic surgeon. Radiation oncologists should be involved in the operation if brachytherapy catheters are to be placed for intracavitary radiation or if intraoperative radiation therapy is planned. Rarely, psychiatric disorders may manifest as colorectal problems (especially functional disorders and chronic pain), and involvement of a psychiatrist or psychologist may be beneficial.

**INFLAMMATORY BOWEL DISEASE**

**General Considerations**

**Epidemiology.** Inflammatory bowel disease includes *ulcerative colitis*, *Crohn’s disease*, and *indeterminate colitis*. *Ulcerative colitis* occurs in 8 to 15 people per 100,000 in the United States and Northern Europe. The incidence is considerably lower in Asia, Africa, and South America, and among the nonwhite population in the United States. Ulcerative colitis incidence peaks during the third decade of life and again in the seventh decade of life. The incidence of Crohn’s disease is slightly lower, 1 to 5 people per 100,000. Crohn’s disease also affects Northern European and Caucasian populations disproportionately. Crohn’s disease has a similar bimodal incidence, with most cases occurring between ages 15 to 30 years and ages 60 to 70 years. In 15% of patients with inflammatory bowel disease, differentiation between ulcerative colitis and Crohn’s colitis is impossible; these patients are classified as having *indeterminate colitis*.

**Etiology.** Inflammatory bowel disease is a multifactorial condition that includes environmental, genetic, and immune causal elements; the variation in disease distributions and severity, as well as the differential responses to medical therapy, reflect a complex pathophysiology that is not reducible to a single cause. Nonetheless, there are several consistent observations among IBD populations that allow some degree of generalization. The consistent differences in IBD incidence between different geographic regions strongly suggest that environmental factors such as diet and exposure to microorganisms have a causal role. Alcohol and oral contraceptive use have also been implicated, as has tobacco use, in the etiology and exacerbation of Crohn’s disease.
IBD is a genetic disease, though one that is polygenic, explaining why IBD frequently affects multiple family members across more than one generation, while also explaining the large number of genes implicated in the development of IBD. Most of the data on the genetics of IBD focus on Crohn’s disease. Although ulcerative colitis has an association with at least 20 genetic loci based on genome-wide association studies (GWAS), this form of IBD has a weaker genetic link than does Crohn’s disease. Approximately 10% to 30% of IBD patients will have at least one other family member also affected by IBD. Additionally, there is 50% disease concordance among monozygotic twins and a 10% disease concordance among dizygotic twins.

Many of the genetic variants most consistently associated with IBD involve loci involved in innate immune function. These include NOD2 (nucleotide-binding oligomerization domain-containing protein 2), which is located on chromosome 16 and which is responsible for coordinating the function of several genes leading to the production of proinflammatory cytokines in response to gut microbes. This genetic variant is arguably the most strongly associated with IBD, being strongly connected with Crohn’s disease, although it is also associated with severe pouchitis in patients with a history of ulcerative colitis. The ATG16L1 gene is located on chromosome 2, and its product is involved in the immune response to muramyl dipeptide, a component of both gram-positive and gram-negative bacteria that is recognized by the immune system. A related gene, IRGM, is located on chromosome 5, and its product is pivotal in the interferon-gamma–mediated clearance of intracellular pathogens. Variants in this gene are associated with a higher incidence of ileocolic resections in Crohn’s disease patients.

Patients with IBD appear to have a chronic immune dysregulation, which may lead to an interplay with gut microbes which are also present in non-IBD patients, but which elicit pathologic immune responses in the IBD population leading to chronic, idiopathic inflammation of the alimentary tract. Bacteria such as Mycobacterium paratuberculosis and Listeria monocytogenes, as well as viruses such as parainfluenza virus and measles virus, have been suggested as having a role in the development of Crohn’s disease. With the decreased cost of sequencing and with the expansion of reference databases for identification of organisms, microbiome studies have been applied to the study of IBD patients, both to learn about the pathogenesis of IBD as well as for disease prognosis. Recent studies have identified an abundance of Serratia marcescens, E coli, and Candida tropicalis in the guts of patients with Crohn’s disease. In a study of 543 stool samples, patients with ulcerative colitis and primary sclerosing cholangitis (PSC) demonstrated a distinct bacterial community structure, with an enrichment of bacteria belonging to the Veillonella genus, which is associated with several diseases characterized by inflammation and fibrosis.

A defect in the gut mucosal barrier, which increases exposure to intestinal microbes as well as proinflammatory substances, is a potential etiologic factor related to immune dysregulation. An autoimmune mechanism has also been postulated. Although there is no clear evidence linking an immunologic disorder to inflammatory bowel disease, the similarity of many of the extraintestinal manifestations to rheumatologic disorders has made this theory attractive. In summary, IBD is primarily characterized by intestinal inflammation, and medical therapy is focused on reducing or preventing that inflammation.

**Pathology and Differential Diagnosis.** Although ulcerative colitis and Crohn’s colitis share many pathologic and clinical similarities, these conditions can be differentiated in 85% of patients. Ulcerative colitis is a mucosal process in which the colonic mucosa and submucosa are infiltrated with inflammatory cells. The mucosa may be atrophic, and crypt abscesses are common. Endoscopically, the mucosa is frequently friable and may possess multiple inflammatory pseudopolyps. In long-standing ulcerative colitis, the colon may be foreshortened and the mucosa replaced by scar. In quiescent ulcerative colitis, the colonic mucosa may appear normal both endoscopically and microscopically. Ulcerative colitis may affect the rectum (proctitis), rectum, and sigmoid colon (proctosigmoiditis), rectum and left colon (left-sided colitis), or the rectum and varying lengths of colon extending proximal to the splenic flexure (pancolitis). Ulcerative colitis does not primarily affect the small intestine, but the terminal ileum may demonstrate inflammatory changes (“backwash ileitis”). A key feature of ulcerative colitis is the continuous involvement of the rectum and colon; rectal sparing or skip lesions suggest a diagnosis of Crohn’s disease. Symptoms are related to the degree of mucosal inflammation and the extent of colitis. Patients typically complain of bloody diarrhea and crampy abdominal pain. Proctitis may produce tenesmus. Severe abdominal pain and fever raise the concern of fulminant colitis or toxic megacolon. Physical findings are nonspecific and range from minimal abdominal tenderness and distention to frank peritonitis. In the nonemergent setting, the diagnosis is typically made by colonoscopy and mucosal biopsy.

In contrast to ulcerative colitis, Crohn’s disease is a transmural inflammatory process that can affect any part of the gastrointestinal tract from mouth to anus. Mucosal ulcerations, an inflammatory cell infiltrate, and noncaseating granulomas are characteristic pathologic findings. Chronic inflammation may ultimately result in fibrosis, strictures, and fistulas in either the colon or small intestine. The endoscopic appearance of Crohn’s colitis is characterized by deep serpiginous ulcers and a “cobblestone” appearance. Skip lesions and rectal sparing are common. Symptoms of Crohn’s disease depend on the severity of inflammation and/or fibrosis and the location of inflammation in the gastrointestinal tract. Acute inflammation may produce diarrhea, crampy abdominal pain, and fever. Strictures may produce symptoms of obstruction. Weight loss is common, both because of obstruction and from protein loss. Perianal Crohn’s disease may present with pain, swelling, and drainage from fistulas or abscesses. Physical findings are also related to the site and severity of disease.

In 15% of patients with colitis from inflammatory bowel disease, differentiation of ulcerative colitis from Crohn’s colitis is impossible either grossly or microscopically (indeterminate colitis). These patients typically present with symptoms similar to ulcerative colitis. Endoscopic and pathologic findings usually include features common to both diseases. Increasingly, serologic markers have been employed to differentiate ulcerative colitis from Crohn’s disease. The anti-Saccharomyces cerevisiae antibody (ASCA) and perinuclear anticytoplasmic antibody (pANCA) may be useful in differentiating these two processes but require prospective study. Further differential diagnoses include infectious colitides, especially cytomegalovirus (CMV), Campylobacter jejuni, Entamoeba histolytica, toxigenic E Coli, C difficile, Neisseria gonorrhoeae, Salmonella, and Shigella species.
Extraintestinal Manifestations. The liver is a common site of extracolonic disease in inflammatory bowel disease. Fatty infiltration of the liver is present in 40% to 50% of patients, and cirrhosis is found in 2% to 5%. Fatty infiltration may be reversed by medical or surgical treatment of colonic disease, but cirrhosis is irreversible. Primary sclerosing cholangitis is a progressive disease characterized by intra- and extrahepatic bile duct strictures. Forty percent to 60% of patients with primary sclerosing cholangitis have ulcerative colitis. Colectomy will not reverse this disease, and the only effective therapy is liver transplantation. Pericholangitis is also associated with inflammatory bowel disease and may be diagnosed with a liver biopsy. Bile duct carcinoma is a rare complication of long-standing inflammatory bowel disease. Patients who develop bile duct carcinoma in the presence of inflammatory bowel disease are, on average, 20 years younger than other patients with bile duct carcinoma.51

Arthritis also is a common extracolonic manifestation of inflammatory bowel disease, and the incidence is 20 times greater than in the general population. Arthritis usually improves with treatment of the colonic disease. Sacroiliitis and ankylosing spondylitis are associated with inflammatory bowel disease, although the relationship is poorly understood. Medical and surgical treatment of the colonic disease does not impact symptoms.51

Erythema nodosum is seen in 5% to 15% of patients with inflammatory bowel disease and usually coincides with clinical disease activity. Women are affected three to four times more frequently than men. The characteristic lesions are raised, red, and predominantly on the lower legs. Pyoderma gangrenosum is an uncommon but serious condition that occurs almost exclusively in patients with inflammatory bowel disease. The lesion begins as an erythematous plaque, papule, or bleb, usually located on the pretibial region of the leg and occasionally near a stoma. The lesions progress and ulcerate, leading to a painful, necrotic wound. Pyoderma gangrenosum may respond to resection of the affected bowel in some patients. In others, this disorder is unaffected by treatment of the underlying bowel disease. One of the challenges in managing pyoderma is that this manifestation of IBD exhibits pathergy, where the disease will manifest and have its severity exacerbated by the creation of surgical sites. Depending on the circumstances, a history or the presence of pyoderma should prompt consideration for avoidance of a stoma.51

Up to 10% of patients with inflammatory bowel disease will develop ocular lesions. These include uveitis, iritis, episcleritis, and conjunctivitis, as well as macular degenerative, hyperpigmented pigmented epithelium (CHRPE). They usually develop during an acute exacerbation of the inflammatory bowel disease. The etiology is unknown.51

Principles of Nonoperative Management. Medical therapy for inflammatory bowel disease focuses on decreasing inflammation and alleviating symptoms, and many of the agents used are the same for both ulcerative colitis and Crohn’s disease. In general, mild to moderate flares may be treated in the outpatient setting. More severe signs and symptoms mandate hospitalization. Pancolitis generally requires more aggressive therapy than limited disease. Because ulcerative proctitis and proctosigmoiditis are limited to the distal large intestine, topical therapy with salicylate and/or corticosteroid suppositories and enemas can be extremely effective. Systemic therapy is rarely required in these patients.

Salicylates Sulfasalazine (Azulfidine), 5-acetyl salicylic acid (5-ASA), and related compounds are first-line agents in the medical treatment of mild to moderate inflammatory bowel disease. These compounds decrease inflammation by inhibition of cyclooxygenase and 5-lipoxygenase in the gut mucosa. They require direct contact with affected mucosa for efficacy. Multiple preparations are available for administration to different sites in the small intestine and colon (sulfasalazine, mesalamine [Pentasa, Asacol, Rowasa]).

Antibiotics Antibiotics are often used to decrease the intraluminal bacterial load in Crohn’s disease. Metronidazole has been reported to improve Crohn’s colitis and perianal disease, but the evidence is weak. Fluoroquinolones may also be effective in some cases. In the absence of fulminating colitis or toxic megacolon, antibiotics are not used to treat ulcerative colitis.

Corticosteroids Corticosteroids (either oral or parenteral) are a key component of treatment for an acute exacerbation of either ulcerative colitis or Crohn’s disease. Corticosteroids are nonspecific inhibitors of the immune system, and 75% to 90% of patients will improve with the administration of these drugs. However, corticosteroids have a number of serious side effects, and use of these agents should be limited to the shortest course possible. In addition, corticosteroids should be used judiciously in children because of the potential adverse effect on growth. Failure to wean corticosteroids is a relative indication for surgery.

Because of the systemic effects of corticosteroids, an effort has been made to develop drugs that act locally and have limited systemic absorption. Agents such as budesonide, beclomethasone dipropionate, and tixocortol pivalate undergo rapid hepatic degradation that significantly decreases systemic toxicity. Budesonide is available as an oral preparation. Corticosteroid enemas provide effective local therapy for proctitis and proctosigmoiditis and have fewer side effects than systemic corticosteroids.

Immunomodulating Agents Azathioprine and 6-mercaptopurine (6-MP) are antimetabolite drugs that interfere with nucleic acid synthesis and thus decrease proliferation of inflammatory cells. These agents are useful for treating ulcerative colitis and Crohn’s disease in patients who have failed salicylate therapy or who are dependent on, or refractory to, corticosteroids. It is important to note, however, that the onset of action of these drugs takes 6 to 12 weeks, and concomitant use of corticosteroids almost always is required.

Cyclosporine is an immunosuppressive agent that interferes with T-lymphocyte function. While cyclosporine is not routinely used to treat inflammatory bowel disease, up to 80% of patients with an acute flare of ulcerative colitis will improve with its use. However, the majority of these patients will ultimately require colectomy. Cyclosporine is also occasionally used to treat exacerbations of Crohn’s disease, and approximately two-thirds of patients will note some improvement. Improvement is generally apparent within 2 weeks of beginning cyclosporine therapy. Long-term use of cyclosporine is limited by its significant toxicities (e.g., nephrotoxicity, hirsutism, gum hypertrophy).

Methotrexate is a folate antagonist that also has been used to treat inflammatory bowel disease. Although the efficacy of this agent is unproven, there are reports that more than 50% of patients will improve with administration of this drug.52
Biologic Agents  In an effort to improve treatment for steroid-refractory inflammatory bowel disease, a class of agents has been developed based on inhibition of tumor necrosis factor alpha (TNF-α). Intravenous infusion of these agents decreases inflammation systemically. Infliximab is a monoclonal antibody directed against TNF-α and was the first biologic agent used to treat Crohn’s disease. More than 50% of patients with moderate to severe Crohn’s disease will improve with infliximab therapy. Infliximab is a chimeric monoclonal antibody directed against TNF-α and it was the first biologic agent used to treat Crohn’s disease. Based on the ACCENT I and II trials, infliximab was associated with a greater incidence of clinical remission, the ability to discontinue corticosteroids, and a longer length of remission compared to placebo. There are also multiple studies demonstrating an improvement in fistulizing perianal Crohn’s disease, although studies define “improvement” in a variety of ways, some of which do not require fistula to completely involute.

Because infliximab is a chimera partially consisting of mouse antibody, human antibodies directed against infliximab can mitigate the efficacy of this drug. For this reason, adalimumab and certolizumab, that have no nonhuman component, were developed. These two drugs can be administered by subcutaneous injection, allowing patients to self-administer these therapies and to avoid the cost and inconvenience of presenting to an infusion center. The ultimate goal of biologic agents, just as with other medical therapies for IBD, is mucosal healing.

The use of biologic agents for the treatment of ulcerative colitis is an area where opinions between gastroenterologists and surgeons are more divergent. Unlike in the case of Crohn’s disease, there is a putative surgical cure for ulcerative colitis, which casts the risks of long-term immunosuppression in a different light than in the setting of Crohn’s where there is no curative therapy available. For patients with moderate to severe ulcerative colitis not responding to other medical therapies, there is evidence supporting the use of infliximab (UC SUCCESS) and adalimumab (ULTRA I AND II). The use of infliximab as rescue therapy for inpatients with severe, steroid-dependent ulcerative colitis has more recently been investigated. Though gastroenterologists may view this intervention as maintenance therapy that begins in an inpatient setting only to be continued following hospital discharge, it is still unclear whether inpatient infliximab is even a reliable bridge to elective surgery. Many of the studies supporting inpatient rescue therapy with infliximab did not focus their analyses on patients with extensive (pan) colitis, the group at highest risk for requiring an unplanned admission due to a disease flare, and the subgroup of ulcerative colitis patients most likely to fail rescue therapy and to require colectomy.

Whether the preoperative use of biologics is associated with a higher incidence of postoperative complications is a matter of contention. There are individual studies demonstrating a higher incidence of postoperative sepsis, intra-abdominal abscess and readmissions for patients undergoing ileocolonectomy for Crohn’s disease who also received preoperative infliximab within three months of surgery. Systematic reviews on this topic, primarily focusing on Crohn’s disease, have concluded that perioperative infliximab may or may not be associated with a higher incidence of postoperative complications, with most of these complications being infectious in nature.

Nutrition  Patients with inflammatory bowel disease are often malnourished. Abdominal pain and obstructive symptoms may decrease oral intake. Diarrhea can cause significant protein loss. Ongoing inflammation produces a catabolic physiologic state. Parenteral nutrition should be strongly considered early in the course of therapy for either Crohn’s disease or ulcerative colitis. The nutritional status of the patient also should be considered when planning operative intervention, and nutritional parameters such as serum albumin, prealbumin, and transferrin should be assessed. In extremely malnourished patients, especially those who are also being treated with corticosteroids, creation of a stoma is often safer than a primary anastomosis.

Ulcerative Colitis  Ulcerative colitis is a dynamic disease characterized by remissions and exacerbations. The clinical spectrum ranges from an inactive or quiescent phase to low-grade active disease to fulminant disease. The onset of ulcerative colitis may be insidious, with minimal bloody stools, or the onset can be abrupt, with severe diarrhea and bleeding, tenesmus, abdominal pain, and fever. The severity of symptoms depends on the degree and extent of inflammation. Although anemia is common, massive hemorrhage is rare. Physical findings are often nonspecific.

The diagnosis of ulcerative colitis is almost always made endoscopically. Because the rectum is invariably involved, proctoscopy may be adequate to establish the diagnosis. The earliest manifestation is mucosal edema, which results in a loss of the normal vascular pattern. In more advanced disease, characteristic findings include mucosal friability and ulceration. Pelvic and rectal mucosae may also be present. While mucosal biopsy is often diagnostic in the chronic phase of ulcerative colitis, biopsy in the acute phase will often reveal only nonspecific inflammation. Evaluation with colonoscopy or barium enema during an acute flare is contraindicated because of the risk of perforation.

Barium enema has been used to diagnose chronic ulcerative colitis and to determine the extent of disease. However, this modality is less sensitive than colonoscopy and may not detect early disease. In long-standing ulcerative colitis, the colon is foreshortened and lacks hastrual markings (“lead pipe” colon). Because the inflammation in ulcerative colitis is purely mucosal, strictures are highly uncommon. Any stricture diagnosed in a patient with ulcerative colitis must be presumed to be malignant until proven otherwise.

Indications for Surgery  Indications for surgery in ulcerative colitis may be emergent or elective. Emergency surgery is required for patients with massive life-threatening hemorrhage, toxic megacolon, or fulminant colitis who fail to respond rapidly to medical therapy. Patients with signs and symptoms of fulminant colitis should be treated aggressively with bowel rest, hydration, broad-spectrum antibiotics, and parenteral corticosteroids. Colonoscopy and barium enema are contraindicated, and anti-diarrheal agents should be avoided. Deterioration in clinical condition or failure to improve within 24 to 48 hours mandates surgery.

Indications for elective surgery include intractability despite maximal medical therapy and high-risk development of major complications of medical therapy such as septic necrosis of joints secondary to chronic steroid use. Elective surgery also is indicated in patients at significant risk of developing colorectal carcinoma. The risk of malignancy increases with pancolonic disease and the duration of symptoms and is approximately 2% after 10 years, 8% after 20 years, and 18% after 30 years. Unlike sporadic colorectal cancers, carcinoma developing in the context of ulcerative colitis is more likely to arise from areas of
flat dysplasia and may be difficult to diagnose at an early stage. For this reason, it is recommended that patients with long-standing ulcerative colitis undergo colonoscopic surveillance with multiple (40–50), random biopsies to identify dysplasia before invasive malignancy develops. However, the adequacy of this type of screening is controversial. Recently, magnifying chromoendoscopy has been used to improve sensitivity. This technique uses topical dyes that are applied to the colonic mucosa at the time of endoscopy (Lugol’s solution, methylene blue, indigo carmine, and others). These dyes highlight contrast between normal and dysplastic epithelium, allowing more precise biopsy of suspicious areas. Surveillance is recommended annually after 8 years in patients with pancolitis, and annually after 15 years in patients with left-sided colitis. Although low-grade dysplasia was long thought to represent minimal risk, recent studies show that invasive cancer may be present in up to 20% of patients with low-grade dysplasia. For this reason, any patient with dysplasia should be advised to undergo proctocolectomy. Controversy exists over whether prophylactic proctocolectomy should be recommended for patients who have had chronic ulcerative colitis for greater than 10 years in the absence of dysplasia. Proponents of this approach note that surveillance colonoscopy with multiple biopsies samples only a small fraction of the colonic mucosa, and dysplasia and carcinoma are often missed. Opponents cite the relatively low risk of progression to carcinoma (approximately 2.4%) if all biopsies lack dysplasia. Neither approach has been shown definitively to decrease mortality from colorectal cancer.

Operative Management

Emergent Operation In a patient with fulminant colitis or toxic megacolon, total abdominal colectomy with ileostomy (with or without a mucus fistula), rather than total proctocolectomy, is recommended. Although the rectum is invariably diseased, most patients improve dramatically after an abdominal colectomy, and this operation avoids a difficult and time-consuming pelvic dissection in a critically ill patient. Rarely, a loop ileostomy and decompressing colostomy may be necessary if the patient is too unstable to withstand colectomy. Definitive surgery may then be undertaken at a later date once the patient has recovered. Complex techniques, such as an ileal pouch–anal reconstruction, generally are contraindicated in the emergent setting. However, massive hemorrhage that includes bleeding from the rectum may necessitate proctectomy and creation of either a permanent ileostomy or an ileal pouch–anal anastomosis.

Elective Operation Elective resection for ulcerative colitis usually is performed for refractory inflammation and/or the risk of malignancy (dysplasia). Because of the risk of ongoing inflammation, the risk of malignancy, and the availability of restorative proctocolectomy, most surgeons recommend operations that include resection of the rectum. Total proctocolectomy with end ileostomy has been the “gold standard” for treating patients with chronic ulcerative colitis. This operation removes the entire affected intestine and avoids the functional disturbances associated with ileal pouch–anal reconstruction. Most patients function well physically and psychologically after this operation. Total proctocolectomy with continent ileostomy (Kock’s pouch) was developed to improve function and quality of life after total proctocolectomy, but morbidity is significant, and restorative proctocolectomy is generally preferred today. Since its introduction in 1980, restorative proctocolectomy with ileal pouch–anal anastomosis has become the procedure of choice for most patients who require total proctocolectomy but wish to avoid a permanent ileostomy (see Figs. 29-11 and 29-12). Abdominal colectomy with ileorectal anastomosis may be appropriate for a patient with indeterminate colitis and rectal sparing.

Crohn’s Disease

Similar to ulcerative colitis, Crohn’s disease is characterized by exacerbations and remissions. Crohn’s disease, however, may affect any portion of the intestinal tract, from mouth to anus. Diagnosis may be made by colonoscopy or esophagogastroduodenoscopy or by barium small bowel study or enema, depending on which part of the intestine is most affected. The presence of skip lesions is key in differentiating Crohn’s colitis from ulcerative colitis, and rectal sparing occurs in approximately 40% of patients. The most common site of involvement of Crohn’s disease is the terminal ileum and cecum (ileocolic Crohn’s disease), followed by the small bowel, and then by the colon and rectum. Perianal and anal canal Crohn’s disease manifest by complex anal fistulae and/or abscesses, anal ulcers, and large skin tags may be the initial site of presentation in up to 4% of cases.

Indications for Surgery. Because Crohn’s disease is currently incurable and because it can affect any part of the gastrointestinal tract, the therapeutic rationale is fundamentally different from that of ulcerative colitis. Ulcerative colitis may be cured by removal of the affected intestinal segment (the colon and rectum). In Crohn’s disease, it is impossible to remove all the at-risk intestine; therefore, surgical therapy is reserved for complications of the disease.

Crohn’s disease may present as an acute inflammatory process or as a chronic fibrotic process. During the acute inflammatory phase, patients may present with intestinal inflammation complicated by fistulae and/or intra-abdominal abscesses. Maximal medical therapy should be instituted, including anti-inflammatory medications, bowel rest, and antibiotics. Parenteral nutrition should be considered if the patient is malnourished. Most intra-abdominal abscesses can be drained percutaneously with the use of CT scan guidance. Although the majority of these patients will ultimately require surgery to remove the diseased segment of bowel, these interventions allow the patient’s condition to stabilize, nutrition to be optimized, and inflammation to decrease prior to embarking on a surgical resection. Once an operation is undertaken, fistulae generally require resection of the segment of bowel with active Crohn’s disease; the secondary sites of the fistula are often otherwise normal and do not generally require resection after division of the fistula. Simple closure of the secondary fistula site usually suffices.

Chronic fibrosis may result in strictures in any part of the gastrointestinal tract. Because the fibrotic process is gradual, free perforation proximal to the obstructing stricture is rare. Chronic strictures almost never improve with medical therapy. Strictures may be treated with resection or strictureplasty. Distal ileal strictures are sometimes amenable to colonoscopic balloon dilatation. Optimal timing for surgery should take into account the patient’s underlying medical and nutritional status.

Once an operation is undertaken for Crohn’s disease, several principles should guide intraoperative decision making. In general, a laparotomy for Crohn’s disease should be performed through a midline incision because of the possible need for a stoma. Laparoscopy is also increasingly used in this setting. Because many patients with Crohn’s disease often will require multiple operations, the length of bowel removed should be...
minimized. Bowel should be resected to an area with grossly normal margins; frozen sections are not necessary. Finally, a primary anastomosis may be created safely if the patient is medically stable, nutritionally replete, and taking few immunosuppressive medications. Creation of a stoma should be strongly considered in any patient who is hemodynamically unstable, septic, malnourished, or receiving high-dose immunosuppressive therapy and among patients with extensive intra-abdominal contamination.

Ileocolic and Small Bowel Crohn’s Disease. The terminal ileum and cecum are involved in Crohn’s disease in up to 41% of patients; the small intestine is involved in up to 35% of patients. The most common indications for surgery are internal fistula or abscess (30–38% of patients) and obstruction (35–37% of patients). Psoas abscess may result from ileocolic Crohn’s disease. Sepsis should be controlled with percutaneous drainage of abscess(es) and antibiotics, if possible. Parenteral nutrition may be necessary in patients with chronic obstruction. The extent of resection depends on the amount of involved intestine. Short segments of inflamed small intestine and right colon should be resected and a primary anastomosis created if the patient is stable, nutrition is adequate, and immunosuppression is minimal. Isolated chronic strictures should also be resected. In patients with multiple fibrotic strictures that would require extensive small bowel resection, stricturoplasty is a safe and effective alternative to resection. Short strictures are amenable to a transverse stricturoplasty, while longer strictures may be treated with a side-to-side small bowel anastomosis (Fig. 29-18).

Figure 29-18. Alternative stricturoplasty techniques. A. A short stricture is opened along the antimesenteric surface of the bowel wall. B. The enterotomy is closed transversely. C. A long stricture is opened along the antimesenteric surface of the bowel wall. D. The bowel is folded into an inverted “U.” E. A side-to-side anastomosis is made. (Reproduced with permission from Corman ML. Colon & Rectal Surgery, 2nd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 1989.)
Risk of recurrence after resection for ileocolic and small bowel Crohn’s disease is high. More than 50% of patients will experience a recurrence within 10 years, and the majority of these will require a second operation.

**Crohn’s Colitis.** Crohn’s disease of the large intestine may present as *fulminant colitis* or *toxic megacolon*. In this setting, treatment is identical to treatment of fulminant colitis and toxic megacolon secondary to ulcerative colitis. Resuscitation and medical therapy with bowel rest, broad-spectrum antibiotics, and parenteral corticosteroids should be instituted. If the patient’s condition worsens or fails to rapidly improve, total abdominal colectomy with end ileostomy is recommended. An elective proctectomy may be required if the patient has refractory Crohn’s proctitis. Alternatively, if the rectum is spared, an ileorectal anastomosis may be appropriate once the patient has recovered.

Other indications for surgery in chronic Crohn’s colitis are intractability, complications of medical therapy, and risk of or development of malignancy. Unlike ulcerative colitis, Crohn’s colitis may be segmental, and rectal sparing is often observed. A segmental colectomy may be appropriate if the remaining colon and/or rectum appear normal. An isolated colonic stricture may also be treated by segmental colectomy. Although it was long thought that Crohn’s disease did not increase the risk of colorectal carcinoma, it is now recognized that Crohn’s colitis (especially pancolitis) carries nearly the same risk for cancer as ulcerative colitis. Annual surveillance colonoscopy with multiple biopsies is recommended for patients with long-standing Crohn’s colitis (>7 years in duration). As in ulcerative colitis, dysplasia is an indication for total proctocolectomy. Ileal pouch–anal reconstruction is not recommended in these patients because of the risk for development of Crohn’s disease within the pouch and the high risk of complications, such as fistula, abscess, stricture, pouch dysfunction, and pouch failure.

**Anal and Perianal Crohn’s Disease.** Anal and perianal manifestations of Crohn’s disease are very common and occur in 35% of all patients with Crohn’s disease. Isolated anal Crohn’s disease is uncommon, affecting only 3% to 4% of patients. Detection of anal Crohn’s disease, therefore, should prompt evaluation of the remainder of the gastrointestinal tract.

The most common perianal lesions in Crohn’s disease are *skin tags* that are minimally symptomatic. *Fissures* are also common. Typically, a fissure from Crohn’s disease is particularly deep or broad and perhaps better described as an anal ulcer. These fissures are often multiple and located in a lateral position rather than anterior or posterior midline as seen in an idiopathic fissure in ano. A classic-appearing fissure in ano located laterally should raise the suspicion of Crohn’s disease. *Perianal abscess* and *fistulas* are common and can be particularly challenging. Fistulas tend to be complex and often have multiple tracts (Fig. 29-19). *Hemorrhoids* are not more common in patients with Crohn’s disease than in the general population, although many patients tend to attribute any anal or perianal symptom to “hemorrhoids.”

Treatment of anal and perianal Crohn’s disease focuses on alleviation of symptoms. Perianal skin irritation from diarrhea often responds to medical therapy directed at small bowel or colonic disease. In general, skin tags and hemorrhoids should not be excised unless they are extremely symptomatic because of the risk of creating chronic, nonhealing wounds. Fissures may respond to local or systemic therapy; sphincterotomy is relatively contraindicated because of the risk of creating a chronic, nonhealing wound and because of the increased risk of incontinence in a patient with diarrhea from underlying colitis or small bowel disease. Anal ulcers associated with Crohn’s disease are usually not very painful unless there is an underlying abscess. Thus, in patients with significant anal pain, an examination under anesthesia is indicated to exclude an underlying abscess or fistula and to assess the rectal mucosa. In the absence of active Crohn’s proctitis, one can proceed cautiously with a partial internal sphincterotomy if the examination under anesthesia reveals a classic-appearing posterior or anterior fissure and anal stenosis.

Recurrent abscess(es) or complex anal fistulae should raise the possibility of Crohn’s disease. Treatment focuses on control of infection, delineation of complex anatomy, treatment of underlying mucosal disease, and sphincter preservation. Abscesses often can be drained locally, and mushroom catheters are useful for maintaining drainage. Endoanal ultrasound and pelvic MRI are useful for mapping complex fistulous tracts. Liberal use of setons can control many fistulas and avoid division of the sphincter. Many patients with anal Crohn’s disease function well with multiple setons left in place for years. Endoanal advancement flaps may be considered for definitive therapy if the rectal mucosa is uninvolved but will not heal due to rectal inflammation. In 10% to 15% of cases, intractable perianal sepsis requires proctectomy.

Rectovaginal fistula can be a particularly difficult problem in these patients. A rectal or vaginal mucosal advancement flap may be used if the rectal mucosa appears healthy and scarring of the rectovaginal septum is minimal. Occasionally, proctectomy is the best option for women with highly symptomatic rectovaginal fistulae. Although proximal diversion is often employed to protect complex perianal reconstruction, there is no evidence that diversion alone increases healing of anal and perianal Crohn’s disease.

Medical treatment of underlying proctitis with salicylate and/or corticosteroid enemas may be helpful; however, control of infection is the primary goal of therapy. Metronidazole has been used with some success in this setting. Anti-TNF-α agents (infliximab and adalimumab) have shown some efficacy in healing chronic fistulas secondary to Crohn’s disease. The success of these agents has led to a concerted effort to identify other
immunomodulators that might prove useful. Proinflammatory cytokines such as interleukin-12 and interferon-γ are potential targets. Inhibition of immune cell migration has also been suggested as an approach. However, it is of paramount importance to drain any and all abscesses before initiating immunosuppressive therapy such as corticosteroids or anti-TNF-α monoclonal antibodies.\(^5^4\)\(^5^5\)

**Indeterminate Colitis**
Approximately 15% of patients with inflammatory bowel disease manifest clinical and pathologic characteristics of both ulcerative colitis and Crohn’s disease. Endoscopy, barium enema, and biopsy may be unable to differentiate ulcerative colitis from Crohn’s colitis in this setting. The indications for surgery are the same as those for ulcerative colitis: **intractability, complications of medical therapy, and risk of or development of malignancy.** In the setting of indeterminate colitis in a patient who prefers a sphincter-sparing operation, a total abdominal colectomy with end ileostomy may be the best initial procedure. Pathologic examination of the entire colon may then allow a more accurate diagnosis. If the diagnosis suggests ulcerative colitis, an ileal pouch–anal anastomosis procedure can be performed. If the diagnosis remains in question, the safest surgical option is completion proctectomy with end ileostomy (similar to Crohn’s colitis). Ileal pouch–anal reconstruction may also be considered with the understanding that the pouch failure rate is between 15% and 20%.\(^5^0\)

**DIVERTICULAR DISEASE**
*Diverticulosis* is a clinical term used to describe the presence of symptomatic diverticula. *Diverticulosis* refers to the presence of diverticula without inflammation. *Diverticulitis* refers to inflammation and infection associated with diverticula. The majority of colonic diverticula are *false diverticula* in which the mucosa and muscularis mucosa have herniated through the colonic wall. These diverticula occur between the teniae coli, at points where nutrient arterial blood vessels penetrate the colonic wall (presumably creating an area of relative weakness in the colonic muscle). They are thought to be *pulsion* diverticula resulting from high intraluminal pressure. *Diverticular bleeding* can be massive but usually is self-limited. *True divertica*, which comprise all layers of the bowel wall, are rare and are usually congenital in origin.

Diverticulosis is extremely common in the United States and Europe. It is estimated that half of the population older than age 50 years has colonic diverticulosis. The sigmoid colon is the most common site of diverticulosis (Fig. 29-20). Diverticulosis is thought to be an acquired disorder, but the etiology is poorly understood. The most accepted theory is that a lack of dietary fiber results in smaller stool volume, requiring high intraluminal pressure and high colonic wall tension for propulsion. Chronic contraction then results in muscular hypertrophy and development of the process of segmentation in which the colon acts like separate segments instead of functioning as a continuous tube. As segmentation progresses, the high pressures are directed radially toward the colon wall rather than to development of propulsive waves that move stool distally. The high radial pressures directed against the bowel wall create pulsion diverticula. A loss of tensile strength and a decrease in elasticity of the bowel wall with age have also been proposed etiologies. Although none of these theories has been proven, a high-fiber diet does appear to decrease the incidence of diverticulosis. Although diverticulosis is common, most cases are asymptomatic, and complications occur in the minority of people with this condition.

**Inflammatory Complications (Diverticulitis)**
*Diverticulitis* refers to inflammation and infection associated with a diverticulum and is estimated to occur in 10% to 25% of people with diverticulosis. Peridiverticular and pericolic infection results from a perforation (either macroscopic or microscopic) of a diverticulum, which leads to contamination, inflammation, and infection. The spectrum of disease ranges from mild, uncomplicated diverticulitis that can be treated in the outpatient setting, to free perforation and diffuse peritonitis that requires emergency laparotomy. Most patients present with left-sided abdominal pain, with or without fever, and leukocytosis. A mass may be present. Plain radiographs are useful for detecting free intra-abdominal air. CT scan is extremely useful for defining pericolic inflammation, phlegmon, or abscess. Contrast enemas and/or endoscopy are relatively contraindicated because of the risk of perforation. The differential diagnosis includes malignancy, ischemic colitis, infectious colitis, and inflammatory bowel disease.

**Uncomplicated Diverticulitis.** Uncomplicated diverticulitis is characterized by left lower quadrant pain and tenderness. CT findings include pericolic soft tissue stranding, colonic wall thickening, and/or phlegmon. Most patients with uncomplicated diverticulitis will respond to outpatient therapy with broad-spectrum oral antibiotics and a low-residue diet. Antibiotics should be continued for 7 to 10 days. About 10% to 20% of patients with more severe pain, tenderness, fever, and leukocytosis are treated in the hospital with parenteral antibiotics.
and bowel rest. Most patients improve within 48 to 72 hours. Failure to improve may suggest abscess formation. CT can be extremely useful in this setting, and many pericolic abscesses can be drained percutaneously. Deterioration in a patient’s clinical condition and the development of peritonitis are indications for laparotomy.

Most patients with uncomplicated diverticulitis will recover without surgery, and 50% to 70% will have no further episodes. It has long been believed that the risk of complications increases with recurrent disease. For this reason, elective sigmoid colectomy has often been recommended after the second episode of diverticulitis, especially if the patient has required hospitalization. Resection has often been recommended after the first episode in very young patients and is often recommended after the first episode of complicated diverticulitis. These general guidelines have been questioned in recent years, and more recent studies suggest that the risk of complications and/or need for emergent resection does not increase with recurrent disease. Moreover, the rate of complications is rare after elective surgery and recurrences do not increase the rate of complications. As such, the rate of resection in all patients, including young patients and those with complicated disease, has decreased.70-72

Many surgeons now will not advise colectomy even after two documented episodes of diverticulitis assuming the patient is completely asymptomatic and that carcinoma has been excluded by colonoscopy. Immunosuppressed patients are generally still advised to undergo colectomy after a single episode of documented diverticulitis. Medical comorbidities should be considered when evaluating a patient for elective resection, and the risks of recurrent disease should be weighed against the risks of the operation.69,70 Because colon carcinoma may present in an identical fashion to diverticulitis (either complicated or uncomplicated), all patients must be evaluated for malignancy after resolution of the acute episode. Colonoscopy is recommended 4 to 6 weeks after recovery. Inability to exclude malignancy is another indication for resection.

In the elective setting, a sigmoid colectomy with a primary anastomosis is the procedure of choice. The resection should always be extended to the rectum distally because the risk of recurrence is high if a segment of sigmoid colon is retained.

The proximal extent of the resection should include all thickened or inflamed bowel; however, resection of all diverticula is unnecessary. Increasingly, laparoscopy is being used for elective sigmoid colectomy for diverticular disease.

**Complicated Diverticulitis.** Complicated diverticulitis includes diverticulitis with abscess, obstruction, diffuse peritonitis (free perforation), or fistulas between the colon and adjacent structures. Colovesical, colovaginal, and coloenteric fistulas are long-term sequelae of complicated diverticulitis. The Hinchey staging system is often used to describe the severity of complicated diverticulitis: Stage I includes colonic inflammation with an associated pericolic abscess; stage II includes colonic inflammation with a retroperitoneal or pelvic abscess; stage III is associated with purulent peritonitis; and stage IV is associated with fecal peritonitis. Treatment depends on the patient’s overall clinical condition and the degree of peritoneal contamination and infection. Small abscesses (<2 cm in diameter) may be treated with parenteral antibiotics. Larger abscesses are best treated with CT-guided percutaneous drainage (Fig. 29-21) and antibiotics.72 Many of these patients will ultimately require resection, but percutaneous drainage may allow a one-stage, elective procedure and may obviate the need for colectomy if full recovery follows the drainage.

Urgent or emergent laparotomy may be required if an abscess is inaccessible to percutaneous drainage, if the patient’s condition deteriorates or fails to improve, or if the patient presents with free intra-abdominal air or peritonitis. In almost all cases, an attempt should be made to resect the affected segment of bowel. Patients with small, localized pericolic or pelvic abscesses (Hinchey stages I and II) may be candidates for a sigmoid colectomy with a primary anastomosis (a one-stage operation). Among patients with larger abscesses, peritoneal soiling, or peritonitis, sigmoid colectomy with end colostomy and Hartmann’s pouch is the most commonly used procedure.71 Success also has been reported after sigmoid colectomy, primary anastomosis, with or without on-table lavage, and proximal diversion (loop ileostomy). This option may be appropriate in stable patients and offers the great advantage that the subsequent operation to restore bowel continuity is simpler than is

---

**Figure 29-21.** A. Computed tomography scan demonstrating pelvic abscess from perforated diverticular disease. B. Posterolateral computed tomography–guided drainage of abdominal abscess from perforated diverticular disease. (Used with permission from Charles O. Finne III, MD, Minneapolis, MN.)
takedown of a Hartmann’s pouch. The presence of inflammation and phlegmon may increase the risk of ureteral damage during mobilization of the sigmoid colon, and preoperative placement of ureteral catheters can be invaluable. In extremely unstable patients, or in the presence of such severe inflammation that resection would harm adjacent organs, proximal diversion and local drainage have been employed. However, this approach is generally avoided because of high morbidity and mortality rates, along with the requirement for multiple operations. More recently, several studies have suggested that laparoscopic lavage and drainage without bowel resection may be safe and effective even in the presence of free perforation. However, a 20% risk of surgical reoperation can be expected.73,74

Obstructive symptoms occur in approximately 67% of patients who develop acute diverticulitis, and complete obstruction occurs in 10%. Patients with incomplete obstruction often respond to fluid resuscitation, nasogastric suction, and gentle, low-volume water or Gastrografin enemas. Relief of obstruction allows full bowel preparation and elective resection. A high-volume oral bowel preparation is contraindicated in the presence of obstructive symptoms. Obstruction that does not rapidly respond to medical management mandates laparotomy. Sigmoid colectomy with end colostomy is the safest procedure to perform in this setting. However, colectomy and primary anastomosis, with or without on-table lavage (depending on extent of fecal load in the proximal colon), and proximal diversion may be appropriate if the patient is stable and the proximal and distal bowel appear healthy.

Approximately 5% of patients with complicated diverticulitis develop fistulas between the colon and an adjacent organ. Colovescial fistulas are most common, followed by colovaginal and colorectal fistulas. Colocutaneous fistulas are a rare complication of diverticulitis. Two key points in the evaluation of fistulas are to define the anatomy of the fistula and exclude other diagnoses. Contrast enema and/or small bowel studies are extremely useful in defining the course of the fistula. CT scan can identify associated abscesses or masses. The differential diagnosis includes malignancy, Crohn’s disease, and radiation-induced fistulas. While Crohn’s disease and radiation injury may be suspected based on the patient’s medical history, colonoscopy or sigmoidoscopy is required to rule out malignancy. In addition, in a patient who has received radiation therapy, a fistula must be considered to be recurrent cancer until proven otherwise. Once the anatomy of the fistula has been defined and other diagnoses excluded, operative management should include resection of the affected segment of the colon involved with diverticulitis (usually with a primary anastomosis) and simple repair of the secondarily involved organ. Suspicion of carcinoma may mandate a wider, en bloc resection.

Hemorrhage

Bleeding from a diverticulum results from erosion of the peri-diverticular arteriole and may result in massive hemorrhage. Most significant lower gastrointestinal hemorrhage occurs in elderly patients in whom both diverticulosis and angiodysplasia are common. Consequently, the exact bleeding source may be difficult to identify. Fortunately, in 80% of patients, bleeding stops spontaneously. Clinical management should focus on resuscitation and localization of the bleeding site as described for lower gastrointestinal hemorrhage. Colonoscopy may occasionally identify a bleeding diverticulum that may then be treated with epinephrine injection or cautery. Angiography may be diagnostic and therapeutic in this setting. In the rare instance in which diverticular hemorrhage persists or recurs, laparotomy and segmental colectomy may be required.

Giant Colonic Diverticulum

Giant colonic diverticula are extremely rare. Most occur on the antimesenteric side of the sigmoid colon. Patients may be asymptomatic or may present with vague abdominal complaints such as pain, nausea, or constipation. Plain radiographs may suggest the diagnosis. Barium enema is usually diagnostic. Complications of a giant diverticulum include perforation, obstruction, and volvulus. Resection of the involved colon and diverticulum is recommended.

Right-Sided Diverticula

The cecum and ascending colon infrequently are involved in diverticulosis coli. Even more uncommon is a true solitary diverticulum, which contains all layers of the bowel wall and is thought to be congenital in origin. Right-sided diverticula occur more often in younger patients than do left-sided diverticula and are more common in people of Asian descent than in other populations. Most patients with right-sided diverticula are asymptomatic. However, diverticulitis does occur occasionally. Because patients are young and present with right lower quadrant pain, they are often thought to suffer from acute appendicitis, and the diagnosis of right-sided diverticulitis is subsequently made in the operating room. If there is a single large diverticulum and minimal inflammation, a diverticulectomy may be performed, but an ileocolic resection is usually the preferred operation in this setting. Hemorrhage rarely occurs and should be treated in the same fashion as hemorrhage from a left-sided diverticulum.

ADENOCARCINOMA AND POLYPY

Incidence

Colorectal carcinoma is the most common malignancy of the gastrointestinal tract. Over 130,000 new cases are diagnosed annually in the United States, and more than 50,000 patients die of this disease each year, making colorectal cancer the third most lethal cancer in the United States.75

The incidence is similar in men and women and has remained fairly constant over the past 20 years; however, the widespread adoption of current national screening programs is gradually decreasing the incidence of this common and lethal disease in people over 50 years of age. However, people younger than 50 have been experiencing and increase incidence and worse mortality.75 Early detection and improvements in medical and surgical care are thought to be responsible for the decreasing mortality of colorectal cancer observed in recent years.

Epidemiology (Risk Factors)

Identification of risk factors for development of colorectal cancer is essential to establish screening and surveillance programs in appropriately targeted populations.

Aging. Aging is the dominant risk factor for colorectal cancer, with incidence rising steadily after age 50 years. More than 90% of cases diagnosed are in people older than age 50 years. This is the rationale for initiating screening tests of asymptomatic patients at average risk of developing colorectal cancer at age 50 years. However, individuals of any age can develop colorectal cancer, so symptoms such as a significant change in bowel
habits, rectal bleeding, melena, unexplained anemia, or weight loss require a thorough evaluation.

**Hereditary Risk Factors.** Approximately 80% of colorectal cancers occur sporadically, while 20% arise in patients with a known family history of colorectal cancer. Advances in the understanding of these familial disorders have led to interest in early diagnosis using genetic testing. Because of the medical, legal, and ethical considerations that are involved in this type of testing, all patients should be offered genetic counseling if a familial syndrome is suspected.

**Environmental and Dietary Factors.** The observation that colorectal carcinoma occurs more commonly in populations that consume diets high in animal fat and low in fiber has led to the hypothesis that dietary factors contribute to carcinogenesis. A diet high in saturated or polyunsaturated fats increases risk of colorectal cancer, while a diet high in oleic acid (olive oil, coconut oil, fish oil) does not increase risk. Animal studies suggest that fats may be directly toxic to the colonic mucosa and thus may induce early malignant changes. In contrast, a diet high in vegetable fiber appears to be protective. A correlation between alcohol intake and incidence of colorectal carcinoma has also been suggested. Ingestion of calcium; selenium; vitamins A, C, and E; carotenoids; and plant phenols may decrease the risk of developing colorectal cancer. Obesity and sedentary lifestyle dramatically increase cancer-related mortality in a number of malignancies, including colorectal carcinoma. This knowledge is the basis for primary prevention strategies to eliminate colorectal cancer by altering diet and lifestyle.

**Inflammatory Bowel Disease.** Patients with long-standing colitis from inflammatory bowel disease are at increased risk for the development of colorectal cancer. It is hypothesized that chronic inflammation predisposes the mucosa to malignant changes, and there is some evidence that degree of inflammation influences risk. In general, the duration and extent of colitis correlate with risk. Other factors thought to increase risk include the presence of primary sclerosing cholangitis and family history of colorectal cancer.

**Other Risk Factors.** Cigarette smoking is associated with an increased risk of colonic adenomas, especially after more than 35 years of use. Patients with ulcerative colitis may also be at increased risk for both adenoma and carcinoma formation.

Acromegaly, which is associated with increased levels of circulating human growth hormone and insulin-like growth factor-1, increases risk as well. Pelvic irradiation may increase the risk of developing rectal carcinoma. However, it is unclear whether this represents a direct effect of radiation damage or is instead a correlation between the development of rectal cancer and a history of another pelvic malignancy; for example, among patients who develop prostate cancer and are treated with radiation, the risk of rectal cancer increases significantly.

**Pathogenesis of Colorectal Cancer**

**Genetic Defects.** An intense research effort has focused on elucidating the genetic defects and molecular abnormalities associated with the development and progression of colorectal adenomas and carcinoma. Mutations may cause activation of oncogenes (K-ras) and/or inactivation of tumor suppressor genes (APC, deleted in colorectal carcinoma [DCC], p53). Colorectal carcinoma is thought to develop from adenomatous polyps by accumulation of these mutations in what has come to be known as the adenoma-carcinoma sequence (Fig. 29-22).

Defects in the APC gene were first described in patients with Familial Adenomatous Polyposis (FAP). By investigating these families, characteristic mutations in the APC gene were identified. They are now known to be present in 80% of sporadic colorectal cancers as well.

The APC gene is a tumor suppressor gene. Mutations in both alleles are necessary to initiate polypl formation. The majority of mutations are premature stop codons, which result in a truncated APC protein. In FAP, the site of mutation correlates with the clinical severity of the disease. For example, mutations in either the 3′ or 5′ end of the gene result in attenuated forms of FAP (AFAP), whereas mutations in the center of the gene result in more virulent disease. Thus, knowledge of the specific mutation in a family may help guide clinical decision-making.

APC inactivation alone does not result in a carcinoma. Instead, this mutation sets the stage for the accumulation of genetic damage that results in malignancy. Additional mutations may include activation or inactivation of a variety of genes.

One of the most commonly involved genes in colorectal cancer is K-ras. K-ras, a signaling molecule in the epidermal growth factor receptor (EGFR) pathway, is classified as a proto-oncogene because mutation of only one allele will perturb the cell cycle. The K-ras gene product is a G-protein involved in intracellular signal transduction. When active, K-ras binds guanosine triphosphate (GTP); hydrolysis of GTP to guanosine diphosphate (GDP) then inactivates the G-protein. Mutation of KRAS results in an inability to hydrolyze GTP, thus leaving the G-protein permanently in the active form. It is thought that this then leads to uncontrolled cell division. Because K-ras mutation results in uncontrolled downstream signaling, anti-EGFR agents are ineffective in treating K-ras mutant tumors. As such, K-ras mutation status is important in deciding when to utilize anti-EGFR therapies. Other EGFR signaling molecules such as BRAF have also been implicated in colorectal cancer pathogenesis and progression, and ongoing research is focusing on elucidating their roles in this disease.

**Figure 29-22.** Schematic showing progression from normal colonic epithelium to carcinoma of the colon.
Another common mutation occurs in the MYH gene on chromosome 1p. MYH is a base excision repair gene, and biallelic deletion results in changes in other downstream molecules. Since its discovery, MYH mutations have been associated with an AFAP phenotype in addition to sporadic cancers. Unlike APC gene mutations that are expressed in an autosomal dominant pattern, the requirement for biallelic mutation in MYH results in an autosomal recessive pattern of inheritance.

The tumor suppressor gene p53 has been well characterized in a number of malignancies. The p53 protein appears to be crucial for initiating apoptosis in cells with irreparable genetic damage. Mutations in p53 are present in 75% of colorectal cancers.

Deletion of the tumor suppressor phosphatase and tensin homolog (PTEN) appears to be involved in a number of hamartomatous polyposis syndromes. Deletions in PTEN have been identified in juvenile polyposis, Peutz-Jeghers syndrome, Cowden’s syndrome, and PTEN hamartoma syndrome, in addition to multiple endocrine neoplasia type IIB. Peutz-Jeghers syndrome is also associated with mutation in STK11, a serine-threonin kinase gene. The genetic changes that underlie serrated polyposis syndrome(s) are currently poorly understood.

**Genetic Pathways.** The mutations involved in colorectal cancer pathogenesis and progression are now recognized to accumulate via one of three major genetic pathways: the loss of heterozygosity (LOH; chromosomal instability) pathway, the microsatellite instability (MSI) pathway, and the CpG island methylation (CIMP; serrated methylated) pathway.

**The Loss of Heterozygosity Pathway** The LOH pathway is characterized by chromosomal deletions and tumor aneuploidy. Eighty percent of colorectal carcinomas appear to arise from mutations in the LOH pathway. This pathway was first described in patients with FAP in whom mutations of the APC gene were found to be inherited.

Another example of LOH occurs in the region of chromosome 18q. This region has been found to be deleted in up to 70% of colorectal cancers. Two tumor suppressor genes, DCC and SMAD4, are located in this region, and as such, deletion of 18q may result in the loss of one or both of these genes. DCC is a tumor suppressor gene, and loss of both alleles is required for malignant degeneration. The main role of this molecule appears to be in the central nervous system, where it is involved in neural differentiation and axonal migration. This observation has led to the hypothesis that DCC may be involved in differentiation and cellular adhesion in colorectal cancer, but this theory remains unproven. DCC mutations are present in more than 70% of colorectal carcinomas and may negatively impact prognosis. SMAD4 functions in the signaling cascade of transforming growth factor beta and beta-catenin (also a downstream effector of the APC gene). Loss of either of these genes is thought to promote cancer progression.

**The Microsatellite Instability Pathway.** Many of the remaining colorectal carcinomas are thought to arise from mutations in the MSI pathway, which is characterized by errors in mismatch repair during DNA replication. These errors in mismatch repair were first described in Hereditary Nonpolyposis Colon Cancer (HNPCC; Lynch syndrome) Lynch 9, but are now recognized to be present in many sporadic tumors as well. A number of genes have been identified that appear to be crucial for recognizing and repairing DNA replication errors. These mismatch repair genes include MSH2, MLH1, PMS1, PMS2, and MSH6/GTBP. A mutation in one of these genes predisposes a cell to mutations, which may occur in proto-oncogenes or tumor suppressor genes. Accumulation of these errors then leads to genomic instability and ultimately to carcinogenesis.

Microsatellites are regions of the genome in which short base-pair segments are repeated several times, regions that are particularly prone to replication error. Consequently, a mutation in a mismatch repair gene produces variable lengths of these repetitive sequences, a finding that has been described as MSI.

Tumors associated with MSI appear to have different biologic characteristics than do tumors that result from the LOH pathway. Tumors with MSI are more likely to be in the right colon and possess diploid DNA and are associated with a better prognosis than tumors that arise from the LOH pathway that are microsatellite stable. Tumors arising from the LOH pathway tend to occur in the more distal colon, often have chromosomal aneuploidy, and are associated with a poorer prognosis.

**CpG Island Methylation Pathway** In the CIMP pathway, genes do not accumulate mutations (deletions or insertions of bases), but instead are activated or inactivated by methylation. This process has been called epigenetic alteration to differentiate it from the more traditional genetic alterations or true mutations. In normal cells, methylation is critical for regulation of gene expression. In cancer, aberrant methylation (either hyper- or hypomethylation), usually of a promoter region, results in abnormal activation or inactivation of genes. This gene silencing or, alternatively, activation results in a phenotype similar to that present with a true gene mutation. This pathway has also been called the serrated methylated pathway because of the observation that serrated polyps often harbor aberrant methylation in contrast to adenomatous polyps that are more often associated with mutations in the APC gene (LOH pathway).

Although these classifications are useful for understanding the mechanisms underlying carcinogenesis, they are not mutually exclusive. For example, a mismatch repair gene may be inactivated by methylation. Errors in mismatch repair may then allow mutations to inactivate a tumor suppressor gene. In addition, there is considerable interest in targeting molecules in each of these pathways in order to design better anticancer agents. Finally, ongoing research is focusing on the utility of molecular profiling in predicting prognosis and/or response to treatment.

**Polyps**

It is now well accepted that the majority of colorectal carcinomas evolve from adenomatous polyps; this sequence of events is the adenoma-carcinoma sequence. Polyp is a nonspecific clinical term that describes any projection from the surface of the intestinal mucosa regardless of its histologic nature. Colorectal polyps may be classified as neoplastic (tubular adenoma, villous adenoma, tubulovillous adenomas, serrated adenomas/polyps), hyperplastic, hamartomatous (juvenile, Peutz-Jeghers, Cronkite-Canada), or inflammatory (pseudopolyp, benign lymphoid polyp).

**Neoplastic Polyps.** Adenomatous polyps are common, occurring in up to 25% of the population older than 50 years of age in the United States. By definition, these lesions are dysplastic. The risk of malignant degeneration is related to both the size and type of polyp. Tubular adenomas are associated with malignancy in only 5% of cases, whereas villous adenomas may harbor cancer in up to 40%. Tubulovillous adenomas are at intermediate risk (22%). Invasive carcinomas are rare in
polyps smaller than 1 cm; the incidence of invasive carcinoma increases with size. The risk of carcinoma in a polyp larger than 2 cm is 35% to 50%. Although most neoplastic polyps do not evolve to cancer, most colorectal cancers originate as a polyp. It is this fact that forms the basis for secondary prevention strategies to eliminate colorectal cancer by targeting the neoplastic polyp for removal before malignancy develops.

Polyps may be pedunculated or sessile. Most pedunculated polyps are amenable to colonoscopic snare excision. Removal of sessile polyps is often more challenging. Special colonoscopic techniques, including saline lift, piecemeal snare excision, and endoscopic mucosal resection facilitate successful removal of many sessile polyps. For rectal sessile polyps, transanal operative excision is preferred because it produces an intact, single pathology specimen that can be used to determine the need for further therapy. Interpretation of the precise depth of invasion of a cancer arising in a sessile polyp after piecemeal excision is often impossible. The site of sessile polypectomies should be marked by tattoo marking to guide subsequent endoscopic surveillance and to facilitate identification of the involved bowel segment should operative resection be necessary.

Complications of polypectomy include perforation and bleeding. A small perforation (microperforation) in a fully prepared, stable patient may be managed with bowel rest, broad-spectrum antibiotics, and close observation. Signs of sepsis, peritonitis, or deterioration in clinical condition are indications for laparotomy. Bleeding may occur immediately after polypectomy or may be delayed. The bleeding will usually stop spontaneously, but colonoscopy may be required to apply endoscopic clips, resnare a bleeding stalk, cauterize the lesion, or inject/apply epinephrine. Occasionally angiography and infusion of vasopressin may be necessary. Rarely, colectomy is required.

Hyperplastic Polyps. Hyperplastic polyps are extremely common in the colon. These polyps are usually small (<5 mm) and show histologic characteristics of hyperplasia without any dysplasia. They are not considered premalignant, but they cannot be distinguished from adenomatous polyps colonoscopically and are therefore often removed. In contrast, large hyperplastic polyps (>2 cm) may have a risk of malignant degeneration. Hyperplastic polyposis is a rare disorder in which multiple large hyperplastic polyps occur in young adults. These patients are at increased risk for the development of colorectal cancer.

Serrated Polyps. Serrated polyps, including sessile serrated adenomas and traditional serrated adenomas, are a recently recognized, histologically distinct group of neoplastic polyps. Endoscopically they are flat lesions and frequently difficult to visualize. These lesions were long thought to be similar to hyperplastic polyps with minimal malignant potential. However, it has become clear that some of these polyps will develop into invasive cancers. In addition, a familial serrated polyposis syndrome has been described. Serrated polyps should be treated like adenomatous polyps.

Hamartomatous Polyps (Juvenile Polyps). In contrast to adenomatous and serrated polyps, hamartomatous polyps (juvenile polyps) usually are not premalignant. These lesions are the characteristic polyps of childhood but may occur at any age. Bleeding is a common symptom, and intussusception and/or obstruction may occur. Because the gross appearance of these polyps is identical to adenomatous polyps, these lesions should also be treated by polypectomy. In contrast to adenomatous polyposis syndromes, these conditions are often associated with mutation in BMPRIA and SMAD4.

Familial juvenile polyposis is an autosomal dominant disorder in which patients develop hundreds of polyps in the colon and rectum. Unlike solitary juvenile polyps, these lesions may degenerate into adenomas and eventually carcinoma. Annual screening should begin between the ages of 10 and 12 years. Treatment is surgical and depends in part on the degree of rectal involvement. If the rectum is relatively spared, a total abdominal colectomy with ileorectal anastomosis may be performed with subsequent close surveillance of the retained rectum. If the rectum is carpeted with polyps, total proctocolectomy is the more appropriate operation. These patients are candidates for ileal pouch–anal reconstruction to avoid a permanent stoma.

Peutz-Jeghers syndrome is characterized by polyposis of the small intestine and, to a lesser extent, polyposis of the colon and rectum. Characteristic melanin spots are often noted on the buccal mucosa and lips of these patients. The polyps of Peutz-Jeghers syndrome are generally considered to be hamartomas and are not thought to be at significant risk for malignant degeneration. However, carcinoma may occasionally develop. Because the entire length of the gastrointestinal tract may be affected, surgery is reserved for symptoms such as obstruction or bleeding or for patients in whom polyps develop adenomatous features. Screening consists of a baseline colonoscopy and upper endoscopy at age 20 years, followed by annual flexible sigmoidoscopy thereafter. Clinicians should ensure patients are screened for associated extraintestinal malignancies (breast, upper gastrointestinal tract, pancreas, cervix, ovaries, and testicles).

Cronkite-Canada syndrome is a disorder in which patients develop gastrointestinal polyposis in association with alopecia, cutaneous pigmentation, and atrophy of the fingernails and toenails. Diarrhea is a prominent symptom, and vomiting, malabsorption, and protein-losing enteropathy may occur. Most patients die of this disease despite maximal medical therapy, and surgery is reserved for complications of polyposis such as obstruction.

Cowden’s syndrome is an autosomal dominant disorder with hamartomas of all three embryonal cell layers. Facial trichilemmomas, breast cancer, thyroid disease, and gastrointestinal polyps are typical of the syndrome. Patients should be screened for cancers. Treatment is otherwise based on symptoms.

Inflammatory Polyps (Pseudopolyps). Inflammatory polyps occur most commonly in the context of inflammatory bowel disease, but they may also occur after amebic colitis, ischemic colitis, and schistosomal colitis. These lesions are not premalignant, but they cannot be distinguished from adenomatous polyps based on gross appearance and therefore should be removed. Microscopic examination shows islands of normal, regenerating mucosa (the polyp) surrounded by areas of mucosal loss. Polyposis may be extensive, especially in patients with severe colitis, and may mimic FAP.

Inherited Colorectal Carcinoma

Many of the genetic defects originally described in hereditary cancers have subsequently been found in sporadic tumors. Although the majority of colorectal cancer is sporadic, several hereditary syndromes provide paradigms for the study of this disease. Insight gained from studying inherited colorectal cancer
syndromes has led to better understanding of the genetics of colorectal carcinoma.

**Familial Adenomatous Polyposis.** This rare autosomal dominant condition accounts for only about 1% of all colorectal adenocarcinomas. Nevertheless, this syndrome has provided tremendous insight into the molecular mechanisms underlying colorectal carcinogenesis. The genetic abnormality in FAP is a mutation in the *APC* gene, located on chromosome 5q. Of patients with FAP, *APC* mutation testing is positive in 75% of cases. While most patients with FAP will have a known family history of the disease, up to 25% present without other affected family members. Clinically, patients develop hundreds to thousands of adenomatous polyps shortly after puberty. The lifetime risk of colorectal cancer in FAP patients approaches 100% by age 50 years.

Flexible sigmoidoscopy of first-degree relatives of FAP patients beginning at age 10 to 15 years has been the traditional mainstay of screening. Today, following genetic counseling, *APC* gene testing may be used to screen family members, providing an *APC* mutation has been identified. If *APC* testing is positive in a relative of a patient with a known *APC* mutation, annual flexible sigmoidoscopy beginning at age 10 to 15 years is done until polyps are identified. If *APC* testing is negative, the relative can be screened starting at age 50 years per average-risk guidelines. If *APC* testing is refused or unavailable, or if a mutation cannot be identified, annual flexible sigmoidoscopy beginning at age 10 to 15 years is performed until age 24 years. Screening flexible sigmoidoscopy is then done every 2 years until age 34 years, every 3 years until age 44 years, and then every 3 to 5 years.

FAP patients are also at risk for the development of adenomas anywhere in the gastrointestinal tract, particularly in the duodenum. Periampullary carcinoma is a particular concern. Upper endoscopy is therefore recommended for surveillance every 1 to 3 years beginning at age 25 to 30 years.

Once the diagnosis of FAP has been made and polyps are developing, treatment is surgical. Four factors affect the choice of operation: age of the patient; presence and severity of symptoms; extent of rectal polyposis; and presence and location of cancer or desmoid tumors. Three operative procedures can be considered: total proctocolectomy with an end (Brooke) ileostomy; total abdominal colectomy with ileorectal anastomosis; and restorative proctocolectomy with ileal pouch–anal anastomosis (IPAA) with or without a temporary ileostomy. Most patients elect to have an ileal pouch–anal anastomosis in the absence of a distal rectal cancer, a mesenteric desmoid tumor that prevents the ileum from reaching the anus, or poor sphincter function. Mucosectomy has been advocated in patients with FAP undergoing ileal pouch–anal anastomosis because of the risk of neoplasia in the anal transition zone, but the requirement for this procedure remains controversial. Although patient satisfaction with IPAA remains high, function may not be ideal, and up to 50% of patients experience some degree of incontinence. Total abdominal colectomy with an ileorectal anastomosis is also an option in these patients, but requires vigilant surveillance of the retained rectum for development of rectal cancer. There is increasing data suggesting that the administration of cyclooxygenase-2 (COX-2) inhibitors (celecoxib, sulindac) may slow or prevent the development of polyps.

FAP may be associated with extraintestinal manifestations such as congenital hypertrophy of the retinal pigment epithelium, desmoid tumors, epidermoid cysts, mandibular osteomas (Gardner’s syndrome), and central nervous system tumors (Turcot’s syndrome). Although they arise in a minority of FAP patients, desmoid tumors in particular can make surgical management difficult and are a source of major morbidity and mortality in these patients. These lesions arise from fibroblasts, and although technically benign, can be highly locally invasive. These tumors often involve the bowel mesentery, pelvis, and abdominal wall. In some cases, surgery is thought to be the “trigger.” Local recurrence after attempted resection is common; therefore, surgery is to be avoided if at all possible. Desmoid tumors are often hormone responsive, and growth may be inhibited in some patients with tamoxifen. COX-2 inhibitors and nonsteroidal, anti-inflammatory drugs may also be beneficial in this setting. A recent report suggests that imatinib may also be effective.

**Attenuated Familial Adenomatous Polyposis.** AFAP is a recognized variant of FAP. Patients present later in life with fewer polyps (usually 10–100) predominantly located in the right colon, when compared to classic FAP. Colorectal carcinoma develops in more than 50% of these patients, but occurs later (average age, 55 years). Patients are also at risk for duodenal polyposis. However, in contrast to FAP, *APC* gene mutations are present in only about 30% of patients with AFAP. When present, these mutations are expressed in an autosomal dominant pattern.

Mutations in *MYH* also result in the AFAP phenotype but are expressed in an autosomal recessive pattern. It has been suggested that *MYH* mutations may be responsible for AFAP in patients who do not have a detectable *APC* gene mutation.

Genetic testing is often offered to patients with suspected AFAP. When positive, genetic counseling and testing may be used to screen at-risk family members. If the family mutation is unknown, screening colonoscopy is recommended beginning at age 13 to 15 years, then every 4 years to age 28 years, and then every 3 years. These patients are often candidates for a total abdominal colectomy with ileorectal anastomosis because the limited polyposis in the rectum can usually be treated by colonoscopic snare excision. Prophylaxis with COX-2 inhibitors also may be appropriate. Because of the more subtle phenotype in these patients, it is important to rule out other familial syndromes such as HNPCC (Lynch syndrome) and the more common familial colorectal cancer.

**Lynch Syndrome (Hereditary Nonpolyposis Colon Cancer; HNPCC).** Lynch syndrome is more common than FAP, but it is still extremely rare (1–3% of all colon cancers). The genetic defects associated with Lynch syndrome arise from errors in mismatch repair, the phenotypic result being MSI. Lynch syndrome is inherited in an autosomal dominant pattern and is characterized by the development of colorectal carcinoma at an early age (average age, 40–45 years). Approximately 70% of affected individuals will develop colorectal cancer. Cancers appear in the proximal colon more often than in sporadic colorectal cancer and have a better prognosis regardless of stage. The risk of synchronous or metachronous colorectal carcinoma is 40%. Lynch syndrome may also be associated with extracolonic malignancies, including endometrial carcinoma, which is most common in women, and ovarian, pancreas, stomach, small bowel, biliary, and urinary tract carcinomas. The diagnosis is made based on family history. The Amsterdam I criteria for clinical diagnosis of Lynch syndrome are three affected relatives...
with histologically verified adenocarcinoma of the large bowel (one must be a first-degree relative of one of the others) in two successive generations of a family with one patient diagnosed before age 50 years. The presence of other related carcinomas should raise the suspicion of this syndrome. Revised criteria Amsterdam II requires three or more relatives with an HNPCC related malignancy in which at least one is a first degree relative of the others, two generations are affected, at least one cancer occurred before age 50, FAP has been excluded, and pathology of the tumors has been reviewed and confirmed. In a patient with an established diagnosis of colorectal cancer, tumor testing for presence of mismatch repair gene products (immunohistochemistry) and/or MSI can sometimes serve as screening for this syndrome.\(^\text{10,93,94}\)

Lynch syndrome results from mutations in mismatch repair genes, and like FAP, specific mutations are associated with different phenotypes. For example, mutations in \textit{PMS2} or \textit{MSH6} result in a more attenuated form of Lynch syndrome when compared to mutations in other genes. \textit{MSH6} inactivation also appears to be associated with a higher risk for endometrial cancer. Further significance of these specific mutations remains to be determined.

Screening colonoscopy is recommended annually for at-risk patients beginning at either age 20 to 25 years or 10 years younger than the youngest age at diagnosis in the family, whichever comes first. Because of the high risk of endometrial carcinoma, transvaginal ultrasound or endometrial aspiration biopsy is also recommended annually after age 25 to 35 years. Because there is a 40% risk of developing a second colon cancer, total colectomy with ileorectal anastomosis is recommended once adenomas or a colon carcinoma is diagnosed. Annual proctoscopy is necessary because the risk of developing rectal cancer remains high. Similarly, prophylactic hysterectomy and bilateral salpingo-oophorectomy should be considered in women who have completed childbearing.\(^\text{93-95}\)

\textbf{Familial Colorectal Cancer.} Nonsyndromic familial colorectal cancer accounts for 10% to 15% of patients with colorectal cancer. The lifetime risk of developing colorectal cancer increases with a family history of the disease. The lifetime risk of colorectal cancer in a patient with no family history of this disease (average-risk population) is approximately 6%, but rises to 12% if one first-degree relative is affected and to 35% if two first-degree relatives are affected. Age of onset also impacts risk, and a diagnosis before the age of 50 years is associated with a higher incidence in family members. Screening colonoscopy is recommended every 5 years beginning at age 40 years or beginning 10 years before the age of the earliest diagnosed patient in the pedigree. While there are no specific genetic abnormalities that are associated with familial colorectal cancer, any of the defects found in either the LOH pathway or MSI pathway may be present in these patients.

\textbf{Prevention: Screening and Surveillance}

Because the majority of colorectal cancers are thought to arise from adenomatous polyps, preventive measures focus on identification and removal of these premalignant lesions. In addition, many cancers are asymptomatic, and screening may detect these tumors at an early and curable stage (Table 29-1). Although screening for colorectal cancer decreases the incidence of cancer and cancer-related mortality, the optimal method of screening remains controversial. Screening guidelines are meant for asymptomatic patients.\(^\text{95-98}\) Any patient with a gastrointestinal

\textbf{Fecal Occult Blood Testing and Fecal Immunohistochemical Testing.} FOBT is known to reduce colorectal cancer mortality by 33% and metastatic disease by 50%. However, FOBT is relatively insensitive, missing up to 50% of cancers and the majority of adenomas. Its specificity is low because 90% of patients with positive tests do not have colorectal cancer. FIT is more sensitive and specific for cancer. Mortality benefits for its use are inferred from FOBT literature. Compliance with annual testing is low and costs are significant if one includes the colonoscopy examinations done to evaluate patients with positive FOBT/FIT. Nonetheless, the direct evidence that FOBT screening is efficacious and decreases both the incidence and mortality of colorectal cancer is so strong that national guidelines recommend annual FOBT/FIT screening for asymptomatic, average-risk Americans older than 50 years of age as one of several accepted strategies. A positive FOBT/FIT should be followed by colonoscopy.\(^\text{97-100}\)

\textbf{Stool DNA.} Neoplastic lesions of the colon shed cells into the lumen posing an opportunity for detection via DNA testing. A commercially available multigene stool DNA test evaluates stool samples for mutant \textit{KRAS}, methylated \textit{BMP3} and the promoter region of \textit{NDRG4}. In a large North American prospective study, this test has recently been found to be 92% sensitive for detection of colorectal cancer. Compared to FIT, stool DNA testing has a lower specificity (74%) raising concerns about how to manage stool DNA-positive patients who have a negative colonoscopic evaluation. Sensitivity for advanced precancerous lesions was 42\%.\(^\text{100,101}\) This test, in combination with FIT, is supported as a screening modality every 1 to 3 years by the U.S. Preventative Task Force, and every 3-year utilization is supported by NCCN guidelines.\(^\text{100}\) Nevertheless, additional studies will be necessary to determine if these tests are comparable or superior to more traditional methods has been techniques.

\textbf{Flexible Sigmoidoscopy.} Screening by flexible sigmoidoscopy every 5 years may lead to a 60% to 70% reduction in mortality from colorectal cancer, chiefly by identifying high-risk individuals with adenomas. However, it is important to recognize that lesions in the proximal colon cannot be identified, and for this reason, flexible sigmoidoscopy has often been paired with \textit{air-contrast barium enema} to detect transverse and right colon lesions. Patients found to have a polyp, cancer, or other lesion on flexible sigmoidoscopy will require colonoscopy.\(^\text{101-102}\)

\textbf{Fecal Occult Blood Testing and Flexible Sigmoidoscopy.} Several trials have shown that FOBT screening is least effective at detecting rectosigmoid cancers.\(^\text{97-99}\) This is precisely the area screened by flexible sigmoidoscopy; thus, the combination of the two tests has been suggested as a reasonable screening strategy. Winawer and colleagues, in a study of 12,479 subjects, showed that the combination of FOBT annually with flexible sigmoidoscopy every 5 years resulted in lower mortality from colorectal cancer and better survival in patients with colorectal cancer.\(^\text{99}\) A similar benefit was confirmed in long-term (11-year) follow-up from the Norwegian Colorectal Cancer Prevention Trial. Such data led to the American Cancer Society recommendations that one of the acceptable screening regimens for average-risk Americans is the combination of FOBT/FIT annually and flexible sigmoidoscopy every 5 years; this combination was preferred over either test alone. Recent NCCN guidelines offer the option of flexible sigmoidoscopy with stool-based testing.
every 10 years. The addition of air-contrast barium enema to assess the proximal colon may improve sensitivity as well.96

**Colonoscopy.** Colonoscopy is currently the most accurate and most complete method for examining the large bowel. This procedure is highly sensitive for detecting even small polyps (<1 cm) and allows biopsy, polypectomy, control of hemorrhage, and dilation of strictures. However, colonoscopy does require mechanical bowel preparation, and the discomfort associated with the procedure requires conscious sedation in most patients. Colonoscopy is also considerably more expensive than other screening modalities and requires a well-trained endoscopist. The risk of a major complication after colonoscopy (perforation and hemorrhage) is extremely low (0.2–0.3%). Nevertheless, deaths have been reported.

**Air-Contrast Barium Enema.** Air-contrast barium enema is also highly sensitive for detecting polyps greater than 1 cm in diameter (90% sensitivity). Unfortunately, there are no studies proving its efficacy for screening large populations. Accuracy is greatest in the proximal colon but may be compromised in the sigmoid colon if there is significant diverticulosis. The major disadvantages of barium enema are the need for mechanical bowel preparation and the requirement for colonoscopy if a lesion is discovered.

**Computed Tomography Colonography (Virtual Colonoscopy).** Advances in imaging technology have created a number of less invasive, but highly accurate tools for screening. CT colonography makes use of helical CT technology and three-dimensional reconstruction to image the intraluminal colon. At present, patients require a mechanical bowel preparation. The colon is then insufflated with air, a spiral CT is performed, and both two-dimensional and three-dimensional images are generated. In the hands of a qualified radiologist, sensitivity appears to be as good as colonoscopy for colorectal cancers and polyps greater than 1 cm in size.103 Colonoscopy is required if a lesion is identified. CT colonography has also been used for imaging the proximal colon in cases of obstruction or if a colonoscopy cannot be completed in selected patients. Limitations of this technique include false-positive results from retained stool, diverticular disease, haustral folds, motion artifacts, and an inability to detect flat adenomas.

**Guidelines for Screening.** Current American Cancer Society guidelines advocate screening for the average-risk population (asymptomatic, no family history of colorectal carcinoma, no personal history of polyps or colorectal carcinoma, no familial syndrome) beginning at age 50 years. Recommended procedures include yearly FOBT/FIT, flexible sigmoidoscopy every...
5 years, FOBT/FIT and flexible sigmoidoscopy in combination, air-contrast barium enema every 5 years, or colonoscopy every 10 years. Patients with other risk factors should be screened earlier and more frequently (Table 29-2).93-96,104,105

### Routes of Spread and Natural History

Carcinoma of the colon and rectum arises in the mucosa. The tumor subsequently invades the bowel wall and eventually adjacent tissues and other viscera. Tumors may become bulky and circumferential, leading to colon obstruction. Local invasion (especially in the rectum or sigmoid colon) may occasionally cause obstruction of other organs such as the ureter.

Regional lymph node involvement is the most common form of spread of colorectal carcinoma and usually precedes distant metastasis or the development of carcinomatosis. The likelihood of nodal metastasis increases with tumor size, poorly differentiated histology, lymphovascular invasion, and depth of invasion. The T stage (depth of invasion) is the single most significant predictor of lymph node spread. Carcinoma in situ (Tis) in which there is no penetration of the muscularis mucosa (basement membrane) has also been called high-grade dysplasia and should carry no risk of lymph node metastasis. Small lesions confined to the bowel wall (T1 and T2) are associated with lymph node metastasis in 5% to 20% of cases, whereas larger tumors that invade through the bowel wall or into adjacent organs (T3 and T4) are likely to have lymph node metastasis in more than 50% of cases. The number of lymph nodes with metastases correlates with the presence of distant disease and inversely with survival. Four or more involved lymph nodes (N2 disease) predict a poor prognosis. In colon cancer, lymphatic spread usually follows the major venous outflow from the involved segment of the colon. Lymphatic spread from the rectum follows two routes. In the upper rectum, drainage ascends along the superior rectal vessels to the inferior mesenteric nodes. In the lower rectum, lymphatic drainage may course along the middle rectal vessels. Nodal spread along the inferior rectal vessels to the internal iliac nodes or groin is rare unless the tumor involves the anal canal or the proximal lymphatics are blocked with tumor (Fig. 29-23).

The most common site of distant metastasis from colorectal cancer is the liver. These metastases arise from hematogenous spread via the portal venous system. Like lymph node metastasis, the risk of hepatic metastasis increases with tumor size and tumor grade. However, even small tumors may produce distant metastasis. The lung is also a site of hematogenous spread, but this rarely occurs in isolation. Carcinomatosis (diffuse peritoneal metastases) occurs by peritoneal seeding and has a dismal prognosis.

### Staging and Preoperative Evaluation

#### Clinical Presentation

Symptoms of colon and rectal cancers are nonspecific and generally develop when the cancer is locally advanced. The classic first symptoms are a change in bowel habits and rectal bleeding. Abdominal pain, bloating, and other signs of obstruction typically occur with larger tumors and
suggest more advanced disease. Because of the caliber of the bowel and the consistency of the stool, left-sided tumors are more likely to cause obstruction than are right-sided tumors. Rectal tumors may cause bleeding, tenesmus, and pain. However, it is important to note that many patients may be asymptomatic and/or present with unexplained anemia, weight loss, or poor appetite.

**Staging.** Colorectal cancer staging is based on tumor depth and the presence or absence of nodal or distant metastases. Older staging systems, such as the Dukes’ Classification and its Astler-Coller modification, have been replaced by the tumor-node-metastasis (TNM) staging system described by the American Joint Committee on Cancer (AJCC). The AJCC TNM classification has recently been updated to reflect survival outcomes based upon the Surveillance Epidemiology and End Results (SEER) registry (Table 29-3). Stage I disease includes adenocarcinomas that are invasive through the muscularis mucosa but are confined to the submucosa (T1) or the muscularis propria (T2) in the absence of nodal metastases. Stage II disease consists of tumors that invade through the bowel wall into the subserosa or nonperitonealized pericolic or perirectal tissues (T3) or into other organs or tissues or through the visceral peritoneum (T4) without nodal metastases. Stage III disease includes any T stage with nodal metastases, and stage IV disease denotes distant metastases.

The preoperative imaging evaluation usually identifies stage IV disease. In colon cancer, differentiating stages I, II, and III depends on histologic examination of the resected specimen. In rectal cancer, endorectal ultrasound or MRI may predict the stage (ultrasound stage, uTxNx) preoperatively, but the final determination depends on pathologic examination of the resected tumor and adjacent lymph nodes (pathologic stage, pTxNx). Disease stage correlates with 5-year survival. Patients with stages I and II disease can expect excellent survival rates. The presence of nodal metastases (stage III) decreases survival. In rectal cancer, staging has been further refined, and outcomes suggest that subgroups of patients within each stage may have very different prognoses. If the mesorectum around a rectal cancer is involved or threatened (only 1–2 mm of clearance), there is a very high likelihood of local recurrence and a poor prognosis. This circumferential or radial margin is probably best assessed preoperatively by MRI. Although nodal involvement is the single most important prognostic factor in colorectal carcinoma, tumor characteristics, such as degree of differentiation, mucinous or signet-ring cell histology, vascular invasion, and DNA aneuploidy, also adversely affect prognosis. Preoperative CEA also has been suggested to be a prognostic indicator. Molecular profiling is currently being studied in an effort to further improve prognostic indicators. The 5-year survival rate with stage IV disease is low. However, in well-selected patients, metastasectomy, especially of isolated liver or lung lesions, can result in cure. In these patients, the involvement of a multidisciplinary team and/or tumor board is highly recommended (Fig. 29-24).

**Preoperative Evaluation.** Once a colon or rectal carcinoma has been diagnosed, a staging evaluation should be undertaken. The colon must be evaluated for synchronous tumors, usually by colonoscopy. Synchronous disease will be present in up to 5% of patients. For rectal cancers, digital rectal examination and rigid or flexible proctoscopy with biopsy should be performed to assess tumor size, location, morphology, histology, and fixation. Endorectal ultrasound or MRI can be invaluable in staging rectal cancer and is used to classify the ultrasound T and N stage of rectal cancers (see Fig. 29-24). A chest/abdominal/pelvic CT scan should be obtained to evaluate for distant metastases. Pelvic CT scan, and sometimes MRI, can be useful in large rectal tumors and in recurrent disease to determine the extent of local invasion. Among patients with obstructive symptoms, a water-soluble contrast study (Gastrografin enema) may be useful for delineating the degree of obstruction. It is important to avoid mechanical bowel preparation (for either colonoscopy or surgery) in a patient who appears to be obstructed. PET scan may be useful in evaluating lesions seen on CT scan and in patients in whom a risky or highly morbid operation is planned (pelvic exenteration, sacrectomy). Preoperative CEA is often obtained and may be useful for postoperative follow-up.

**Therapy for Colonic Carcinoma**

**Principles of Resection.** The objective in treatment of carcinoma of the colon is to remove the primary tumor along with its lymphovascular supply. Because the lymphatics of the colon accompany the main arterial supply, the length of bowel resected depends on which vessels are supplying the segment involved with the cancer. Any adjacent organ or tissue, such as the omentum, that has been invaded should be resected en bloc with the tumor. If all of the tumor cannot be removed, a palliative procedure should be considered, although it important to note that “debulking” is rarely effective in colorectal adenocarcinoma.

The presence of synchronous cancers or adenomas or a strong family history of colorectal neoplasms suggests that the entire colon is at risk for carcinoma (often called a field defect), and a subtotal or total colectomy should be considered. Metachronous tumors (a second primary colon cancer) identified during follow-up studies should be treated similarly. However, the surgeon must be aware of which mesenteric vessels have been ligated at the initial studies because this may influence the viability of the remaining colon and the choice of procedure.

The number of lymph nodes recovered in the surgical specimen has long served as a proxy for the oncologic adequacy of resection. A number of studies previously have suggested that...
<table>
<thead>
<tr>
<th>Stage</th>
<th>Tumor Stage (T)</th>
<th>Nodal Stage (N)</th>
<th>Distant Metastasis (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Tis</td>
<td>N0</td>
<td>M0</td>
</tr>
<tr>
<td>I</td>
<td>T1, T2</td>
<td>N0</td>
<td>M0</td>
</tr>
<tr>
<td>IIA</td>
<td>T3</td>
<td>N0</td>
<td>M0</td>
</tr>
<tr>
<td>IIB</td>
<td>T4a</td>
<td>N0</td>
<td>M0</td>
</tr>
<tr>
<td>IIC</td>
<td>T4b</td>
<td>N0</td>
<td>M0</td>
</tr>
<tr>
<td>IIIA</td>
<td>T1–T2</td>
<td>N1/N1c</td>
<td>M0</td>
</tr>
<tr>
<td>IIIA</td>
<td>T1</td>
<td>N2a</td>
<td>M0</td>
</tr>
<tr>
<td>IIIA</td>
<td>T3–T4a</td>
<td>N1/N1c</td>
<td>M0</td>
</tr>
<tr>
<td>IIIB</td>
<td>T2–T3</td>
<td>N2a</td>
<td>M0</td>
</tr>
<tr>
<td>IIIB</td>
<td>T1–T2</td>
<td>N2b</td>
<td>M0</td>
</tr>
<tr>
<td>IIIC</td>
<td>T4a</td>
<td>N2a</td>
<td>M0</td>
</tr>
<tr>
<td>IIIC</td>
<td>T3–T4a</td>
<td>N2b</td>
<td>M0</td>
</tr>
<tr>
<td>IIVA</td>
<td>Any T</td>
<td>Any N</td>
<td>M1a</td>
</tr>
<tr>
<td>IVB</td>
<td>Any T</td>
<td>Any N</td>
<td>M1b</td>
</tr>
<tr>
<td>IVC</td>
<td>Any T</td>
<td>Any N</td>
<td>M1c</td>
</tr>
</tbody>
</table>

**Tumor Stage (T)**
- **TX**: Primary tumor cannot be assessed
- **T0**: No evidence of primary tumor
- **Tis**: Carcinoma *in situ*, intramucosal carcinoma (involvement of lamina propria with no extension through muscularis mucosae)
- **T1**: Tumor invades the submucosa (through the muscularis mucosa but not into the muscularis propria)
- **T2**: Tumor invades the muscularis propria
- **T3**: Tumor invades through the muscularis propria into pericolorectal tissues
- **T4**: Tumor invades the visceral peritoneum or invades or adheres to adjacent organ or structure
  - **T4a**: Tumor invades through the visceral peritoneum (including gross perforation of the bowel through tumor and continuous invasion of tumor through areas of inflammation to the surface of the visceral peritoneum)
  - **T4b**: Tumor directly invades or adheres to adjacent organs or structures

**Nodal Stage (N)**
- **NX**: Regional lymph nodes cannot be assessed
- **N0**: No regional lymph node metastasis
- **N1**: One to three regional lymph nodes are positive (tumor in lymph nodes measuring ≥0.2 mm), or any number of tumor deposits is present and all identifiable lymph nodes are negative
  - **N1a**: One regional lymph node is positive
  - **N1b**: Two or three regional lymph nodes are positive
  - **N1c**: No regional lymph nodes are positive, but there are tumor deposits in the
    - subserosa
    - mesentery
    - or nonperitonealized pericolonic, or perirectal/mesorectal tissues.
- **N2**: Four or more regional nodes are positive
  - **N2a**: Four to six regional lymph nodes are positive
  - **N2b**: Seven or more regional lymph nodes are positive

**Distant Metastasis (M)**
- **M0**: No distant metastasis by imaging, etc.; no evidence of tumor in distant sites or organs (This category is not assigned by pathologists.)
- **M1**: Metastasis to one or more distant sites or organs or peritoneal metastasis is identified
  - **M1a**: Metastasis to one site or organ is identified without peritoneal metastasis
  - **M1b**: Metastasis to two or more sites or organs is identified without peritoneal metastasis
  - **M1c**: Metastasis to the peritoneal surface is identified alone or with other site or organ metastases
a minimum of 12 lymph nodes in the resected specimen are necessary for adequate staging. In addition, patients in whom more nodes are harvested have better long-term outcome. As such, a 12-node minimum has been suggested as an appropriate benchmark for assessing quality of care. However, several investigators recently have called this into question, noting that the number of lymph nodes examined does not correlate with staging, use of adjuvant chemotherapy, or patient survival. Others have suggested that the number of negative lymph nodes and/or the lymph node ratio (positive lymph nodes to total lymph nodes) may further improve staging.  

If unexpected metastatic disease is encountered at the time of a laparotomy, the decision about whether to proceed with resection of the primary tumor depends on the volume of distant disease, location and size of the primary tumor, the operation required to remove the primary tumor, and the operative approach. If the metastatic disease is low volume (isolated or potentially resectable liver lesions) and the resection of the primary tumor is straightforward (segmental abdominal colectomy), it is probably reasonable to proceed with resection. On the other hand, if the metastatic disease is high volume (metastases), especially if the primary tumor is minimally symptomatic, the operation should be aborted in order to facilitate early systemic chemotherapy. Some centers favor starting the operation with a diagnostic laparoscopy in cases where risk of discovering metastasis is high in order to minimize the magnitude of the operation should surgery be aborted. With recent advances in chemotherapy, many of these patients will never develop a complication from the primary tumor requiring surgical intervention. Other palliative approaches include a bypass or proximal stoma for obstructing lesions.

Stage-Specific Therapy

Stage 0 (Tis, N0, M0) Polyps containing carcinoma in situ (high-grade dysplasia) carry no risk of lymph node metastasis. However, the presence of high-grade dysplasia increases the risk of finding an invasive carcinoma within the polyp. For this reason, these polyps should be excised completely, and pathologic margins should be free of dysplasia. Most pedunculated polyps and many sessile polyps may be completely removed endoscopically. These patients should be followed with frequent colonoscopy to ensure that the polyp has not recurred and that an invasive carcinoma has not developed. In cases where the polyp cannot be removed entirely, a segmental resection is recommended.

Stage I: The Malignant Polyp (T1, N0, M0) Occasionally a polyp that was thought to be benign will be found to harbor invasive carcinoma after polypectomy. Treatment of a malignant polyp is based on the risk of local recurrence and the risk of lymph node metastasis. The risk of lymph node metastases depends primarily on the depth of invasion. Invasive carcinoma in the head of a pedunculated polyp with no stalk involvement carries a low risk of metastasis (<1%) and may be completely resected endoscopically. For sessile polyps, the depth of invasion predicts risk of lymphovascular spread. A recent classification stratifies risk by depth of submucosal spread. Superficial lesions (submucosa 1; Sm1) are low risk, whereas Sm2 and Sm3 are intermediate and high risk. Lymphovascular invasion, poorly differentiated histology, tumor budding, or tumor within 1 mm of the resection margin greatly increases the risk of local recurrence and metastatic spread. Segmental colectomy is then indicated. Invasive carcinoma arising in a sessile polyp extending into the submucosa is and is usually best treated with segmental colectomy (Fig. 29-25).

Stage I and II: Localized Colon Carcinoma (T1-3, N0, M0) The majority of patients with stages I and II colon cancer will be cured with surgical resection. Few patients with completely resected stage I disease will develop either local or distant recurrence, and adjuvant chemotherapy does not improve survival in these patients. However, up to 46% of patients with completely resected stage II disease will ultimately die from colon cancer. For this reason, adjuvant chemotherapy has been suggested for selected patients with stage II disease (young patients, tumors with “high-risk” histologic findings). It remains controversial as to whether chemotherapy improves survival rates in these patients. In some cases, molecular profiling may predict prognosis, although it is important to note that these tools have not been shown to predict response to therapy. At present,
molecular profiling for selecting patients to receive chemotherapy remains unproven.

**Stage III: Lymph Node Metastasis (Tany, N1, M0)** Patients with lymph node involvement are at significant risk for both local and distant recurrence, and adjuvant chemotherapy has been recommended routinely in these patients. 5-Fluorouracil–based regimens (with leucovorin) and oxaliplatin (FOLFOX) reduce recurrences and improve survival in this patient population. It is important to note, however, that a subgroup of patients with stage III disease will do well without chemotherapy. MSI status in particular predicts good prognosis. Subset analysis from the CRYSTAL trial has shown that patients with MSI-high stage III disease do not benefit from 5-fluorouracil–based chemotherapy. Molecular profiling, therefore, may be helpful in determining which stage III patients can safely avoid systemic chemotherapy.\(^\text{118}\)

**Stage IV: Distant Metastasis (Tany, Nany, M1)** Survival is extremely limited in stage IV colon carcinoma. Systemic chemotherapy is recommended in almost all cases of distant spread. However, unlike many other malignancies, highly selected patients with isolated, resectable metastases may benefit from resection (metastasectomy). The most common site of metastasis is the liver. Of patients with systemic disease, approximately 15% will have metastases limited to the liver. Of these, 20% are potentially resectable for cure. Survival is improved in these patients (20–40% 5-year survival) when compared to patients who do not undergo resection. Hepatic resection of synchronous metastases from colorectal carcinoma may be performed as a combined procedure or in two stages. The second most common site of metastasis is the lung, occurring in approximately 20% of patients with colorectal carcinoma. Although very few of these patients will be potentially resectable, among those who are (about 1–2% of all colorectal cancer patients), long-term survival benefit is approximately 30% to 40%. There are limited reports of successful resection of metastases in other sites (ovary and retroperitoneum are most common). Cytoreductive surgery and intraperitoneal chemotherapy (HIPEC) has been suggested for patients with carcinomatosis, but remains unproven for colorectal cancer and carries high morbidity.\(^\text{119,120}\)

The remainder of patients with stage IV disease cannot be cured surgically, and therefore, the focus of treatment should be palliation. Methods such as colonic stenting for obstructing lesions of the left colon also provide good palliation. More limited surgical intervention such as a diverting stoma or bypass procedure may be appropriate in patients with stage IV disease who develop obstruction. Hemorrhage in an unresectable tumor can sometimes be controlled with angiographic embolization. External beam radiation also has been used for palliation. The involvement of a palliative care team in the management of these patients is critical.\(^\text{116}\)

**Therapy for Rectal Carcinoma**

**Principles of Resection.** The biology of rectal adenocarcinoma is thought to be similar to the biology of colonic adenocarcinoma, and the operative principles of complete resection of the primary tumor, its lymphatic bed, and any other involved organ apply to surgical resection of rectal carcinoma. However, the anatomy of the pelvis and proximity of other structures (ureters, bladder, prostate, vagina, iliac vessels, and sacrum) make resection more challenging and often require a different approach than for colonic adenocarcinoma. Moreover, it is more difficult to achieve negative radial margins in rectal cancers that extend through the bowel wall because of the anatomic limitations of the pelvis. Therefore, local recurrence is higher than with similar stage colon cancers. However, unlike the intraperitoneal colon, the relative paucity of small bowel and other radiation-sensitive structures in the pelvis makes it easier to treat rectal tumors with radiation. Therapeutic decisions, therefore, are based on the location and depth of the tumor and its relationship to other structures in the pelvis.

**Local Therapy.** The distal 10 cm of the rectum are accessible transanally. For this reason, several local approaches have been proposed for treating rectal neoplasms. *Transanal excision* (full thickness or mucosal) is an excellent approach for noncircumferential, benign, villous adenomas of the rectum. *Transanal endoscopic microsurgery* (TEM) and *transanal minimally invasive surgery* (TAMIS) make use of a specially designed proctoscope, magnifying system, and instruments similar to those used in laparoscopy to allow local excision of lesions higher in the rectum (up to 15 cm). Although this technique has been used for selected T1, and some T2, carcinomas, local excision does not allow pathologic examination of the lymph nodes and might therefore understage patients. Moreover, local recurrence rates are high after transanal excision, and salvage surgery, while often curative, has been reported to be associated with poorer survival than with initial radical surgery. Current recommendation is to limit local excision of T1 lesions to patients with well to moderately differentiated small lesions (<3 cm) and/or in patients medically unfit for radical resection. In general, local excision of any rectal neoplasm should be considered an *excisional biopsy* because final pathologic examination of the specimen may reveal an invasive carcinoma that then mandates more radical therapy.\(^\text{120,121}\)

Ablative techniques, such as electrocautery or endocavitary radiation, also have been used. The disadvantage of these techniques is that no pathologic specimen is retrieved to confirm the tumor stage. Fulguration is generally reserved for extremely high-risk, symptomatic patients with a limited life span who cannot tolerate more radical surgery.\(^\text{120}\)

**Radical Resection.** Radical resection is preferred to local therapy for most rectal carcinomas. Radical resection involves removal of the involved segment of the rectum along with its lymphovascular supply. Although any microscopically negative margin has been suggested to be adequate, most surgeons

---

*Figure 29-25.* Levels of invasive carcinoma in pedunculated and sessile polyps. ca = carcinoma.
SPECIFIC CONSIDERATIONS
PART II
still attempt to obtain a 2-cm distal mural margin for curative resections.

Total mesorectal excision (TME) is a technique that uses sharp dissection along anatomic planes to ensure complete resection of the rectal mesentery during low and extended low anterior resections. For upper rectal or rectosigmoid resections, a partial mesorectal excision of at least 5 cm distal to the tumor appears adequate. TME both decreases local recurrence rates and improves long-term survival rates. Moreover, this technique is associated with less blood loss and less risk to the pelvic nerves and presacral plexus than is blunt dissection. The principles of TME should be applied to all radical resections for rectal cancer.

Recurrence of rectal cancer generally has a poor prognosis. Extensive involvement of other pelvic organs (usually occurring in the setting of tumor recurrence) may require a pelvic exenteration. The rectal and perineal portions of this operation are similar to an APR, but en bloc resection of the ureters, bladder, and prostate or uterus and vagina are also performed. A permanent colostomy and an ileal conduit to drain the urinary tract may be necessary. The sacrum may also be resected if necessary (sacrectomy) up to the level of the S2-S3 junction. These operations are best performed in tertiary centers with multidisciplinary teams consisting of a colon and rectal surgeon, urologist, neurosurgeon, and plastic surgeon.

Stage-Specific Therapy (Fig. 29-26). Pretreatment staging of rectal carcinoma often relies on endorectal ultrasound or MRI to determine the T and N status of a rectal cancer. Ultrasound is highly accurate at assessing tumor depth, but it is less accurate in diagnosing nodal involvement. Ultrasound evaluation can guide choice of therapy in most patients. MRI is useful to assess mesorectal involvement. When the radial margin is threatened or involved, neoadjuvant chemoradiation is recommended.122

Stage 0 (Tis, N0, M0) Villous adenomas harboring carcinoma in situ (high-grade dysplasia) are ideally treated with local excision. A 1-cm margin should be obtained. Rarely, radical resection will be necessary if transanal excision is not technically possible (large circumferential lesions).

Stage I: Localized Rectal Carcinoma (T1-2, N0, M0) Although local excision has been used for small, favorable sessile uT1N0 and uT2N0 rectal cancers, local recurrence rates may be as high as 20% and 40%, respectively. Local excision increasingly is offered to patients with small, low-risk lesions, but it does not allow physicians to assess regional lymph nodes. For this reason, radical resection is recommended in all good-risk patients. Lesions with unfavorable histologic characteristics and those located in the distal third of the rectum, in particular, are prone to recurrence. In high-risk patients and in patients who refuse radical surgery because of the risk of need for a permanent colostomy, local excision may be adequate, but strong consideration should be given to adjuvant or neoadjuvant chemoradiation to improve local control. The efficacy of adjuvant or neoadjuvant chemoradiation followed by transanal excision in patients who can tolerate radical surgery has been

Figure 29-26. Diagnostic algorithm for rectal cancer. CT = computed tomography; MRI = magnetic resonance imaging; U/S = ultrasound.
controversial. Early results from ACOSOG Z6041, in which patients with T2 rectal cancers received neoadjuvant chemoradiation followed by transanal excision, showed a pathologic complete response rate of 44%. At 3 years, disease-free survival was 88%, which is comparable to cancer outcomes after a formal resection. However, population-based data suggests that survival after local excision for rectal cancer is suboptimal and should not be offered as a matter of routine.

**Locally Advanced Rectal Cancer (Stages II and III)**

**Stage II: Localized Rectal Carcinoma (T3-4, N0, MO).** Larger rectal tumors, especially if located in the distal rectum, are more likely to recur locally. There are two schools of thought, each differing in their approach, concerning how to control local recurrences. Advocates of total mesorectal resection suggest that optimization of operative technique will obviate the need for any adjuvant chemoradiation to control local recurrence after resection of stages I, II, and III rectal cancers. The opposing school suggests that stages II and III rectal cancers will benefit from chemoradiation. They argue that such therapy reduces local recurrences and prolongs survival whether given preoperatively or postoperatively. The advantages of preoperative chemoradiation include tumor shrinkage, increased likelihood of resection and of a sphincter-sparing procedure, tumor downstaging by treating locally involved lymph nodes, and decreased risk to the small intestine. Disadvantages include possible overtreatment of early-stage tumors, impaired wound healing, and pelvic fibrosis increasing the risk of operative complications. Postoperative radiation allows accurate pathologic staging of the resected tumor and lymph nodes and avoids the wound healing problems associated with preoperative radiation. However, bulky tumors, tumors involving adjacent organs, and very low rectal tumors may be much more difficult to resect without preoperative radiation and may require a more extensive operation.

**Stage III: Lymph Node Metastasis (Tany, N1, MO).** Many surgeons now recommend chemotherapy and radiation either pre- or postoperatively for node-positive rectal cancers. The advantages and disadvantages are similar to those listed for stage II disease, except that the likelihood of overtreating an early-stage lesion is considerably less.

Over the past two decades, a wide variety of studies have addressed the issue of adjuvant and neoadjuvant therapy for locally advanced rectal cancer. Many of these studies demonstrated both improved local control and prolonged survival and resulted in the 1990 National Institutes of Health (NIH) consensus conference recommendation for postoperative chemoradiation therapy in these patients. There is little controversy regarding chemoradiation therapy for stage III (node-positive) disease. However, advances in surgical technique, such as TME, for locally advanced node-negative cancers (T3-4, N0; stage II) have improved local control with surgery alone, prompting some to abandon adjuvant chemoradiation in these patients, especially for those with cancers in the proximal rectum. Although the data from these studies are intriguing, other reports have shown that chemoradiation improves local control and survival even in patients who undergo TME. Thus, most colorectal surgeons in the United States continue to recommend adjuvant or neoadjuvant therapy for patients with locally advanced disease. Many European surgeons now rely heavily on MRI staging to determine the need for neoadjuvant chemoradiation. They use neoadjuvant chemoradiation if the radial margin is threatened or involved by the cancer or if anal sphincter or other local organ invasion is present. In the United States, chemoradiation therapy is still recommended for all patients with stage III disease and the majority of patients with stage II disease. In select patients with T3 tumors, favorable histology, and negative radical margins, chemoradiation may not be necessary, but larger prospective studies are required before this approach can be recommended.

Appropriate timing of chemoradiation for locally advanced rectal cancer has been debated. Historically, preoperative chemoradiation has been advocated based on tumor shrinkage/downstaging, improved resectability, and the possibility of performing a sphincter-sparing operation in some patients. In addition, the absence of small bowel adhesions in the pelvis may decrease toxicity. However, preoperative radiation therapy may increase operative complications and impair wound healing. Although preoperative endorectal ultrasound and MRI have improved our ability to stage rectal cancer, clinical “overstaging” can be problematic, and neoadjuvant therapy may therefore overtreat patients with pT1-2, N0 tumors. Advocates of postoperative radiation therapy cite more accurate pathologic staging and fewer operative/postoperative complications. However, large, bulky tumors may be unresectable or require a more extensive operation (APR, pelvic exenteration) without preoperative therapy. In addition, postoperative pelvic radiation may compromise function of the neorectum.

Comparisons of perioperative toxicity and oncologic outcome have been addressed by the German CAO/ARO/AIO-94 trial. In this study, pre- and postoperative chemoradiation were associated with equivalent acute toxicity and equivalent postoperative complication rates. Postoperative chemoradiation, however, doubled the risk of postoperative stricture formation. In addition, preoperative chemoradiation halved the risk of local recurrence (6% vs. 12%). Based on these data, most surgeons consider preoperative chemoradiation to be the most appropriate therapy for locally advanced rectal cancer. In the United States, this generally consists of 5-FU based chemotherapy and 5 to 6 weeks of external beam radiation (“long course”) followed by surgery 6 to 8 weeks later. It is important to note, however, that many European centers utilize a “short course” preoperative radiation regimen consisting of 5 days of radiation followed by surgery within 1 to 2 weeks. At present, these modalities have not been compared in any randomized, prospective trial.

With advances in chemoradiation, an increasing number of patients with locally advanced rectal cancer will have complete shrinkage of their tumor (a clinical complete response; cCR). In light of the potential morbidity of proctectomy, it has been suggested that select patients can be managed nonoperatively (“watch and wait”). However, data from current studies are contradictory and concern remains about the ability to predict which patients with clinical complete response actually have a pathologic complete response. Patients selected for nonoperative management must be examined by a surgeon at a frequent intervals. Additional adjuvant chemotherapy administered after the decision for a nonoperative approach is another important consideration. At present, this approach is not recommended outside of a specialty center and/or clinical trial.

**Stage IV: Distant Metastasis (Tany, Nany, M1)** Like stage IV colon carcinoma, survival is limited in patients with distant metastasis from rectal carcinoma. Isolated hepatic
and/or pulmonary metastases are rare, but when present may be resected for cure in selected patients. Some patients will require palliative procedures. Radical resection may be required to control pain, bleeding, or tenesmus, but highly morbid procedures such as pelvic exenteration and sacrectomy should generally be avoided in this setting. Local therapy using cautery, endocavitary radiation, or laser ablation may be adequate to control bleeding or prevent obstruction. Intraluminal stents may be useful in the uppermost rectum but often cause pain and tenesmus lower in the rectum. Occasionally, a proximal diverting colostomy will be required to alleviate obstruction. A mucus fistula should be created if possible to vent the distal colon. It is critical that the morbidity of any procedure be realistically weighed against potential benefit in these patients with limited life expectancy. The assistance of a palliative care team can be invaluable in this setting.

Follow-Up and Surveillance

Patients who have been treated for one colorectal cancer are at risk for the development of recurrent disease (either locally or systemically) or metachronous disease (a second primary tumor). In theory, metachronous cancers should be preventable by using surveillance colonoscopy to detect and remove polyps before they progress to invasive cancer. For most patients, a colonoscopy should be performed within 12 months after the diagnosis of the original cancer (or sooner if the colon was not examined in its entirety prior to the original resection). If that study is normal, colonoscopy should be repeated every 3 to 5 years thereafter.

The optimal method of following patients for recurrent cancer remains controversial. The goal of close follow-up observation is to detect resectable recurrence and to improve survival. Re-resection of local recurrence and resection of distant metastasis to liver, lung, or other sites are often technically challenging and highly morbid, with only a limited chance of achieving long-term survival. Thus, only selected patients who would tolerate such an approach should be followed intensively. Because most recurrences occur within 2 years of the original diagnosis, surveillance focuses on this time period. Patients who have undergone local resection of rectal tumors also should be followed with frequent endoscopic examinations (every 3–6 months for 3 years, then every 6 months for 2 years). CEA is often followed every 3 to 6 months for 2 years. CT scans are often performed annually for 5 years, but there are few data to support this practice. More intensive surveillance is appropriate in high-risk patients such as those with possible Lynch syndrome or T3, N+ cancers. Although intensive surveillance improves detection of resectable recurrences, it is important to note that a survival benefit has never been proven. Therefore, the risks and benefits of intensive surveillance must be weighed and treatment individualized.

Treatment of Recurrent Colorectal Carcinoma

Between 20% and 40% of patients who have undergone curative intent surgery for colorectal carcinoma will eventually develop recurrent disease. Most recurrences occur within the first 2 years after the initial diagnosis, but preoperative chemoradiation therapy may delay recurrence. While most of these patients will present with distant metastases, a small proportion will have isolated local recurrence and may be considered for salvage surgery. Recurrence after colon cancer resection usually occurs at the local site within the abdomen or in the liver or lungs. Resection of other involved organs may be necessary. Recurrence of rectal cancer can be considerably more difficult to manage because of the proximity of other pelvic structures. If the patient has not received chemotherapy and radiation, then adjuvant therapy should be administered prior to salvage surgery. Radical resection may require extensive resection of pelvic organs (pelvic exenteration with or without sacrectomy). Ideally, the aim of a salvage operation should be to resect all of the tumor with negative margins. However, if the ability to achieve a negative margin is in question, the addition of intraoperative radiation therapy (usually brachytherapy) can help improve local control. Pelvic MRI is useful for identifying tumor extension that would prevent successful resection (extension of tumor into the pelvic sidewall, involvement of the iliac vessels or bilateral sacral nerves, sacral invasion above the S2–S3 junction). Patients should also undergo a thorough preoperative evaluation to identify distant metastases (CT of chest, abdomen, and pelvis, and PET scan) before undergoing such an extensive procedure. Nevertheless, radical salvage surgery can prolong survival in selected patients.

Minimally Invasive Techniques for Resection

Laparoscopic colectomy for cancer has been controversial. Early reports of high port site recurrence dampened enthusiasm for this technique. The ability to perform an adequate oncologic resection for cancer has also been questioned. Several trials have laid to rest many of these fears. The Clinical Outcomes of Surgical Therapy (COST) Study Group, the Colon Cancer Laparoscopic or Open Resection (COLOR) trial, and the United Kingdom Medical Research Council Conventional versus Laparoscopic-Assisted Surgery in Colorectal Cancer (CLASSICC) trial all have shown oncologic equivalence between open and laparoscopic techniques. In these multi-institutional studies, the rates of cancer recurrence, survival, and quality of life were similar, suggesting that, in the hands of an appropriately trained surgeon, laparoscopic colectomy is appropriate for colon cancer. The recent introduction of robotic surgery offers an additional minimally invasive approach. Early studies suggest that robotic surgery may be the oncologic equivalent to laparoscopic surgery for colon cancer.

Laparoscopic surgery for rectal cancer has been more controversial. Multiple studies of laparoscopic total mesorectal excision for rectal cancer have demonstrated decreased blood loss, earlier return of bowel function and shorter length of stay compared with open TME. While a laparoscopic approach to mobilization of the left colon and splenic flexure can be used in rectal resection procedures, laparoscopic TME refers to completion of the pelvic dissection laparoscopically and not through the abdominal extraction site. Technical challenges in the retraction of the rectum and surrounding soft tissues and transection of the distal rectum mandate careful evaluation oncologic outcomes. Two recent randomized controlled trials from the United States (ACOSOG Z6051) and Australia and New Zealand (ALaCaRT: Australasian Laparoscopic Cancer of the Rectum Trial) have shown that laparoscopic surgery is not superior to open surgery in this regard. When the totality of evidence from nine randomized trials is evaluated, CRM was positive in 7.9% of laparoscopic and 6.1% of open rectal resections, a difference that was not statistically significant. In trials that reported the completeness of mesorectal excision (n = 5), inadequate TME was significantly more likely in laparoscopic (13.2%) compared with open (10.4%) resections. Long-term recurrence and
survival data from ACOSOG and ALaCaRT studies will be needed to determine what role laparoscopy should play in rectal cancer surgery. Two earlier trials, COLOR (ColoNcancer Laparoscopic or Open Resection) II \(^{134}\) and COREAN (Comparison of Open versus laparoscopic surgery for mid and low Rectal cancer After Neoadjuvant chemoradiotherapy) \(^{135}\) have shown equivalent 3-year recurrence and survival. It is possible that specific subsets of patients are more appropriate for open surgery. For example, ALaCaRT authors thought laparoscopic surgery may be less successful in patients who had neoadjuvant therapy, larger T3 tumors, or higher BMIs, but their study was underpowered to make definitive conclusions in these patients. A laparoscopic approach should not be considered for T4 tumors.

The inferior quality of laparoscopic rectal cancer surgery suggested by recent randomized trials is attributed to technical limitations, leading to the suggestion that robotic surgery, with its wristed instruments, fixed platform, and potentially improved visualization could offer an improvement. The results of the ROLARR (Robotic vs Laparoscopic Resection for Rectal Cancer) randomized trial will offer more data on the quality of TME and margins after robotic surgery; however, a trial comparing open to robotic dissection will also be important.\(^{141}\)

**OTHER NEOPLASMS**

**Rare Colorectal Tumors**

**Neuroendocrine Tumors.** Neuroendocrine tumors occur most commonly in the gastrointestinal tract, and up to 25% of these tumors are found in the rectum. Well-differentiated neuroendocrine tumors (also known as carcinoid tumors) are commonly identified on colonoscopy as small (<1 cm) nodules and have a benign clinical course, with overall survival is greater than 80%. However, the risk of malignancy increases with size, and more than 60% of tumors greater than 2 cm in diameter are associated with distant metastases. Rectal neuroendocrine tumors appear to be less likely to secrete vasoactive substances than carcinoids in other locations, and carcinoid syndrome is uncommon in the absence of hepatic metastases. Small carcinoids can be locally resected transanally. Larger tumors, poorly differentiated tumors (such as small cell or large cell neuroendocrine carcinomas), and those with obvious invasion into the muscularis require more radical surgery. Neuroendocrine tumors in the proximal colon are less common and are more likely to be malignant. Size also correlates with risk of malignancy, and tumors less than 2 cm in diameter rarely metastasize. However, the majority of neuroendocrine tumors in the proximal colon present as bulky lesions, and up to two-thirds of patients will have metastatic spread at the time of diagnosis. These tumors should usually be treated with radical resection. Because well-differentiated neuroendocrine tumors are typically slow growing, patients with distant metastases may expect reasonably long survival. Symptoms of carcinoid syndrome can often be alleviated with somatostatin analogues (octreotide) and/or interferon-α. Tumor debulking can offer effective palliation in selected patients.\(^{142-144}\)

**Mixed Adenoneuroendocrine Carcinomas.** Mixed adenoneuroendocrine carcinomas, also known as composite carcinoid carcinomas, adenosarcomatous tumors, amphiocarcinomas or collision tumors, have histologic features of both neuroendocrine tumors and adenocarcinomas. The natural history of these tumors more closely parallels that of adenocarcinomas than neuroendocrine tumors, and regional and systemic metastases are common. Carcinoid carcinoma of the colon and rectum should be treated according to the same oncologic principles as followed for management of adenocarcinoma.

**Lipomas.** Lipomas are benign lesions that occur most commonly in the submucosa of the colon and rectum. The majority of lipomas are asymptomatic and discovered incidentally. Small asymptomatic lesions do not require resection. However, larger lesions may occasionally cause bleeding, obstruction, or intussusception, especially when greater than 2 cm in diameter. Larger lipomas should be resected by colonoscopic techniques or by a colotomy and enucleation or limited colectomy.\(^{144}\)

**Lymphoma.** Gastrointestinal lymphoma may be primary or generalized/secondary. Primary GI lymphomas occur most frequently in the terminal ileum and cecum. Lymphoma involving the colon and rectum is rare, but it accounts for about 10% of all gastrointestinal lymphomas. Presentation, treatment and prognosis differ between patients with lymphoma occurring as a localized entity in the colon and rectum versus those incurring in patients who have generalized lymphoma with colorectal involvement. Symptoms in isolated rectal lymphoma include bleeding, obstruction, and pain, and these tumors may be clinically indistinguishable from adenocarcinomas. The cecum is most often involved, probably as a result of spread from the terminal ileum. Symptoms include bleeding and obstruction. Bowel resection is the treatment of choice for isolated colorectal lymphoma. Adjuvant therapy may be given based on the stage of disease.\(^{144}\)

**Leiomyoma and Leiomyosarcoma.** Leiomyomas are benign tumors of the smooth muscle of the bowel wall and occur most commonly in the upper gastrointestinal tract. Most patients are asymptomatic, and lesions are often diagnosed incidentally when a mass is seen on endoscopy or felt on digital rectal examination. However, large lesions can cause bleeding or obstruction. Because it is difficult to differentiate a benign leiomyoma from a malignant leiomyosarcoma, these lesions should be resected. Recurrence is common after local resection, but most small leiomyomas can be adequately treated with limited resection. Lesions larger than 5 cm should be treated with radical resection because the risk of malignancy is high.

Leiomyosarcoma is rare in the gastrointestinal tract. When this malignancy occurs in the large intestine, the rectum is the most common site. Leiomyosarcoma of the rectum is usually low grade, and, as such, can be difficult to differentiate from leiomyoma. Definitive diagnosis is usually made after resection. Symptoms, when they occur, are usually bleeding or obstruction. A radical resection is indicated for most of these tumors. Despite complete resection, recurrence is not uncommon, and prognosis is generally poor.\(^{144}\)

**Gastrointestinal Stromal Tumor (GIST).** Gastrointestinal Stromal Tumors (GIST) are most common in the proximal GI tract but do occasionally occur in the colorectum (5–10%) and may be mistaken for leiomyomas. GISTs are mesenchymal tumors that arise from the interstitial cells of Cajal. The vast majority (>95%) of GISTs express CD117 (KIT), and as such, are sensitive to tyrosine kinase inhibitors (TKIs), such as imatinib mesylate and sunitinib malate. Risk stratification is based on tumor size and mitotic activity, and 30% to 50% are malignant. Although small GISTs may be asymptomatic and discovered incidentally, larger lesions can cause bleeding, obstruction, or abdominal pain. Treatment of choice is surgical resection (either local excision or radical resection) with
microscopically negative margins, if possible; however, local recurrence is common. For larger marginally resectable tumors, TKIs (imatinib) can be used to shrink the tumor. These agents can also be considered for adjuvant therapy after resection and are useful for treating metastatic disease.\textsuperscript{144}

**Retrorectal/Presacral Tumors**

Tumors occurring in the retrorectal space are rare. This region lies between the upper two-thirds of the rectum and the sacrum above the rectosacral fascia. It is bound by the rectum anteriorly, the presacral fascia posteriorly, and the endopelvic fascia laterally (lateral ligaments). The retrorectal space contains multiple embryologic remnants derived from a variety of tissues (neuroectoderm, notochord, and hindgut). As such, tumors that develop in this space are often heterogeneous.

Congenital lesions are most common, comprising almost two-thirds of retrorectal lesions. The remainder are classified as neurogenic, osseous, inflammatory, or miscellaneous lesions. Malignancy is more common in the pediatric population than in adults, and solid lesions are more likely to be malignant than are cystic lesions. Inflammatory lesions may be solid or cystic (abscess) and usually represent extensions of infection either in the perirectal space or in the abdomen.

Developmental cysts constitute the majority of congenital lesions and may arise from all three germ cell layers. Dermoid and epidermoid cysts are benign lesions that arise from the ectoderm. Enterogenous cysts arise from the primitive gut. Anterior meningocele and myelomeningocele arise from herniation of the dural sac through a defect in the anterior sacrum. A “scimitar sign” (sacrum with a rounded, concave border without any bony destruction) is the pathognomonic radiographic appearance of this condition.

Solid lesions include teratomas, chordomas, neurologic tumors, or osseous lesions. Teratomas are true neoplasms and contain tissue from each germ cell layer. They often contain both cystic and solid components. Teratomas are more common in children than in adults, but when found in adults, 30% are malignant. Chordomas arise from the notochord and are the most common malignant tumor in this region. These are slow-growing, invasive cancers that show characteristic bony destruction. Neurogenic tumors include neurofibromas and sarcomas, neurilemomas, ependymomas, and ganglioneuromas. Osseous lesions include osteomas and bone cysts, as well as neoplasms such as osteogenic sarcoma, Ewing’s tumor, chordomyxosarcoma, and giant cell tumors.

Patients may present with pain (lower back, pelvic, or lower extremity), gastrointestinal symptoms, or urinary tract symptoms. Most lesions are palpable on digital rectal examination. While plain X-rays and CT scans are often used to evaluate these lesions, pelvic MRI is the most sensitive and specific imaging study. Myelogram is occasionally necessary if there is central nervous system involvement. Treatment is almost always surgical resection. The approach depends in part on the nature of the lesion and its location. Lesions high in the pelvis may be approached via a transabdominal route, whereas low lesions may be resected transsacrally. Intermediate lesions may require a combined abdominal and sacral operation. Although survival is excellent after resection of benign lesions, local recurrence is not uncommon. Prognosis after resection of malignant lesions is highly variable and reflects the biology of the underlying tumor.

The role of biopsy in this setting has been controversial. Historically, the recommendation was to avoid biopsy because of the risk of infection or needle tract seeding. This recommendation has recently been challenged, especially for large and/or unusual tumors that would be better treated with multimodality neoadjuvant therapy (GIST, sarcoma, metastatic adenocarcinoma). A recent study confirmed the utility of needle biopsy of solid lesions and refuted concerns about needle tract seeding. As such, most solid lesions should be biopsied regardless of resectability. Biopsy or aspiration of cystic lesions, especially meningoceles, should still be avoided because of the risk of infection.\textsuperscript{144-147}

**Anal Canal and Perianal Tumors**

Cancers of the anal canal are uncommon and account for approximately 2% of all colorectal malignancies. Neoplasms of the anal canal have traditionally been divided into those affecting the anal margin (distal to the dentate line) and those affecting the anal canal (proximal to the dentate line) based on lymphatic drainage patterns. Lymphatics from the anal canal proximal to the dentate line drain cephalad via the superior rectal lymphatics to the inferior mesenteric nodes and laterally along both the middle rectal vessels and inferior rectal vessels through the ischiorectal fossa to the internal iliac nodes. Lymph from the anal canal distal to the dentate line usually drains to the inguinal nodes. It can also drain to the superior rectal lymph nodes or along the inferior rectal lymphatics to the ischiorectal fossa if primary drainage routes are blocked with tumor (Fig. 29-27).

A more clinically useful classification divides anal lesions into those that are perianal (can be completely visualized with gentle eversion of the buttocks) and those that are intra-anal (cannot be completely visualized with gentle eversion of the buttocks). In many cases, therapy depends on whether the tumor is perianal or intra-anal.

**Squamous Intraepithelial Lesions.** Anal canal and perianal dysplasia have long had a potpourri of nomenclature. Anal intraepithelial neoplasia (AIN), Bowen’s disease, and carcinoma in situ all refer to human papillomavirus (HPV)–induced
dysplasia. Because these entities are pathologically identical, there has been an effort to standardize nomenclature. High-grade squamous intraepithelial lesions (HSIL) include high- and intermediate-grade dysplasia, AINII and AINIII, Bowen’s disease, and carcinoma in situ. Low-grade squamous intraepithelial lesions (LSIL) includes low-grade dysplasia and AINI. Recently, the terms High-grade AIN (HGAIN; AINII) and low-grade AIN (LGAIN; AIN I/III) have been suggested. Both high- and low-grade lesions are associated with infection with HPV, especially types 16 and 18.

High-grade lesions are precursors to invasive squamous cell carcinoma (epidermoid carcinoma) and may appear as a plaque or may only be apparent with high-resolution anoscopy and application of acetic acid and/or Lugol’s iodine solution. The incidence of both squamous intraepithelial lesions and epidermoid carcinoma of the anus has increased dramatically among human immunodeficiency virus (HIV)–positive men who have sex with men. This increase is thought to result from increased rates of HPV infection along with HIV-induced immunosuppression. Treatment of high-grade dysplasia is ablation. Because of a high recurrence and/or reinfestation rate, these patients require close surveillance. High-risk patients should be followed with frequent anal Papanicolaou (Pap) smears every 3 to 6 months. An abnormal Pap smear should be followed by an examination under anesthesia and anal mapping using high-resolution anoscopy. High-resolution anoscopy shows areas with abnormal telangiectasias that are consistent with high-grade dysplasia. Many centers now consider this technique for repeated ablation of high-grade lesions to be the optimal method for following these patients. It should be noted, however, that the practice has not been universally adopted, and it is unclear whether close surveillance in lower-risk (nonimmunosuppressed) patients is necessary. Rarely, extensive disease may require resection with flap closure. Medical therapy for HPV has also been proposed. Topical immunomodulators such as imiquimod (Aldara) have been shown to induce regression in some series. Topical 5-fluorouracil has also been used in this setting. Finally, the introduction of a vaccine against HPV may help decrease the incidence of this disease in the future.

Epidermoid Carcinoma. Epidermoid carcinoma of the anus includes squamous cell carcinoma, cloacogenic carcinoma, transitional carcinoma, and basaloid carcinoma. The clinical behavior and natural history of these tumors are similar. Epidermoid carcinoma is a slow-growing tumor and usually presents as an intra-anal or perianal mass. Pain and bleeding may be present. Perianal epidermoid carcinoma may be treated in a similar fashion as squamous cell carcinoma of the skin in other locations because wide local excision can usually be achieved without resecting the anal sphincter. Intra-anal epidermoid carcinoma cannot be excised locally, and first-line therapy relies on chemotheraphy and radiation (the Nigro protocol: 5-fluorouracil, mitomycin C, and 30 Gy of external beam radiation). This regimen cures 70% to 80% of these tumors. It is important to note that epidermoid carcinomas continue to respond after completion of chemoradiation. Lesions that persist greater than 3 to 6 months after therapy may represent persistent disease and should be biopsied. Recurrence usually requires radical resection (APR). Metastasis to inguinal lymph nodes is a poor prognostic sign.

Verrucous Carcinoma (Buschke-Lowenstein Tumor, Giant Condyoma Acuminata). Verrucous carcinoma is a locally aggressive form of condyloma acuminata. Although these lesions do not metastasize, they can cause extensive local tissue destruction and may be grossly indistinguishable from epidermoid carcinoma. Wide local excision is the treatment of choice when possible, but radical resection may sometimes be required. Topical immunomodulators such as imiquimod (Aldara) may shrink some tumors, but they are almost never curative. Very large lesions may respond to external beam radiation, but resection is almost always required.

Basal Cell Carcinoma. Basal cell carcinoma of the anus is rare and resembles basal cell carcinoma elsewhere on the skin (raised, pearly edges with central ulceration). This is a slow-growing tumor that rarely metastasizes. Wide local excision is the treatment of choice, but recurrence occurs in up to 30% of patients. Radical resection and/or radiation therapy may be required for large lesions.

Adenocarcinoma. Adenocarcinoma of the anus is extremely rare and usually represents downward spread of a low rectal adenocarcinoma. Adenocarcinoma may occasionally arise from the anal glands or may develop in a chronic fistula. Radical resection, usually after neoadjuvant chemoradiation, is usually required.

Extramammary perianal Paget’s disease is adenocarcinoma in situ arising from the apocrine glands of the perianal area. The lesion is typically plaque-like and may be indistinguishable from high-grade intraepithelial lesions. Characteristic Paget’s cells are seen histologically. These tumors are often associated with a synchronous gastrointestinal adenocarcinoma, so a complete evaluation of the intestinal tract should be performed. Wide local excision is usually adequate treatment for perianal Paget’s disease.

Melanoma. Anorectal melanoma is rare, comprising less than 1% of all anorectal malignancies and 1% to 2% of melanomas. Diagnosis is often delayed, and symptoms are attributed to hemorrhoidal disease. Despite many advances in the treatment of cutaneous melanoma, prognosis for patients with anorectal disease remains poor. Overall 5-year survival is less than 10%, and many patients present with systemic metastasis and/or deeply invasive tumors at the time of diagnosis. A few patients with anorectal melanoma, however, present with isolated local or locoregional disease that is potentially resectable for cure, and both radical resection (APR) and wide local excision have been advocated. Recurrence is common and usually occurs systemically regardless of the initial surgical procedure. Local resection with free margins does not increase the risk of local or regional recurrence, and APR offers no survival advantage over local excision. Because of the morbidity associated with APR, wide local excision is recommended for initial treatment of localized anal melanoma. In some patients, wide local excision may not be technically feasible, and APR may be required if the tumor involves a significant portion of the anal sphincter or is circumferential. The addition of adjuvant chemotherapy, biochemotherapy, vaccines, or radiotherapy may be of benefit in some patients, but efficacy remains unproven.

OTHER BENIGN COLORECTAL CONDITIONS

Rectal Prolapse and Solitary Rectal Ulcer Syndrome

Rectal Prolapse. Rectal prolapse refers to a circumferential, full-thickness protrusion of the rectum through the anus and has also been called “first-degree” prolapse, “complete” prolapse, or
procidentia. Internal prolapse occurs when the rectal wall intussuscepts but does not protrude and is probably more accurately described as internal intussusception. Mucosal prolapse is a partial-thickness protrusion often associated with hemorrhoidal disease and is usually treated with banding or hemorrhoidectomy.

In adults, this condition is far more common among women, with a female-to-male ratio of 6:1. Prolapse becomes more prevalent with age in women and peaks in the seventeenth decade of life. In men, prevalence is unrelated to age. Symptoms include tenesmus, a sensation of tissue protruding from the anus that may or may not spontaneously reduce, and a sensation of incomplete evacuation. Mucus discharge and leakage may accompany the protrusion. Patients also present with a myriad of functional complaints, from incontinence and diarrhea to constipation and outlet obstruction.

A thorough preoperative evaluation, including colonic transit studies, anorectal manometry, tests of pudendal nerve terminal motor latency, EMG, and cine-defecography, may be useful. The colon should be evaluated by colonoscopy, air-contrast barium enema, or CT colonography to exclude neoplasms or diverticular disease. Cardiopulmonary condition should be thoroughly evaluated because comorbidities may influence the choice of surgical procedure.

The primary therapy for rectal prolapse is surgery, and more than 100 different procedures have been described to treat this condition. Operations can be categorized as either abdominal or perineal. Abdominal operations have taken three major approaches: (a) reduction of the perineal hernia and closure of the cul-de-sac (Moschowitz repair); (b) fixation of the rectum, either with a prosthetic mesh or fascia lata sling (Ripstein and Wells rectopexy; ventral rectopexy) or by suture rectopexy; or (c) resection of redundant sigmoid colon (Fig. 29-28). In some cases, resection is combined with rectal fixation (resection rectopexy). The recently described ventral rectopexy involves dissection of the anterior rectum down to the pelvic floor. Mesh is sutured to the anterior rectum at one end and anchored to the sacral promontory at the other end. Abdominal procedures for rectal prolapse are increasingly performed laparoscopically or robotically. Perineal approaches have focused on tightening the anus with a variety of prosthetic materials, reefing the rectal mucosa (Delorme procedure), or resecting the prolapsed bowel from the perineum (perineal rectosigmoidectomy or Altemeier procedure) (Fig. 29-29).

Because rectal prolapse occurs most commonly in elderly women, the choice of operation depends in part on the patient’s overall medical condition. Abdominal rectopexy (with or without sigmoid resection) offers the most durable repair, with recurrence occurring in less than 10% of patients. Perineal rectosigmoidectomy avoids an abdominal operation and may be preferable in high-risk patients but is associated with a higher recurrence rate. Reefing the rectal mucosa is effective for patients with limited prolapse. Anal encirclement procedures generally have been abandoned.

**Solitary Rectal Ulcer Syndrome.** Solitary rectal ulcer syndrome and colitis cystica profunda are commonly associated with internal intussusception. Patients may complain of pain, bleeding, mucus discharge, or outlet obstruction. In solitary rectal ulcer syndrome, one or more ulcers are present in the distal rectum, usually on the anterior wall. In colitis cystica profunda, nodules or a mass may be found in a similar location. Evaluation should include anorectal manometry, defecography, and either colonoscopy or barium enema to exclude other diagnoses. Biopsy of an ulcer or mass is mandatory to exclude malignancy or infection due to cytomegalovirus (CMV) in an immunosuppressed patient. Nonoperative therapy (high-fiber diet, defecation training to avoid straining, and laxatives or enemas) is effective in the majority of patients. Biofeedback has also been reported to be effective in some patients. Surgery (either abdominal or perineal repair of prolapse as described earlier) is reserved for highly symptomatic patients who have failed all medical interventions.

**Volvulus**

Volvulus occurs when an air-filled segment of the colon twists about its mesentry. The sigmoid colon is involved in up to 90% of cases, but volvulus can involve the cecum (<20%) or transverse colon. A volvulus may reduce spontaneously, but more commonly produces bowel obstruction, which can progress to strangulation, gangrene, and perforation. Chronic constipation may produce a large, redundant colon (chronic megacolon) that predisposes to volvulus, especially if the mesenteric base is narrow.

The symptoms of volvulus are those of acute bowel obstruction. Patients present with abdominal distention, nausea, and vomiting. Symptoms rapidly progress to generalized abdominal pain and tenderness. Fever and leukocytosis are heralds of gangrene and/or perforation. Occasionally, patients will report a long history of intermittent obstructive symptoms and distention, suggesting intermittent chronic volvulus.

**Sigmoid Volvulus.** Sigmoid volvulus can often be differentiated from cecal or transverse colon volvulus by the appearance of plain X-rays of the abdomen. Sigmoid volvulus produces a characteristic bent inner tube or coffee bean appearance, with the convexity of the loop lying in the right upper quadrant (opposite the site of obstruction). Gastrografin enema shows a
narrowing at the site of the volvulus and a pathognomonic bird’s beak (Fig. 29-30).

Unless there are obvious signs of gangrene or peritonitis, the initial management of sigmoid volvulus is resuscitation followed by endoscopic detorsion. Detorsion is usually most easily accomplished by using a rigid proctoscope, but a flexible sigmoidoscope or colonoscope may also be effective. A rectal tube may be inserted to maintain decompression. Although these techniques are successful in reducing sigmoid volvulus in the majority of patients, the risk of recurrence is high (up to 40%). For this reason, an elective sigmoid colectomy should be performed after the patient has been stabilized and undergone an adequate bowel preparation.

Clinical evidence of gangrene or perforation mandates immediate surgical exploration without an attempt at endoscopic decompression. Similarly, the presence of necrotic mucosa, ulceration, or dark blood noted on endoscopy examination suggests strangulation and is an indication for operation. If dead bowel is present at laparotomy, a sigmoid colectomy with end colostomy (Hartmann’s procedure) may be the safest operation to perform.

Cecal Volvulus. Cecal volvulus results from nonfixation of the right colon. In the majority of cases, rotation occurs around the ileocolic blood vessels and vascular impairment occurs early, although 10% to 30% of the cecum folds upon itself (cecal bascule). Plain X-rays of the abdomen show a characteristic kidney-shaped, air-filled structure in the left upper quadrant (opposite the site of obstruction), and a Gastrografin enema confirms obstruction at the level of the volvulus.

Unlike sigmoid volvulus, cecal volvulus can almost never be detorsed endoscopically. Moreover, because vascular compromise occurs early in the course of cecal volvulus, surgical exploration is necessary when the diagnosis is made. Right hemicolectomy with a primary ileocolic anastomosis can usually be performed safely and prevents recurrence. Simple detorsion or detorsion and cecopexy are associated with a high rate of recurrence.

Transverse Colon Volvulus. Transverse colon volvulus is extremely rare. Nonfixation of the colon and chronic constipation with megacolon may predispose to transverse colon

Figure 29-29. Perineal rectosigmoidectomy shown in lithotomy position. A. A circular incision is made 2 cm proximal to the dentate line. B. The anterior peritoneal reflection is opened. C. The mesentery is divided and ligated. D. The peritoneum may be sutured to the bowel wall. E. The bowel is resected. F. A hand-sewn anastomosis is performed.
volvulus. The radiographic appearance of transverse colon volvulus resembles sigmoid volvulus, but Gastrografin enema will reveal a more proximal obstruction. Although colonoscopic detorsion is occasionally successful in this setting, most patients require emergent exploration and resection.

Megacolon

Megacolon describes a chronically dilated, elongated, hypertrophied large bowel. Megacolon may be congenital or acquired and is usually related to chronic mechanical or functional obstruction. In general, the degree of megacolon is related to the duration of obstruction. Evaluation must always include examination of the colon and rectum (either endoscopically or radiographically) to exclude a surgically correctable mechanical obstruction.

Congenital megacolon caused by Hirschsprung’s disease results from the failure of migration of neural crest cells to the distal large intestine. The resulting absence of ganglion cells in the distal colon results in a failure of relaxation and causes a functional obstruction. The proximal, healthy bowel becomes progressively dilated. Surgical resection of the aganglionic segment is curative. Although Hirschsprung’s disease primarily is a disease of infants and children, it occasionally presents later in adulthood, especially if an extremely short segment of the bowel is affected (ultrashort-segment Hirschsprung’s disease).

Acquired megacolon may result from infection or chronic constipation. Infection with the protozoan Trypanosoma cruzi (Chagas’ disease) destroys ganglion cells and produces both megacolon and megaesophagus. Chronic constipation from slow transit or secondary to medications (especially anticholinergic medications) or neurologic disorders (paraplegia, poliomyelitis, amyotrophic lateral sclerosis, multiple sclerosis) may produce progressive colonic dilatation. Diverting ileostomy or subtotal colectomy with an ileorectal anastomosis is occasionally necessary in these patients.

Colonic Pseudo-Obstruction (Ogilvie’s Syndrome)

Colonic pseudo-obstruction (Ogilvie’s syndrome) is a functional disorder in which the colon becomes massively dilated in the absence of mechanical obstruction. Pseudo-obstruction most commonly occurs in hospitalized patients and is associated with the use of narcotics, bed rest, and comorbid disease. Pseudo-obstruction is thought to result from autonomic dysfunction and severe adynamic ileus. The diagnosis is made based on the presence of massive dilatation of the colon (usually predominantly the right and transverse colon) in the absence of a mechanical obstruction. Initial treatment consists of cessation of narcotics, anticholinergics, or other medications that may contribute to ileus. Strict bowel rest and intravenous hydration are crucial. Most patients will respond to these measures. In patients who fail to improve, colonoscopic decompression often is effective. However, this procedure is technically challenging, and great care must be taken to avoid causing perforation. Up to 40% of patients recur. Intravenous neostigmine (an acetylcholinesterase inhibitor), administered as a single 2 mg dose, also is extremely effective in decompressing the dilated colon and is associated with a low rate of recurrence (20%).155 However, neostigmine may produce transient but profound bradycardia and may be inappropriate in patients with cardiopulmonary

Figure 29-30. Sigmoid volvulus: (A) Illustration and (B) Gastrografin enema showing “bird-beak” sign (arrow). (B. Reproduced with permission from James EC, Corry RJ, Perry JCF: Basic Surgical Practice. Philadelphia, PA: Hanley & Belfus; 1987.)
disease. Because the colonic dilatation is typically greatest in the proximal colon, placement of a rectal tube is rarely effective. It is crucial to exclude mechanical obstruction (usually with a Gastrografin enema) prior to medical or endoscopic treatment.

**Ischemic Colitis**

Intestinal ischemia occurs most commonly in the colon. Unlike small bowel ischemia, colonic ischemia rarely is associated with major arterial or venous occlusion. Instead, most colonic ischemia appears to result from low flow and/or small vessel occlusion. Risk factors include vascular disease, diabetes mellitus, vasculitis, hypotension, and tobacco use. In addition, ligation of the inferior mesenteric artery during aortic surgery predisposes to colonic ischemia. Occasionally, thrombosis or embolism may cause ischemia. Although the splenic flexure is the most common site of ischemic colitis, any segment of the colon may be affected. The rectum is relatively spared because of its rich collateral circulation.

Signs and symptoms of ischemic colitis reflect the extent of bowel ischemia. In mild cases, patients may have diarrhea (usually bloody) without abdominal pain. With more severe ischemia, intense abdominal pain (often out of proportion to the clinical examination), tenderness, fever, and leukocytosis are present. Peritonitis and/or systemic toxicity are signs of full-thickness necrosis and perforation.

The diagnosis of ischemic colitis is often based on the clinical history and physical examination. Plain films may reveal *thumb printing*, which results from mucosal edema and submucosal hemorrhage. CT often shows nonspecific colonic wall thickening and pericolic fat stranding. Angiography is usually not helpful because major arterial occlusion is rare. While sigmoidoscopy may reveal characteristic dark, hemorrhagic mucosa, the risk of precipitating perforation is high. For this reason, *sigmoidoscopy is relatively contraindicated in any patient with significant abdominal tenderness. Contrast studies (Gastrografin or barium enema) are similarly contraindicated during the acute phase of ischemic colitis.*

Treatment of ischemic colitis depends on clinical severity. Unlike ischemia of the small bowel, the majority of patients with ischemic colitis can be treated medically. Bowel rest and broad-spectrum antibiotics are the mainstay of therapy, and 80% of patients will recover with this regimen. Hemodynamic parameters should be optimized, especially if hypotension and low flow appear to be the inciting cause. Long-term sequelae include stricture (10–15%) and chronic segmental ischemia (15–20%). Colonoscopy should be performed after recovery to evaluate strictures and to rule out other diagnoses such as inflammatory bowel disease or malignancy. Failure to improve after 2 to 3 days of medical management, progression of symptoms, and deterioration in clinical condition are indications for surgical exploration. In this setting, all necrotic bowel should be resected. Primary anastomosis should be avoided. Occasionally, repeated exploration (a *second-look operation*) may be necessary.

**Infectious Colitis**

**Pseudomembranous Colitis** (*Clostridium difficile* Colitis). *Pseudomembranous colitis* is caused by *C. difficile*, a gram-positive anaerobic bacillus. *Clostridium difficile* colitis is extremely common and is the leading cause of nosocomially acquired diarrhea. The spectrum of disease ranges from watery diarrhea to fulminant, life-threatening colitis. *Clostridium difficile* is carried in the large intestine of many healthy adults. Colitis is thought to result from overgrowth of this organism after depletion of the normal commensal flora of the gut with the use of antibiotics. Although clindamycin was the first antimicrobial agent associated with *C difficile* colitis, almost any antibiotic may cause this disease. Moreover, although the risk of *C difficile* colitis increases with prolonged antibiotic use, even a single dose of an antibiotic may cause the disease. Immunosuppression, medical comorbidities, prolonged hospitalization or nursing home residence, and bowel surgery increase the risk.

The pathogenic changes associated with *C difficile* colitis result from production of two toxins: *toxin A* (an enterotoxin) and *toxin B* (a cytotoxin). Diagnosis of this disease was traditionally made by culturing the organism from the stool. Detection of one or both toxins (either by cytotoxic assays or by immunoassays) has proven to be more rapid, sensitive, and specific.156 The diagnosis may also be made endoscopically by detection of characteristic ulcers, plaques, and pseudomembranes.157

Management should include immediate cessation of the offending antimicrobial agent. Patients with mild disease (diarrhea but no fever or abdominal pain) may be treated as outpatients with a 10-day course of oral metronidazole. Oral vancomycin is a second-line agent used in patients allergic to metronidazole or in patients with recurrent disease. More severe diarrhea associated with dehydration and/or fever and abdominal pain is best treated with bowel rest, intravenous hydration, and oral metronidazole or vancomycin. Proctosigmoiditis may respond to vancomycin enemas. Recurrent colitis occurs in up to 20% of patients and may be treated by a longer course of oral vancomycin (up to 1 month) or rifaximin (a rapamycin derivative). Reintroduction of normal flora by ingestion of *probiotics* or *stool transplantation* has been suggested as a possible treatment for recurrent or refractory disease. Fulminant colitis, characterized by septicemia and/or evidence of perforation, requires emergent laparotomy. A total abdominal colectomy with end ileostomy may be lifesaving. Over the past decade, *C difficile* colitis has increased in prevalence, and new, more virulent strains have appeared, making this disease increasingly challenging to treat.157,158

**Other Infectious Colitides.** A variety of other infections with bacteria, parasites, fungi, or viruses may cause colonic inflammation. Common bacterial infections include enterotoxic *E coli*, *C jejuni*, *Yersinia enterocolitica*, *S typhi*, *Shigella*, and *N gonorrhoeae*. Less commonly, *Mycobacterium tuberculosis*, *M bovis*, *Actinomyces israelii*, or *Treponema pallidum* (syphilis) may cause colitis or proctitis. Parasitic infections such as amebiasis, cryptosporidiosis, and giardiasis are also relatively common. Fungal infections (*Candida* species, histoplasmosis) are extremely rare in otherwise healthy individuals. The most common viral infections that produce colitic symptoms are HIV, herpes simplex viruses, and CMV.

Most symptoms are nonspecific and consist of diarrhea (with or without bleeding), crampy abdominal pain, and malaise. A thorough history may offer clues to the etiology (other medical conditions, especially immunosuppression; recent travel or exposures; and ingestions). Diagnosis is usually made by identification of a pathogen in the stool, either by microscopy or culture. Serum immunoassays may also be useful (amebiasis, HIV, CMV). Occasionally, endoscopy with biopsy may be required. Treatment is tailored to the infection.
Any patient with anal/perianal symptoms requires a careful history and physical, including a digital rectal examination. Other studies such as defecography, manometry, CT scan, MRI, contrast enema, endoscopy, endoanal ultrasound, or exam under anesthesia may be required to arrive at an accurate diagnosis.

Hemorrhoids

Hemorrhoids are cushions of submucosal tissue containing venules, arterioles, and smooth muscle fibers that are located in the anal canal (see Fig. 29-4). Three hemorrhoidal cushions are found in the left lateral, right anterior, and right posterior positions. Hemorrhoids are thought to function as part of the continence mechanism and aid in complete closure of the anal canal at rest. Because hemorrhoids are a normal part of anorectal anatomy, treatment is only indicated if they become symptomatic. Excessive straining, increased abdominal pressure, and hard stools increase venous engorgement of the hemorrhoidal plexus and cause prolapse of hemorrhoidal tissue. Bleeding, thrombosis, and symptomatic hemorrhoidal prolapse may result.

External hemorrhoids are located distal to the dentate line and are covered with anoderm. Because the anoderm is richly innervated, thrombosis of an external hemorrhoid may cause significant pain. It is for this reason that external hemorrhoids should not be ligated or excised without adequate local anesthetic. A skin tag is redundant fibrotic skin at the anal verge, often persisting as the residua of a thrombosed external hemorrhoid. Skin tags are often confused with symptomatic hemorrhoids. External hemorrhoids and skin tags may cause itching and difficulty with hygiene if they are large. Treatment of external hemorrhoids and skin tags is only indicated for symptomatic relief.

Internal hemorrhoids are located proximal to the dentate line and covered by insensate anorectal mucosa. Internal hemorrhoids may prolapse or bleed, but they rarely become painful unless they develop thrombosis and necrosis (usually related to severe prolapse, incarceration, and/or strangulation). Internal hemorrhoids are graded according to the extent of prolapse. First-degree hemorrhoids bulge into the anal canal and may prolapse beyond the dentate line on straining. Second-degree hemorrhoids prolapse through the anus but reduce spontaneously. Third-degree hemorrhoids prolapse through the anal canal and require manual reduction. Fourth-degree hemorrhoids prolapse but cannot be reduced and are at risk for strangulation.

Combined internal and external hemorrhoids straddle the dentate line and have characteristics of both internal and external hemorrhoids. Hemorrhoidectomy is often required for large, symptomatic, combined hemorrhoids. Postpartum hemorrhoids result from straining during labor, which results in edema, thrombosis, and/or strangulation. Hemorrhoidectomy is often the treatment of choice, especially if the patient has had chronic hemorrhoidal symptoms. Portal hypertension was long thought to increase the risk of hemorrhoidal bleeding because of the anastomoses between the portal venous system (middle and upper hemorrhoidal plexuses) and the systemic venous system (inferior rectal plexuses). It is now understood that hemorrhoidal disease is no more common in patients with portal hypertension than in the normal population. Rectal varices, however, may occur and may cause hemorrhage in these patients. In general, rectal varices are best treated by lowering portal venous pressure. Rarely, suture ligation may be necessary if massive bleeding persists. Surgical hemorrhoidectomy should be avoided in these patients because of the risk of massive, difficult-to-control variceal bleeding.

Treatment

Medical Therapy Bleeding from first- and second-degree hemorrhoids often improves with the addition of dietary fiber, stool softeners, increased fluid intake, and avoidance of straining. Associated pruritus often may improve with improved hygiene. Many over-the-counter topical medications are desiccants and are relatively ineffective for treating hemorrhoidal symptoms.

Rubber Band Ligation Persistent bleeding from first-, second-, and selected third-degree hemorrhoids may be treated by rubber band ligation.

Mucosa located 1 to 2 cm proximal to the dentate line is grasped and pulled into a rubber band applier. After firing the ligator, the rubber band strangulates the underlying tissue, causing scarring and preventing further bleeding or prolapse (Fig. 29-31). In general, only one or two quadrants are banded per visit. Severe pain will occur if the rubber band is placed at or distal to the dentate line where sensory nerves are located. Other complications of rubber band ligation include urinary retention, infection, and bleeding. Urinary retention occurs in approximately 1% of patients and is more likely if the ligation has inadvertently included a portion of the internal sphincter. Necrotizing infection is an uncommon, but life-threatening complication. Severe pain, fever, and urinary retention are early signs of infection and should prompt immediate evaluation of the patient usually with an exam under anesthesia. Treatment includes debridement of necrotic tissue, drainage of associated abscesses, and broad-spectrum antibiotics. Bleeding may occur approximately 7 to 10 days after rubber band ligation, at the time when the ligated pedicle necroses and sloughs. Bleeding is usually self-limited, but persistent hemorrhage may require exam under anesthesia and suture ligation of the pedicle.

Infrared Photocoagulation Infrared photocoagulation is an effective office treatment for small first- and second-degree hemorrhoids. The instrument is applied to the apex of each hemorrhoid to coagulate the underlying plexus. All three quadrants may be treated during the same visit. Larger hemorrhoids and hemorrhoids with a significant amount of prolapse are not effectively treated with this technique.

Sclerotherapy The injection of bleeding internal hemorrhoids with sclerosing agents is another effective office technique for treatment of first-, second-, and some third-degree hemorrhoids. One to 3 mL of a sclerosing solution (phenol in olive oil, sodium morrhuate, or quinine urea) is injected into the submucosa of each hemorrhoid. Few complications are associated with sclerotherapy, but infection and fibrosis have been reported.

Excision of Thrombosed External Hemorrhoids Acutely thrombosed external hemorrhoids generally cause intense pain and a palpable perianal mass during the first 24 to 72 hours after thrombosis. The thrombosis can be effectively treated with an elliptical excision performed in the office under local anesthesia. Because the clot is usually loculated, simple incision and drainage is rarely effective. After 72 hours, the clot begins to resorb, and the pain resolves spontaneously. Excision is unnecessary, but sitz baths and analgesics are often helpful.

Operative Hemorrhoidectomy A number of surgical procedures have been described for elective resection of symptomatic hemorrhoids. All are based on decreasing blood flow to the hemorrhoidal plexuses and excising redundant anoderm and mucosa.
Closed Submucosal Hemorrhoidectomy  The Parks or Ferguson hemorrhoidectomy involves resection of hemorrhoidal tissue and closure of the wounds with absorbable suture. The procedure may be performed in the prone or lithotomy position under local, regional, or general anesthesia. The anal canal is examined and an anal speculum inserted. The hemorrhoid cushions and associated redundant mucosa are identified and excised using an elliptical incision starting just distal to the anal verge and extending proximally to the anorectal ring. It is crucial to identify the fibers of the internal sphincter and carefully brush these away from the dissection in order to avoid injury to the sphincter. The apex of the hemorrhoidal plexus is then ligated and the hemorrhoid excised. The wound is then closed with a running absorbable suture. All three hemorrhoidal cushions may be removed using this technique; however, care should be taken to avoid resecting a large area of perianal skin in order to avoid postoperative anal stenosis (Fig. 29-32).

Open Hemorrhoidectomy  This technique, often called the Milligan and Morgan hemorrhoidectomy, follows the same principles of excision described earlier, but the wounds are left open and allowed to heal by secondary intention.

Whitehead’s Hemorrhoidectomy  Whitehead’s hemorrhoidectomy involves circumferential excision of the hemorrhoidal cushions just proximal to the dentate line. After excision, the rectal mucosa is then advanced and sutured to the dentate line. While some surgeons still use Whitehead’s hemorrhoidectomy, most have abandoned this approach because of the risk of ectropion (Whitehead’s deformity).

Procedure for Prolapse and Hemorrhoids/Stapled Hemorrhoidectomy  Procedure for prolapse and hemorrhoids (PPH) is also referred to as a stapled hemorrhoidopexy. Best suited for patients with second- and third-degree hemorrhoids, this outpatient procedure uses a stapling device similar in appearance and mechanism of action to an end-to-end anastomotic (EEA) stapling device used for rectal surgery. Just as with an EEA stapler, proximal and distal tissue donuts, in this case consisting of mucosa and submucosa, are generated by the PPH stapler though the primary means by which this procedure provides relief for internal hemorrhoids is by pexying the redundant hemorrhoidal tissue, ligating the venules feeding the hemorrhoidal plexus and fixing redundant mucosa proximal to the dentate line. Several studies suggest that this procedure is safe and effective, that it is associated with less postoperative pain and disability, and that it has an equivalent risk of postoperative complications when compared to excisional hemorrhoidectomy. Complications associated with this procedure include chronic anal pain, bacteremia, rectovaginal fistula, formation of an obstructing rectal stricture and even rectal perforation. In at least one systematic review comparing outcomes between PPH and excisional hemorrhoidectomy, overall incidence of complications was similar,
though the incidence of recurrent hemorrhoids was lower following excisional hemorrhoidectomy.\textsuperscript{159,160}

**Doppler-Guided Hemorrhoidal Artery Ligation** Another recent approach to treating symptomatic hemorrhoids is Doppler-guided hemorrhoidal artery ligation (also called trans-anal hemorrhoidal dearterialization). In this procedure, a Doppler probe is used to identify the artery or arteries feeding the hemorrhoidal plexus. These vessels are then ligated. Early reports have shown promise, but long-term durability remains to be determined.\textsuperscript{161}

**Complications of Hemorrhoidectomy.** Postoperative pain following excisional hemorrhoidectomy requires analgesia usually with oral narcotics. Nonsteroidal anti-inflammatory drugs, muscle relaxants, topical analgesics, and comfort measures, including sitz baths, are often useful as well. Urinary retention is a common complication following hemorrhoidectomy and has been reported to be as high as 10\% to 50\% in some series. The risk of urinary retention can be minimized by limiting intraoperative and perioperative intravenous fluids and by providing adequate analgesia. Pain can also lead to \textit{fecal impaction}. Risk of impaction may be decreased by preoperative enemas or a limited mechanical bowel preparation, liberal use of laxatives postoperatively, and adequate pain control. While a small amount of \textit{bleeding}, especially with bowel movements, is to be expected, massive hemorrhage can occur after hemorrhoidectomy. Bleeding may occur in the immediate postoperative period (often in the recovery room) as a result of inadequate ligation of the vascular pedicle. This type of hemorrhage mandates an urgent return to the operating room where suture ligation of the bleeding vessel will often solve the problem. Bleeding may also occur 7 to 10 days after hemorrhoidectomy when the necrotic mucosa overlying the vascular pedicle sloughs. While some of these patients may be safely observed, others will require an exam under anesthesia to ligate the bleeding vessel or to oversew the wounds if no specific site of bleeding is identified. \textit{Infection} is uncommon after hemorrhoidectomy; however, necrotizing soft tissue infection can occur with devastating consequences. Severe pain, fever, and urinary retention may be early signs of infection. If infection is suspected, emergent examination under anesthesia, drainage of abscess, and/or debridement of all necrotic tissue are required.

Long-term sequelae of hemorrhoidectomy include \textit{incontinence}, \textit{anal stenosis}, and \textit{ectropion} (Whitehead’s deformity). Many patients experience transient incontinence to flatus, but these symptoms are usually short-lived, and few patients have

\begin{figure} [h]
\centering
\includegraphics[width=\textwidth]{image}
\caption{Technique of closed submucosal hemorrhoidectomy. \textbf{A.} The patient is in prone jackknife position. \textbf{B.} A Fansler anoscope is used for exposure. \textbf{C.} A narrow ellipse of anoderm is excised. \textbf{D.} A submucosal dissection of the hemorrhoidal plexus from the underlying anal sphincter is performed. \textbf{E.} Redundant mucosa is anchored to the proximal anal canal, and the wound is closed with a running absorbable suture. \textbf{F.} Additional quadrants are excised to complete the procedure.}
\end{figure}
permanent fecal incontinence. Anal stenosis may result from scarring after extensive resection of perianal skin. Ectropion may occur after Whitehead’s hemorrhoidectomy.

**Anal Fissure**

A fissure in ano is a tear in the anoderm distal to the dentate line. The pathophysiology of anal fissure is thought to be related to trauma from either the passage of hard stool or prolonged diarrhea. A tear in the anoderm causes spasm of the internal anal sphincter, which results in pain, increased tearing, and decreased blood supply to the anoderm. This cycle of pain, spasm, and ischemia contributes to development of a poorly healing wound that becomes a *chronic fissure*. The vast majority of anal fissures occur in the posterior midline. Ten percent to 15% occur in the anterior midline. Less than 1% of fissures occur off midline.

**Symptoms and Findings.** Anal fissure is extremely common. Characteristic symptoms include tearing pain with defecation and hematochezia (usually described as blood on the toilet paper). Patients may also complain of a sensation of intense and painful anal spasm lasting for several hours after a bowel movement. On physical examination, the fissure can often be seen in the anoderm by gently separating the buttocks. Patients are often too tender to tolerate digital rectal examination, anoscopy, or proctoscopy. An *acute fissure* is a superficial tear of the distal anoderm and almost always heals with medical management. *Chronic fissures* develop ulceration and heaped-up edges with the white fibers of the internal anal sphincter visible at the base of the ulcer. There often is an associated external skin tag and/or a hypertrophied anal papilla internally. These fissures are more challenging to treat and may require surgery. A lateral location of a chronic anal fissure may be evidence of an underlying disease such as Crohn’s disease, HIV, syphilis, tuberculosis, or leukemia. If the diagnosis is in doubt or there is suspicion of another cause for the perianal pain such as abscess or fistula, an examination under anesthesia may be necessary.

**Treatment.** Therapy focuses on breaking the cycle of pain, spasm, and ischemia thought to be responsible for development of fissure in ano. First-line therapy to minimize anal trauma includes bulk agents, stool softeners, and warm sitz baths. The addition of 2% lidocaine jelly or other analgesic creams can provide additional symptomatic relief. Nitroglycerin ointment has been used locally to improve blood flow but often causes a burning sensation and may have fewer side effects than topical nitrates. Newer agents, such as arginine (a nitric oxide donor) and topical bethanechol (a muscarinic agonist), have also been used to treat fissures. Medical therapy is effective in most acute fissures, but it will heal only approximately 50% of chronic fissures.

Botulinum toxin (Botox) causes temporary muscle paralysis by preventing acetylcholine release from presynaptic nerve terminals. Injection of botulinum toxin is used in some centers as an alternative to surgical sphincterotomy for chronic fissure. Although there are few long-term complications from the use of botulinum toxin, healing appears to be equivalent to other medical therapies.

Surgical therapy has traditionally been recommended for chronic fissures that have failed medical therapy, and lateral internal sphincterotomy is the procedure of choice. The aim of this procedure is to decrease spasm of the internal sphincter by dividing a portion of the muscle. Approximately 30% of the internal sphincter fibers are divided laterally by using either an open (Fig. 29-33) or closed (Fig. 29-34) technique. Healing is achieved in more than 95% of patients using this technique, and most patients experience immediate pain relief. Recurrence occurs in less than 10% of patients, and the risk of incontinence (usually to flatus) ranges from 5% to 15%. Advancement flaps (VY) with or without sphincterotomy have also been reported to successfully treat chronic fissures.

**Anorectal Sepsis and Cryptoglandular Abscess**

**Relevant Anatomy.** The majority of anorectal suppurative disease results from infections of the anal glands (cryptoglandular infection) found in the intersphincteric plane. Their ducts traverse the internal sphincter and empty into the anal crypts at the level of the dentate line. Infection of an anal gland results in the formation of an abscess that enlarges and spreads along one of several planes in the perianal and perirectal spaces. The *perianal space* surrounds the anus and laterally becomes continuous with the fat of the buttocks. The *intersphincteric space* separates the internal and external anal sphincters. It is continuous with the perianal space distally and extends cephalad into the rectal wall. The *ischiorectal space* is located lateral and posterior to the anus and is bounded medially by the external sphincter, laterally by the ischium, superiorly by the levator ani, and inferiorly by the transverse septum. The ischiorectal space contains the inferior rectal vessels and lymphatics. The two ischiorectal spaces connect posteriorly above the anococcygeal ligament but below the levator ani muscle, forming the *deep postanal space*. The *suprallevator spaces* lie above the levator ani on either side of the rectum and communicate posteriorly. The anatomy of these spaces influences the location and spread of cryptoglandular infection (Fig. 29-35).

As an abscess enlarges, it spreads in one of several directions. A *perianal abscess* is the most common manifestation and appears as a painful swelling at the anal verge. Spread through the external sphincter below the level of the puborectalis produces an *ischiorectal abscess*. These abscesses may become extremely large and may not be visible in the perianal region.
Digital rectal exam will reveal a painful swelling laterally in the ischiorectal fossa. *Intersphincteric abscesses* occur in the intersphincteric space and are notoriously difficult to diagnose, often requiring an examination under anesthesia. *Pelvic* and *supralevator abscesses* are uncommon and may result from extension of an intersphincteric or ischiorectal abscess upward or extension of an intraperitoneal abscess downward (Fig. 29-36).

**Diagnosis.** Severe anal pain is the most common presenting complaint. A palpable mass is often detected by inspection of the perianal area or by digital rectal examination. Occasionally, patients will present with fever, urinary retention, or life-threatening sepsis. The diagnosis of a perianal or ischiorectal abscess can usually be made with physical exam alone (either in the office or in the operating room). However, complex or atypical presentations may require imaging studies such as CT or MRI to fully delineate the anatomy of the abscess.

**Treatment.** Anorectal abscesses should be treated by drainage as soon as the diagnosis is established. If the diagnosis is in question, an examination and drainage under anesthesia are often the most expeditious ways both to confirm the diagnosis and to treat the problem. Delayed or inadequate treatment may occasionally cause extensive and life-threatening suppuration with massive tissue necrosis and sepsis. Antibiotics are only indicated if there is extensive overlying cellulitis or if the patient is immunocompromised, has diabetes mellitus, or has valvular heart disease. Antibiotics alone are ineffective at treating perianal or perirectal infection.

### Perianal Abscess

Most perianal abscesses can be drained under local anesthesia in the office, clinic, or emergency department. Larger, more complicated abscesses may require drainage in the operating room. A skin incision is created, and a disk of skin excised to prevent premature closure. No packing is necessary, and sitz baths are started the next day (Fig. 29-37).

### Ischiorectal Abscess

An ischiorectal abscess causes diffuse swelling in the ischiorectal fossa that may involve one or both sides, forming a “horseshoe” abscess. Simple ischiorectal abscesses are drained through an incision in the overlying skin. Horseshoe abscesses require drainage of the deep postanal space and often require counterincisions over one or both ischiorectal spaces (Fig. 29-38).

### Intersphincteric Abscess

Intersphincteric abscesses are notoriously difficult to diagnose because they produce little swelling and few perianal signs of infection. Pain is typically described as being deep and “up inside” the anal area and is usually exacerbated by coughing or sneezing. The pain is so intense that it usually precludes a digital rectal examination. The diagnosis is made based on a high index of suspicion and usually requires an examination.

---

**Figure 29-34.** A through D. Closed lateral internal sphincterotomy for fissure in ano.
Figure 29-35. Anatomy of perianorectal spaces. (A) Anterior view and (B) lateral view. m = muscle.

Figure 29-36. A and B. Pathways of anorectal infection in perianal spaces. m = muscle.
under anesthesia. Once identified, an intersphincteric abscess can be drained through a limited, usually posterior, internal sphincterotomy.

**Supralevator Abscess**

This type of abscess is uncommon and can be difficult to diagnose. Because of its proximity to the peritoneal cavity, supralevator abscesses can mimic intra-abdominal conditions. Digital rectal examination may reveal an indurated, bulging mass above the anorectal ring. It is essential to identify the origin of a supralevator abscess prior to treatment. If the abscess is secondary to an upward extension of an intersphincteric abscess, it should be drained through the rectum. If it is drained through the ischiorectal fossa, a complicated, suprasphincteric fistula may result. If a supralevator abscess arises from the upward extension of an ischiorectal abscess, it should be drained through the ischiorectal fossa. Drainage of this type of abscess through the rectum may result in an extrasphincteric fistula. If the abscess is secondary to intra-abdominal disease, the primary process requires treatment and the abscess is drained via the most direct route (transabdominally, rectally, or through the ischiorectal fossa).

**Perianal Sepsis in the Immunocompromised Patient**

The immunocompromised patient with perianal pain presents a diagnostic dilemma. Because of leukopenia, these patients may develop serious perianal infection without any of the cardinal signs of inflammation. While broad-spectrum antibiotics may cure some of these patients, an exam under anesthesia should not be delayed because of neutropenia. An increase in pain or fever and/or clinical deterioration mandates an exam under anesthesia. Any indurated area should be incised and drained, biopsied to exclude a leukemic infiltrate, and cultured to aid in the selection of antimicrobial agents.\(^{165}\)

**Necrotizing Soft Tissue Infection of the Perineum**

Necrotizing soft tissue infection of the perineum is a rare, but lethal, condition. Most of these infections are polymicrobial and synergistic. The source of sepsis is commonly an undrained or inadequately drained cryptoglandular abscess or a urogenital infection. Occasionally, these infections may be encountered postoperatively (e.g., after hemorrhoidectomy). Immunocompromised patients and diabetic patients are at increased risk.

Physical examination may reveal necrotic skin, bullae, or crepitus. Patients often have signs of systemic toxicity and may be hemodynamically unstable. A high index of suspicion is necessary because perineal signs of severe infection may be minimal and prompt surgical intervention can be lifesaving.

Surgical debridement of all nonviable tissue is required to treat all necrotizing soft tissue infections. Multiple operations may be necessary to ensure that all necrotic tissue has been resected. Broad-spectrum antibiotics are frequently employed, but adequate surgical debridement remains the mainstay of therapy. Colostomy may be required if extensive resection of the sphincter is required or if stool contamination of the perineum makes wound management difficult. Despite early recognition and adequate surgical therapy, the mortality of necrotizing perineal soft tissue infections remains approximately 50%.

---

**Figure 29-37.** A through C. Technique of drainage of perianal abscess.

**Figure 29-38.** Drainage of horseshoe abscess. The deep postanal space is entered, incising the anococcygeal ligament. Counter drainage incisions are made for each limb of the ischiorectal space.
Fistula In Ano

Drainage of an anorectal abscess results in cure for about 50% of patients. The remaining 50% develop a persistent fistula in ano. The fistula usually originates in the infected crypt (internal opening) and tracks to the external opening, usually the site of prior drainage. The course of the fistula can often be predicted by the anatomy of the previous abscesses.

While the majority of fistulas are cryptoglandular in origin, trauma, Crohn’s disease, malignancy, radiation, or unusual infections (tuberculosis, actinomycosis, and chlamydia) may also produce fistulas. A complex, recurrent, or nonhealing fistula should raise the suspicion of one of these diagnoses.

**Diagnosis.** Patients present with persistent drainage from the internal and/or external openings. An indurated tract is often palpable. Although the external opening is often easily identifiable, identification of the internal opening may be more challenging. Goodsell’s rule can be used as a guide in determining the location of the internal opening (Fig. 29-39). In general, fistulas with an external opening anteriorly connect to the internal opening by a short, radial tract. Fistulas with an external opening posteriorly track in a curvilinear fashion to the posterior midline. However, exceptions to this rule often occur if an anterior external opening is greater than 3 cm from the anal margin. Such fistulas usually track to the posterior midline.

Fistulas are categorized based on their relationship to the anal sphincter complex, and treatment options are based on these classifications. An intersphincteric fistula tracks through the distal internal sphincter and intersphincteric space to an external opening near the anal verge (Fig. 29-40A). A transsphincteric fistula often results from an ischiorectal abscess and extends through both the internal and external sphincters (Fig. 29-40B). A suprasphincteric fistula originates in the intersphincteric plane and tracks up and around the entire external sphincter (Fig. 29-40C). An extraspincteric fistula originates in the rectal wall and tracks around both sphincters to exit laterally, usually in the ischiorectal fossa (Fig. 29-40D).

**Treatment.** The goal of treatment of fistula in ano is eradication of sepsis without sacrificing continence. Because fistulous tracks encircle variable amounts of the sphincter complex, surgical treatment is dictated by the location of the internal and external openings and the course of the fistula. The external opening is usually visible as a red elevation of granulation tissue with or without concurrent drainage. The internal opening may be more difficult to identify. Injection of hydrogen peroxide or dilute methylene blue may be helpful. Care must be taken to avoid creating an artificial internal opening (thus often converting a simple fistula into a complex fistula).

Simple intersphincteric fistulas can often be treated by fistulotomy (opening the fistulous tract), curettage, and healing by secondary intention (see Fig. 29-40A). “Horseshoe” fistulas usually have an internal opening in the posterior midline and extend anteriorly and laterally to one or both ischiorectal spaces by way of the deep postanal space. Treatment of a transspincteric fistula depends on its location in the sphincter complex. Fistulas that include less than 30% of the sphincter muscles can often be treated by sphincterotomy without significant risk of major incontinence (see Fig. 29-40B). High transspincteric fistulas, which encircle a greater amount of muscle, are more safely treated by initial placement of a seton. Similarly, suprasphincteric fistulas are usually treated with seton placement (see Fig. 29-40C). Extraposphincteric fistulas are rare, and treatment depends on both the anatomy of the fistula and its etiology. In general, the portion of the fistula outside the sphincter should be opened and drained. A primary tract at the level of the dentate line may also be opened if present. Complex fistulas with multiple tracts may require numerous procedures to control sepsis and facilitate healing. Liberal use of drains and setons is helpful. Failure to heal may ultimately require fecal diversion (see Fig. 29-40D). Complex and/or nonhealing fistulas may result from Crohn’s disease, malignancy, radiation proctitis, or unusual infection. Proctoscopy should be performed in all cases of complex and/or nonhealing fistulas to assess the health of the rectal mucosa. Biopsies of the fistula tract should be taken to rule out malignancy.

A seton is a drain placed through a fistula to maintain drainage and/or induce fibrosis. Cutting setons consist of a suture or a rubber band that is placed through the fistula and intermittently tightened in the office. Tightening the seton results in fibrosis and gradual division of the sphincter, thus eliminating the fistula while maintaining continuity of the sphincter. A noncutting seton is a soft plastic drain (often a vessel loop) placed in the fistula to maintain drainage. The fistula tract may subsequently be laid open with less risk of incontinence because scarring prevents retraction of the sphincter. Alternatively, the seton may be left in place for chronic drainage. Higher fistulas may be treated by an endorectal advancement flap. Fibrin glue and a variety of collagen-based plugs also have been used to treat persistent fistulas with variable results. A more recent technique, ligation of the intersphincteric fistula tract (LIFT), also shows promise. In this procedure, the fistula is identified in the intersphincteric plane (usually by placement of a lacrimal probe), divided, and the two ends ligated. Early reports have shown success with this technique, but long-term outcome is not yet known.

**Rectovaginal Fistula**

A rectovaginal fistula is a connection between the vagina and the rectum or anal canal proximal to the dentate line. Rectovaginal fistulas are classified as low (rectal opening close to the dentate line and vaginal opening in the fourchette), middle (vaginal opening between the fourchette and cervix), or high (vaginal opening near the cervix). Low rectovaginal fistulas are
commonly caused by obstetric injuries or trauma from a foreign body. Mid-rectovaginal fistulas may result from more severe obstetric injury, but they also occur after surgical resection of a mid-rectal neoplasm, radiation injury, or extension of an undrained abscess. High rectovaginal fistulas result from operative or radiation injury. Complicated diverticulitis may cause a colovaginal fistula. Crohn’s disease can cause rectovaginal fistulas at all levels, as well as colovaginal and enterovaginal fistulas.

**Diagnosis.** Patients describe symptoms varying from the sensation of passing flatus from the vagina to the passage of solid stool from the vagina. Most patients experience some degree of fecal incontinence. Contamination may result in vaginitis. Large fistulas may be obvious on anoscopy and/or vaginal speculum examination, but smaller fistulas may be difficult to locate. Occasionally, a barium enema or vaginogram may identify these fistulas. Endorectal ultrasound may also be useful. With the patient in the prone position, installation of methylene blue into the rectum while a tampon is in the vagina may confirm the presence of a small fistula.

**Treatment.** The treatment of rectovaginal fistula depends on the size, location, etiology, and condition of surrounding tissues. Because up to 50% of fistulas caused by obstetric injury heal spontaneously, it is prudent to wait 3 to 6 months before embarking on surgical repair in these patients. If the fistula was caused by a cryptoglandular abscess, drainage of the abscess may allow spontaneous closure.

Low and mid-rectovaginal fistulas are usually best treated with an endorectal advancement flap. The principle of this procedure is based on the advancement of healthy mucosa, submucosa, and circular muscle over the rectal opening (the high-pressure side of the fistula) to promote healing (Fig. 29-41). If a sphincter injury is present, an overlapping sphincteroplasty

Figure 29-40. The four major categories of fistula in ano (left side of drawings) and the usual operative procedure to correct the fistula (right side of drawings). A. Intersphincteric fistula with simple low tract. B. Uncomplicated transsphincteric fistula. C. Uncomplicated suprasphincteric fistula. D. Extrasphincteric fistula secondary to anal fistula.
should be performed concurrently. Fecal diversion is rarely required. High rectovaginal, colovaginal, and enterovaginal fistulas are usually best treated via a transabdominal approach. The diseased tissue, which caused the fistula (upper rectum, sigmoid colon, or small bowel), is resected and the hole in the vagina closed. Healthy tissue, such as omentum or muscle, frequently is interposed between the bowel anastomosis and the vagina to prevent recurrence.

Rectovaginal fistulas caused by Crohn’s disease, radiation injury, or malignancy almost never heal spontaneously. In Crohn’s disease, treatment is based on adequate drainage of perianal sepsis and nutritional support. An endorectal advancement flap may be performed if the rectum is spared from active Crohn’s disease. Fistulas resulting from radiation damage are not amenable to local repair with an advancement flap because of damage to the surrounding rectal and vaginal tissues. Such mid- and high rectovaginal fistulas are occasionally repaired successfully with a transabdominal approach in which healthy tissue (omentum, muscle, or nonradiated bowel) is interposed between the damaged rectum and vagina. Fistulas caused by malignancy should be treated with resection of the tumor. Because differentiating radiation damage from malignancy can be extremely difficult, all fistulas resulting from radiation should be biopsied to rule out the presence of cancer.

Perianal Dermatitis

Pruritus Ani. Pruritus ani (severe perianal itching) is a common problem with a multitude of etiologies. Surgically correctable (anatomic) causes include prolapsing hemorrhoids, ectropion, fissure, fistula, and neoplasms. Perianal infection may also present with pruritus ani. Infections may be caused by fungus (Candida species and Epidermophyton organisms), parasites (Enterobius vermicularis [pinworms], Pediculus pubis [a louse], and Sarcoptes scabiei [scabies]), bacteria
(Corynebacterium minutissimum [erythrasma] and T pallidum [syphilis]), or viruses (HPV [condyloma acuminata]). Antibiotic use may also cause itching, usually by precipitating fungal infection. Noninfectious dermatologic causes include seborrhea, psoriasis, and contact dermatitis. Contact dermatitis can be particularly troublesome because many over-the-counter topical agents used by patients to relieve itching may exacerbate the problem. Occasionally, systemic diseases such as jaundice and diabetes may present with pruritus ani.

Despite the myriad of causes, the majority of pruritus ani is idiopathic and probably related to local hygiene, neurogenic, or psychogenic causes. Treatment focuses on removal of irritants, improving perianal hygiene, dietary adjustments, and avoiding scratching. Biopsy and/or culture may be required to rule out an infectious or dermatologic cause. Hydrocortisone ointment 0.5% to 1.0% can provide symptomatic relief but should not be used for prolonged periods of time because of the risk of dermal atrophy. Skin barriers such as Calmoseptine can also provide relief. Systemic antihistamines or tricyclic antidepressants have also been used with some success.

Nonpruritic Lesions. Several perianal skin conditions may present with perianal skin changes. Leprosy, amebiasis, actinomycosis, and lymphogranuloma venereum produce characteristic perianal lesions. Neoplasms such as squamous intraepithelial lesions, Paget’s disease, and invasive carcinomas may also appear first in the perianal skin. Biopsy can usually distinguish these diagnoses.

Sexually Transmitted Diseases

Bacterial Infections. Proctitis is a common symptom of anorectal bacterial infection. Neisseria gonorrhoeae is the most common bacterial cause of proctitis and causes pain, tenesmus, rectal bleeding, and mucus discharge. Chlamydia trachomatis infection may be asymptomatic or may produce similar symptoms. Treponema pallidum, the microbe causing syphilis, causes a chancre at the site of inoculation, which may be asymptomatic or may present as an atypical fissure (primary syphilis). Condyloma lata are characteristic of secondary syphilis. Chancroid, caused by Haemophilus ducreyi, is a disease manifested by multiple painful, bleeding lesions. Inguinal lymphadenopathy and fluctuant, draining lymph nodes are characteristic. Donovania granulomatis infection produces shiny, red masses on the perineum (granuloma inguinale). Diarrheal illnesses caused by organisms such as Campylobacter or Shigella may also be sexually transmitted. Treatment consists of antimicrobial agents directed against the infecting organism.

Parasitic Infections. Entamoeba histolytica is an increasingly common sexually transmitted disease. Amebas produce ulcerations in the gastrointestinal mucosa and can infect any part of the gut. Symptoms include diarrhea, abdominal pain, and tenesmus. Giardia lamblia is also common and produces diarrhea, abdominal pain, and malaise.

Viral Infections

Herpes Simplex Virus. Herpes proctitis is extremely common. Proctitis is usually caused by type 2 herpes simplex virus and less commonly by type 1 herpes simplex virus. Patients complain of severe, intractable perianal pain and tenesmus. Pain often precedes the development of characteristic vesicles, and these patients may require an examination under anesthesia to exclude another diagnosis such as an interspinchteric abscess. Diagnosis is confirmed by viral culture of tissue or vesicular fluid.

Human Papillomavirus. HPV causes condyloma acuminata (anogenital warts) and is associated with squamous intraepithelial lesions and squamous cell carcinoma (see previous section, “Anal Canal and Perianal Tumors”). Condylomas occur in the perianal area or in the squamous epithelium of the anal canal. Occasionally, the mucosa of the lower rectum may be affected. There are approximately 30 serotypes of HPV. As previously mentioned, HPV types 16 and 18, in particular, appear to predispose to malignancy and often cause flat dysplasia in skin unaffected by warts. In contrast, HPV types 6 and 11 commonly cause warts, but do not appear to cause malignant degeneration.

Treatment of anal condyloma depends on the location and extent of disease. Small warts on the perianal skin and distal anal canal may be treated in the office with topical application of bichloracetic acid or podophyllin. Although 60% to 80% of patients will respond to these agents, recurrence and reinfection are common. Imiquimod (Aldara) is an immunomodulator that was recently introduced for topical treatment of several viral infections, including anogenital condyloma. Initial reports suggest that this agent is highly effective in treating condyloma located on the perianal skin and distal anal canal. Larger and/or more numerous warts require excision and/or fulguration in the operating room. Excised warts should be sent for pathologic examination to rule out dysplasia or malignancy. It is important to note that prior use of podophyllin may induce histologic changes that mimic dysplasia. The recent introduction of a vaccine against HPV holds promise for preventing anogenital condylomas.

Human Immunodeficiency Virus. See later section, “The Immunocompromised Patient.”

Pilonidal Disease

Pilonidal disease (cyst, infection) consists of a hair-containing sinus or abscess occurring in the intergluteal cleft. Although the etiology is unknown, it is speculated that the cleft creates a suction that draws hair into the midline pits when a patient sits. These ingrown hairs may then become infected and present acutely as an abscess in the sacrococcygeal region. Once an acute episode has resolved, recurrence is common.

An acute abscess should be incised and drained as soon as the diagnosis is made. Because these abscesses are usually very superficial, this procedure can often be performed in the office, clinic, or emergency department under local anesthetic. Because midline wounds in the region heal poorly, some surgeons recommend using an incision lateral to the intergluteal cleft. A number of procedures have been proposed to treat a chronic pilonidal sinus. The simplest method involves unroofing the tract, curetting the base, and marsupializing the wound. The wound must then be kept clean and free of hair until healing is complete (often requiring weekly office visits for wound care). Alternatively, a small lateral incision can be created and the pit excised. This method is effective for most primary pilonidal sinuses. In general, extensive resection should be avoided. Complex and/or recurrent sinus tracts may require more extensive resection and closure with a Z-plasty, advancement flap, or rotational flap.

Hidradenitis Suppurativa

Hidradenitis suppurativa is an infection of the cutaneous apocrine sweat glands. Infected glands rupture and form subcutaneous sinus tracts. The infection may mimic complex anal fistula disease, but stops at the anal verge because there are no apocrine
glands in the anal canal. Treatment involves incision and drainage of acute abscesses and unroofing of all chronically inflamed fistulas and debridement of granulation tissue. Radical excision and skin grafting are almost never necessary.

TRAGMA

Penetrating Colorectal Injury

Colorectal injury is common following penetrating trauma to the abdomen and has historically been associated with high mortality. In the first half of the 20th century, the mortality rate from colorectal injury was as high as 90%. The introduction of exteriorization of colonic injuries and fecal diversion during World War II dramatically decreased mortality, and this principle has governed the management of large bowel injury for over 50 years. Recently, however, this practice was challenged, and trauma surgeons are increasingly performing primary repairs in selected patients.

Management of colonic injury depends on the mechanism of injury, the delay between the injury and surgery, the overall condition and stability of the patient, the degree of peritoneal contamination, and the condition of the injured colon. A primary repair may be considered in hemodynamically stable patients with few additional injuries and minimal contamination if the colon appears otherwise healthy. Contraindications to primary repair include shock, injury to more than two other organs, mesenteric vascular damage, and extensive fecal contamination. A delay of greater than 6 hours between the injury and the operation also is associated with increased morbidity and mortality and is a relative contraindication to primary repair. Injuries caused by high-velocity gunshot wounds or blast injuries are often associated with multiple intra-abdominal injuries and tissue loss and therefore are usually treated by fecal diversion after debridement of all nonviable tissue. Patient factors, such as medical comorbidities, advanced age, and the presence of tumor or radiation injury, must also be considered (Table 29-4).

Like injuries to the intraperitoneal colon, penetrating trauma to the rectum traditionally has been associated with high morbidity and mortality. Primary repair of the rectum is more difficult than primary repair of the colon, however, and most rectal injuries are associated with significant contamination. For that reason, the majority of penetrating rectal injuries should be treated with proximal fecal diversion. Distal washout (copious irrigation of the rectum) and presacral drains are not routinely recommended. Small, clean rectal injuries may be closed primarily without fecal diversion in an otherwise stable patient. Intractable rectal bleeding may require angiographic embolization. Very rarely, hemorrhage or extensive tissue loss (especially if the anal sphincter is severely damaged) may require an emergent APR. However, this operation should be avoided, if at all possible, because of the morbidity associated with an extensive pelvic dissection in a severely injured patient.

Blunt Colorectal Injury

Blunt injury to the colon and rectum is considerably less common than penetrating injury. Nevertheless, blunt trauma can cause colon perforation, and shear injury to the mesentery can devascularize the intestine. Management of these injuries should follow the same principles outlined for management of penetrating injuries. Small perforations with little contamination in a stable patient may be closed primarily; more extensive injury requires fecal diversion. A serosal hematoma alone does not mandate resection, but the bowel should be carefully inspected to ensure that there is not an associated perforation or significant bowel ischemia.

Blunt injury to the rectum may result from significant trauma, such as a pelvic crush injury, or may result from local trauma caused by an enema or foreign body. Crush injuries, especially with an associated pelvic fracture, are often associated with significant rectal damage and contamination. These patients require debridement of all nonviable tissue, proximal fecal diversion, and a distal rectal washout, with or without drain placement. Blunt trauma from an enema or foreign body may produce a mucosal hematoma, which requires no surgical treatment if the mucosa is intact. Small mucosal tears may be closed primarily if the bowel is relatively clean and there is little contamination.

Iatrogenic Injury

Intraoperative Injury. The colon and rectum are at risk for inadvertent injury during other procedures, especially during pelvic operations. The key to managing these injuries is early recognition. The vast majority of iatrogenic colorectal injuries may be closed primarily if there is little contamination and if the patient is otherwise stable. Delayed recognition of colorectal injuries may result in significant peritonitis and life-threatening sepsis. In these cases, fecal diversion is almost always required, and the patient may need repeated exploration for drainage of abscesses.

Injury From Barium Enema. Colorectal injury from a barium enema is an extremely rare complication associated with a high rate of morbidity and mortality. Perforation with spillage
of barium, especially above the peritoneal reflection, may result in profound peritonitis, sepsis, and a systemic inflammatory response. If the perforation is recognized early, it may be closed primarily and the abdomen irrigated to remove stool and barium. However, if the patient has developed sepsis, fecal diversion (with or without bowel resection) is almost always required. Rarely, a small mucosal injury to the extraperitoneal rectum may be managed with bowel rest, broad-spectrum antibiotics, and close observation.

**Colonoscopic Perforation.** Perforation is the most common major complication after either diagnostic or therapeutic colonoscopy. Fortunately, this complication is rare and occurs in less than 1% of procedures. Perforation may result from trauma from the tip of the instrument, from shear forces related to the formation of a “loop” in the colonoscope, or from barotrauma from insufflation. Biopsy or fulguration can also cause perforation. Polypectomy using electrocautery may produce a full-thickness burn, resulting in *postpolypectomy syndrome* in which a patient develops abdominal pain, fever, and leukocytosis without evidence of diffuse peritonitis.

Management of colonoscopic perforation depends on the *size of the perforation*, the *duration of time* since the injury, the *overall condition of the patient*, and the *underlying diagnosis*. A large perforation recognized during the procedure requires surgical exploration. Because the bowel has almost always been prepared prior to the colonoscopy, there is usually little contamination associated with these injuries, and most can be repaired primarily. If there is significant contamination, if there has been a delay in diagnosis with resulting peritonitis, or if the patient is hemodynamically unstable, proximal diversion with or without resection is the safest approach. It is also important to know the indication for and findings at the time of colonoscopy. If the patient has an underlying neoplasm and is stable, definitive resection is best. Occasionally, a patient will develop abdominal pain and localized signs of perforation after what was thought to be an uneventful colonoscopy. Many of these patients will have a “microperforation,” which will resolve with bowel rest, broad-spectrum antibiotics, and close observation. Evidence of peritonitis or any deterioration in clinical condition mandates exploration. Similarly, free retroperitoneal or intraperitoneal air may be discovered incidentally after colonoscopy. In a completely asymptomatic patient, this finding is thought to result from barotrauma and dissection of air through tissue planes without a free perforation. Many of these patients can be successfully treated with bowel rest and broad-spectrum antibiotics. Surgical exploration is indicated if any clinical deterioration occurs.

**Anal Sphincter Injury and Incontinence**

The most common cause of anal sphincter injury is obstetric trauma during vaginal delivery. The risk of sphincter injury is increased by a laceration that extends into the rectum (fourth-degree tear), infection of an episiotomy or laceration repair, prolonged labor, and possibly by use of a midline episiotomy. Sphincter damage may also result from hemorrhoidectomy, sphincterotomy, abscess drainage, or fistulotomy. Patients with incontinence and a suspected sphincter injury can be evaluated with anal manometry, EMG, pudendal nerve motor latency, and endoanal ultrasound. Mild incontinence, even in the presence of a sphincter defect, may respond to dietary changes and/or biofeedback. More severe incontinence may require surgical repair.

The anal sphincter can also be injured by penetrating or blunt mechanisms (impalement, blast injury, crush injuries of the pelvis). Because damage to the anal sphincter is not lifethreatening, definitive repair of the sphincter is often deferred until other injuries have been repaired and the patient’s clinical condition is stable. Isolated sphincter injuries that do not involve the rectum may be repaired primarily. Rectal injury accompanied by sphincter injury should be treated with fecal diversion and distal rectal washout, with or without drain placement. Significant perineal tissue loss may require extensive debridement and a diverting colostomy.

**Surgical Repairs.** The most common method of repair of the anal sphincter is a *wrap-around sphincteroplasty* (Fig. 29-42). The procedure involves mobilization of the divided sphincter muscle and reapproximation without tension. The internal and external sphincters may be overlapped together or separately. *Postanal intersphincteric levatorplasty* is less commonly used to repair sphincter defects but may be useful for incontinence caused by prolapse and/or loss of the anorectal angle (see “Continence”). The approach is via the intersphincteric plane posteriorly. It may be performed concomitantly with a perineal repair of rectal prolapse. The levator ani muscle is approximated to restore the anorectal angle, and the puborectalis and external sphincter muscles are tightened with sutures. These elective procedures usually do not require a diverting colostomy.

In cases where there has been significant loss of sphincter muscle or in which prior repairs have failed, more complex techniques, such as *gracilis muscle transposition* with or without

---

**Figure 29-42.** Overlapping sphincteroplasty for incontinence from sphincter disruption. A. The external sphincter muscle with scar at site of injury is mobilized. B. The muscle edges are aligned in an overlapping fashion. C. Mattress sutures are used to approximate the muscle. D. The completed operation.
chronic, low-frequency electrostimulation, have been used with some success. In this procedure, the gracilis muscle is mobilized from the thigh, detached from its insertion on the tibial tuberosity, tunneled through the perineum, and wrapped around the anal canal. Another alternative in patients who have failed other repairs is the artificial anal sphincter. This device consists of an inflatable silastic cuff, a pressure-regulating balloon, and a control pump. Patients deflate the cuff manually to open the anal canal; the cuff then reinflates spontaneously to maintain closure of the anal canal. Frequent infections and erosion can lead to device loss. Sacral nerve stimulation via an implanted pulse generator is a technique used for neurogenic incontinence when the sphincter is intact. In some patients, an end stoma provides the best relief for intractable incontinence.

Foreign Body
Foreign body entrapment in the rectum is not uncommon. Depending on the level of entrapment, a foreign body may cause damage to the rectum, rectosigmoid, or descending colon. Generalized abdominal pain suggests intraperitoneal perforation. Evaluation of the patient includes inspection of the perineum and a careful abdominal examination to detect any evidence of perforation. Plain films of the abdomen are mandatory to detect free intra-abdominal air.

Foreign bodies lodged low in the rectum may often be removed under conscious sedation with or without a local anesthetic block. Objects impacted higher in the rectum may require regional or general anesthesia for removal. Only rarely will a laparotomy be required to remove the object, either through manual manipulation of the object to expel from the anus, or via colotomy. After removal of the foreign body, it is crucial to evaluate the rectum and sigmoid colon for injury. Proctoscopy and/or flexible sigmoidoscopy should be performed. A hematoma without evidence of perforation requires no surgical treatment. Perforation of the rectum or sigmoid colon should be managed as described in the preceding sections.

THE IMMUNOCOMPROMISED PATIENT

Human Immunodeficiency Virus
Patients infected with HIV may present with a myriad of gastrointestinal symptoms. Diarrhea, in particular, is extremely common. The severity of gastrointestinal disease depends in part on the degree of immunosuppression; however, both ordinary and opportunistic pathogens may affect patients at any stage of the disease. Opportunistic infections with bacteria (Salmonella, Shigella, Campylobacter, Chlamydia, and Mycobacterium species), fungi (histoplasmosis, coccidiosis, Cryptococcus), protozoa (toxoplasmosis, cryptosporidiosis, isosporiasis), and viruses (CMV, herpes simplex virus) can cause diarrhea, abdominal pain, and weight loss. CMV in particular may cause severe enterocolitis and is the most common infectious cause of emergency laparotomy in acquired immunodeficiency syndrome (AIDS) patients. Clostridium difficile colitis is a major concern in these patients, especially because many patients are maintained on suppressive antibiotic therapy. The incidence of gastrointestinal malignancy is also increased in patients with HIV infection. Kaposi’s sarcoma is the most common malignancy in AIDS patients and can affect any part of the gastrointestinal tract. Patients may be asymptomatic or may develop bleeding or obstruction. Gastrointestinal lymphoma (usually non-Hodgkin’s lymphoma) is also common. The incidence of colorectal carcinoma may also be increased in this population, although definitive data are lacking.

Perianal disease is extremely common in patients with HIV infection. Because HIV is sexually transmitted, it is common to find concomitant infection with other sexually transmitted diseases such as Chlamydia, herpes simplex virus, and HPV (anal condyloma). Anal condyloma in particular are very common, and the incidence of dysplasia (HGAIN) is high in the HIV-infected population. Abscesses and fistulas may be more difficult to diagnose in these patients and may be complex. Many patients require an examination under anesthesia with biopsy and cultures to determine the etiology of many of these perianal problems. The introduction of highly active antiretroviral therapy (HAART) has changed the natural history of HIV infection, but it remains to be seen how these medications will affect the incidence and outcome of colorectal disease in this patient population.

IMMUNOSUPPRESSION FOR TRANSPLANTATION
The gastrointestinal tract is a common site for posttransplantation complications that are responsible for significant morbidity and mortality. In these patients, infection and medication are the most common causes of diarrhea. Immunosuppressant medications, in particular, may cause diarrhea. CMV infection is common and may be severe. Clostridium difficile colitis also occurs commonly. Diverticulitis appears to be more common in some populations of transplant patients and may be more likely to present with abscess or free perforation. Elective resection after recovery from one episode of confirmed diverticulitis may be indicated in the transplant population. Graft-versus-host disease is unique to transplant patients and often requires endoscopy and biopsy to diagnose gastrointestinal involvement. Patients are subject to the same opportunistic infections outlined earlier; however, sexually transmitted infections and Kaposi’s sarcoma are somewhat less prevalent. Perianal disease is somewhat less common in the transplant population than in patients infected with HIV; however, similar infections may occur, and immunosuppression often makes diagnosis and treatment challenging.

With increasing long-term survival among transplant recipients, the development of posttransplant malignancy has become a major concern. Posttransplant lymphoproliferative disease is increasingly common and may occur anywhere in the gastrointestinal tract. The risk of colorectal carcinoma is increased in patients with predisposing conditions such as ulcerative colitis. However, immunosuppression alone does not appear to increase the incidence of colorectal cancer, and current screening recommendations are similar to those for the average risk population. In contrast, the incidence of anal squamous cell carcinoma is dramatically increased in transplant patients, and patients with known HPV infection should undergo more vigorous screening.

THE NEUTROPENIC PATIENT

Neutropenic enterocolitis (typhilitis) is a life-threatening problem with a mortality rate of greater than 50%. This syndrome is characterized by abdominal pain and distention, fever, diarrhea (often bloody), nausea, and vomiting in a patient with fewer than 1000 neutrophils/μL blood from any cause (bone marrow transplantation, solid-organ transplantation, or chemotherapy). Its etiology is poorly understood. Histologic features can be
seen on biopsy or surgical resection and include a paucity of inflammatory and leukemic infiltrates but with mucosal and submucosal edema, villous sloughing, stromal hemorrhage, and patchy-to-complete epithelial necrosis. CT scan of the abdomen often shows a dilated cecum with pericolic stranding. However, a normal-appearing CT scan does not exclude the diagnosis. Some patients will respond to bowel rest, broad-spectrum antibiotics, parenteral nutrition, and granulocyte infusion or colony-stimulating factors. Evidence of perforation, generalized peritonitis, and deterioration in clinical condition are indications for operation.

Neutropenic patients often develop perianal pain, and diagnosis may be difficult because of a lack of inflammatory response to infection. While broad-spectrum antibiotics may cure some of these patients, an examination under anesthesia should not be delayed because of neutropenia. An increase in pain or fever and/or clinical deterioration mandates an exam under anesthesia. Any indurated area should be incised and drained, biopsied to exclude a leukemic infiltrate, and cultured to aid in the selection of antimicrobial agents.

REFERENCES

Entries highlighted in bright blue are key references.

10. Lynch HT, Lynch JF, Lynch PM, et al. Hereditary colorectal cancer syndromes: molecular genetics, genetic counseling, diagnosis and management. Fam Cancer. 2008;7:27-39. This reference is included in an issue of Familial Cancer that is devoted to discussion of hereditary forms of colorectal cancer. The paper summarizes the evaluation of patients suspected to have familial colorectal cancer with a special emphasis on HNPCC (Lynch Syndrome) and FAP. Recommendations are provided for genetic testing, surveillance strategies, and treatment.
COLON, RECTUM, AND ANUS

CHAPTER 29


36. Scarborough JE, Mantyh CR, Sun Z, Migaly J. Combined mechanical and oral antibiotic bowel preparation reduces incisional surgical site infection and anastomotic leak rates after elective colorectal resection. Ann Surg. 2015;262(2):331-337. This study assessed the effect of combined mechanical and oral antibiotic bowel preparation (prep) and 30-day outcomes after colorectal resection using data from the 2012 Colectomy-Targeted American College of Surgeons National Surgical Quality Improvement (ACS NSQIP) data (4999 patients). The combination of oral and mechanical bowel preparation decreased the incidence of surgical site infection, anastomotic leak, and hospital readmission when compared to mechanical prep alone, oral antibiotic prep alone, or no prep.


81. Lefevre JH, Rodrigue CM, Mourra N, et al. Implication of MYH in colorectal polyposis. Ann Surg. 2006;244:874-879. MYH is a base excision repair gene that removes adenines from DNA that have been mispaired with a nucleotide product of oxidative reaction (8oxoG). Without MYH, mispaired adenines will lead to somatic transversions G:C→T:A in genes such as APC or K-ras. The authors reviewed 433 patients who underwent prophylactic surgery for colorectal polyposis over a 25-year time frame and evaluated 31 patients who did not have APC, finding biallelic MYH mutations in 6 patients. They conclude that MYH is a new gene responsible for approximately 1.4% of all adenomatous polyposis and 20% of polyposis without an APC mutation identified.
85. Rex DK, Ahnen DJ, Baron JA, et al. Serrated lesions of the colorectum: review and recommendations from an expert panel. Am J Gastroenterol. 2012;107:1315-1329. An expert panel provide their recommendation after a thorough literature review and 2-day conference. Hyperplastic polyps, sessile serrated adenomas/polyps, and traditional serrated adenomas are all included in the category of serrated lesions of the colon. This review outlines the histologic features of these lesions. Sessile serrated adenomas are predominantly flat, often with indistinct borders, right sided. They are associated with interval colon adenocarcinomas and are thought to progress to cancer, possibly through a promoter CpG island hypermethylation pathway. The panel recommends that these lesions are completely excised and points to the variability in detection rates between endoscopists, hypothesizing that difficulty in endoscopic detection is related to the high rate of subsequent cancer diagnosis in these patients.


132. Jayne DG, Thorpe HC, Copeland J, Quirke P, Brown JM, Guillou PJ. Five-year follow-up of the Medical Research Council CLASICC trial of laparoscopically assisted versus open surgery for colorectal cancer. Br J Surg. 2010;97(11):1638-1645. doi: 10.1002/bjs.7160. This study presents 5-year follow-up from the UK Medical Research Council CLASICC Trial comparing the safety and efficacy of laparoscopic vs. open surgery for colorectal cancer. Outcomes included overall and disease free survival and local, distant, and port site recurrences. The investigators found no differences between the two groups. They conclude, therefore, that long-term outcomes are not compromised by a laparoscopic approach.


150. Messick CA, Rodriguez-Bigas. Anal dysplasia. Surg Oncol Clin North Am. 2017;26:33-43. The authors review the literature on anal dysplasia with a focus on the natural history and the impact of ablative techniques with high resolution anoscopy on recurrence rates. High-risk patients are categorized as HIV-positive, immunosuppressed transplant recipients, and men who have sex with men. The authors emphasize the lack of literature in patients who are not non–high risk and overall emphasize close follow-up in all patients with high grade squamous intraepithelial lesion (HSIL)/high grade anal intraepithelial neoplasia (HGAIN).


160. Shao WJ, Li GC, Zhang ZH, Yang BL, Sun GD, Chen YQ. Systematic review and meta-analysis of randomized


166. Jacob TJ, Perakath B, Keighley MR. Surgical intervention for anorectal fistula. *Cochrane Database Syst Rev.* 2010;5:CD006319. This meta-analysis evaluated 10 randomized trials on outcomes after anorectal fistula procedures, with the ideal procedure stated to be associated with low recurrence rates, minimal incontinence, and good quality of life. Evaluated surgical interventions included: fistulotomy, fistulectomy, myo-mucosal advancement flap, fibrin glue, and cutting or loose setons. There were no differences in the procedures compared, except that fibrin glue had more recurrences. Trial sample sizes were noted to be small, and more recently discussed procedures such as the anal fistula plug and Ligation of Intersphincteric Fistula Tract (LIFT) were not evaluated in this review.


The Appendix
Fadi S. Dahdaleh, David Heidt, and Kiran K. Turaga

HISTORY
Although anatomists such as Vesalius and Leonardo Da Vinci had written about the appendix, Claudius Amyand in the early 18th century was the first surgeon to describe a successful appendectomy.1 In subsequent centuries, significant progress was made in the diagnosis and management of appendicitis, especially after Chester McBurney advocated for early appendectomy in his 1889 publication.2 Famously, the magician Harry Houdini died of a ruptured appendix after suffering a blow to his abdomen. Following the introduction and widespread use of antibiotics in the 1940s, mortality rates improved further. In 1982, Kurt Semm, a gynecologist, reported on the first laparoscopic appendectomy, which is now the most widely adopted technique.

EMBRYOLOGY, ANATOMY, AND HISTOLOGY
Previously considered a vestigial organ, the appendix is now linked to the development and preservation of gut-associated lymphoid tissue (GALT) and to the maintenance of intestinal flora. It has been suggested that appendectomy is associated with increased Clostridium difficile infections and increased subsequent cancer (colon, esophageal) as a result of microbial alteration, although this is currently unproven.3 The protective effect of an early appendectomy against development of ulcerative colitis has been proposed to be mechanistically linked to the release of dimeric forms of IgA from plasma B cells and the Th2 response mediated by IL-13–producing natural killer T cells.4

The appendix, along with the ileum and the colon, develops from the midgut and first appears at 8 weeks of gestation. As the gut rotates medially, the cecum becomes fixed in the right lower quadrant, thus determining the final position of the appendix. The appendix is a true diverticulum of the cecum as it contains all the histological layers of the colon, although certain differences in the irregularity of crypts remain. The average appendix measures 6 to 9 cm and derives its blood supply from the appendicular branch of the ileocolic artery. Visceral innervation occurs along the superior mesenteric plexus (T10-L1) and the vagus nerves. The appendix is intraperitoneal and retrocecal in location, but it can be pelvic (30%) and retroperitoneal (7%).5

ACUTE APPENDICITIS
Inflammation of the appendix is a significant public health problem with a lifetime incidence of 8.6% in men and 6.7% in women, with the highest incidence occurring in the second and third decade of life.6 While the rate of appendectomy in developed countries has decreased over the last several decades, it remains one of the most frequent emergent abdominal operations.7

1 Early obstruction leads to bacterial overgrowth of aerobic organisms in the early period, and subsequently, it leads to mixed flora. Obstruction generally leads to increased intraluminal pressure and referred visceral pain to the periumbilical region.8 It is postulated that this leads to impaired venous drainage, mucosal ischemia leading to bacterial translocation, and subsequent gangrene and intraperitoneal infection. Escherichia coli and Bacteroides fragilis are the most common aerobic and anaerobic bacteria isolated in perforated appendicitis.9,10
Key Points

1. Inflammation of the appendix is a significant public health problem with a lifetime incidence of 8.6% in men and 6.7% in women, with the highest incidence in the second and third decade of life. While the rate of appendectomy in developed countries has decreased over the last several decades, it remains one of the most frequent emergent abdominal operations.

2. The natural history of appendicitis is unclear, but it appears that progression to perforation is not predictable and that spontaneous resolution is common, suggesting that nonperforated and perforated appendicitis may, in fact, be different diseases.

3. C-reactive protein, bilirubin, IL-6, and procalcitonin have all been suggested to be helpful in the diagnosis of appendicitis, specifically in predicting perforated appendicitis.

4. Perforated appendicitis can be managed either operatively or nonoperatively. Immediate surgery is necessary in patients that appear septic, but this is usually associated with higher complications, including abscesses and enterocutaneous fistulae, due to dense adhesions and inflammation.

5. Single incision appendectomy has not been shown to improve outcomes, including cosmetic outcomes, in prospective randomized studies and has been suggested to have a higher incisional hernia rate.

6. While there is no evidence clearly evaluating long-term outcomes of patients undergoing incidental appendectomy with an asymptomatic appendix, the risk of adhesions and future complications after an appendectomy has been suggested to be higher than the risk of future appendicitis and increased economic costs. An incidental appendectomy is currently not advocated.

7. Older adult patients are at a higher risk for complications due to their premorbid conditions, and it is prudent to obtain definitive diagnostic imaging prior to taking patients to the operating room.

8. Patients with uncomplicated appendicitis do not require further antibiotics after an appendectomy, while patients with perforated appendicitis are treated with 3 to 7 days of antibiotics.

9. The most common mode of presentation for appendiceal carcinoma is that of acute appendicitis. Patients also may present with ascites or a palpable mass, or the neoplasm may be discovered during an operative procedure for an unrelated cause.

This sequence is not inevitable, however, and some episodes of acute appendicitis may resolve spontaneously. Due to differences in epidemiology, nonperforated and perforated appendicitis are considered different diseases. Additionally, since not all nonperforated appendicitis progresses to perforations, it is suggested that the pathogenesis of the two conditions may be different.

CLINICAL DIAGNOSIS

History

It is important to elicit an accurate history from the patient and/or family, in the case of pediatric patients. Inflammation of the visceral peritoneum usually progresses to the parietal peritoneum, presenting with migratory pain, which is a classic sign of appendicitis (likelihood ratio+, 2.06 [1.63–2.60]). Inflammation can often result in anorexia, nausea, vomiting, and fever (Table 30-1). Regional inflammation can also present with an ileus, diarrhea, small bowel obstruction, and hematuria. Pertinent negative history (including menstrual) must be obtained to rule out other etiologies of abdominal pain.

Physical Examination

Most patients lay quite still due to parietal peritonitis. Patients are generally warm to the touch (with a low-grade fever, ∼38.0°C [100.4°F]) and demonstrate focal tenderness with guarding. McBurney’s point, which is found one-third of the distance between the anterior superior iliac spine and the umbilicus, is often the point of maximal tenderness in a patient with an anatomically normal appendix. Certain physical signs with their respective eponyms can be helpful in discerning the location of the appendix: Rovsing’s sign, pain in the right lower quadrant after release of gentle pressure on left lower quadrant (normal position); Dunphy’s sign, pain with coughing (retrocecal appendix); obturator sign, pain with internal rotation of the hip (pelvic appendix); iliopsoas sign, pain with flexion of the hip (retrocecal appendix). In addition, pain with rectal or cervical examinations is also suggestive of pelvic appendicitis.

Laboratory Findings

Patients with appendicitis usually have leukocytosis of 10,000 cells/mm³, with a higher leukocytosis associated with gangrenous and perforated appendicitis (∼17,000 cells/mm³). C-reactive protein, bilirubin, IL-6, and procalcitonin have all been suggested to help in the diagnosis of appendicitis, specifically in predicting perforated appendicitis. The authors believe that a white blood cell (WBC) count and a C-reactive protein are two appropriate lab tests to obtain in the initial work up of appendicitis; a pregnancy test is also essential in women of childbearing age. Lastly, a urinalysis can be valuable in ruling out nephrolithiasis or pyelonephritis.

Imaging

Imaging is often utilized to confirm a diagnosis of appendicitis because a negative operation rate is acceptable in <10% of male patients and <20% of female patients. Routine use of cross-sectional imaging somewhat reduces the rate of negative laparotomies. Imaging studies are most appropriate for patients in whom a diagnosis of appendicitis is unclear or who are at high risk from operative intervention and general anesthesia, such as pregnant patients or patients with multiple comorbidities. Commonly utilized imaging modalities include computerized tomography (CT), ultrasonography (US), and magnetic resonance imaging (MRI).

CT Scan. A contrast-enhanced CT scan has a sensitivity of 0.96 (95% CI 0.95–0.97) and specificity of 0.96 (95% CI 0.93–0.97) in diagnosing acute appendicitis. Features on a CT scan that suggest appendicitis include enlarged lumen and double wall thickness (greater than 6 mm),
### Signs and symptoms of appendicitis

<table>
<thead>
<tr>
<th>Duration of symptoms (hours)</th>
<th>TRUE POSITIVE LIKELIHOOD RATIO</th>
<th>95% CONFIDENCE INTERVAL</th>
<th>TRUE NEGATIVE LIKELIHOOD RATIO</th>
<th>95% CONFIDENCE INTERVAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;9</td>
<td>1.01</td>
<td>0.97–1.05</td>
<td>0.94</td>
<td>0.62–1.42</td>
</tr>
<tr>
<td>&gt;12</td>
<td>0.96</td>
<td>0.90–1.04</td>
<td>1.19</td>
<td>0.87–1.63</td>
</tr>
<tr>
<td>&gt;24</td>
<td>0.65</td>
<td>0.47–0.90</td>
<td>1.47</td>
<td>1.14–1.90</td>
</tr>
<tr>
<td>&gt;48</td>
<td>0.49</td>
<td>0.36–0.67</td>
<td>1.20</td>
<td>1.08–1.34</td>
</tr>
</tbody>
</table>

Fever

<table>
<thead>
<tr>
<th>TRUE POSITIVE LIKELIHOOD RATIO</th>
<th>95% CONFIDENCE INTERVAL</th>
<th>TRUE NEGATIVE LIKELIHOOD RATIO</th>
<th>95% CONFIDENCE INTERVAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fever</td>
<td>1.64</td>
<td>0.89–3.01</td>
<td>0.61</td>
</tr>
</tbody>
</table>

Gastrointestinal dysfunction

<table>
<thead>
<tr>
<th>TRUE POSITIVE LIKELIHOOD RATIO</th>
<th>95% CONFIDENCE INTERVAL</th>
<th>TRUE NEGATIVE LIKELIHOOD RATIO</th>
<th>95% CONFIDENCE INTERVAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anorexia</td>
<td>1.27</td>
<td>1.14–1.41</td>
<td>0.59</td>
</tr>
<tr>
<td>Nausea</td>
<td>1.15</td>
<td>1.04–1.36</td>
<td>0.72</td>
</tr>
<tr>
<td>Vomiting</td>
<td>1.63</td>
<td>1.45–1.84</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Pain

<table>
<thead>
<tr>
<th>TRUE POSITIVE LIKELIHOOD RATIO</th>
<th>95% CONFIDENCE INTERVAL</th>
<th>TRUE NEGATIVE LIKELIHOOD RATIO</th>
<th>95% CONFIDENCE INTERVAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain migration</td>
<td>2.06</td>
<td>1.63–2.60</td>
<td>0.52</td>
</tr>
<tr>
<td>Pain progression</td>
<td>1.39</td>
<td>1.29–1.50</td>
<td>0.46</td>
</tr>
<tr>
<td>Direct tenderness</td>
<td>1.29</td>
<td>1.06–1.57</td>
<td>0.25</td>
</tr>
<tr>
<td>Indirect tenderness</td>
<td>2.47</td>
<td>1.38–4.43</td>
<td>0.71</td>
</tr>
<tr>
<td>Psoas sign</td>
<td>2.31</td>
<td>1.36–3.91</td>
<td>0.85</td>
</tr>
<tr>
<td>Rebound</td>
<td>1.99</td>
<td>1.61–2.45</td>
<td>0.39</td>
</tr>
<tr>
<td>Percussion tenderness</td>
<td>2.86</td>
<td>1.95–4.21</td>
<td>0.49</td>
</tr>
<tr>
<td>Guarding</td>
<td>2.48</td>
<td>1.60–3.84</td>
<td>0.57</td>
</tr>
<tr>
<td>Rigidity</td>
<td>2.96</td>
<td>2.43–3.59</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Temperature (degrees centigrade)

<table>
<thead>
<tr>
<th>TRUE POSITIVE LIKELIHOOD RATIO</th>
<th>95% CONFIDENCE INTERVAL</th>
<th>TRUE NEGATIVE LIKELIHOOD RATIO</th>
<th>95% CONFIDENCE INTERVAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;37.7</td>
<td>1.57</td>
<td>0.90–2.76</td>
<td>0.65</td>
</tr>
<tr>
<td>&gt;38.5</td>
<td>1.87</td>
<td>0.66–5.32</td>
<td>0.89</td>
</tr>
</tbody>
</table>

White blood cells (10⁹/L)

<table>
<thead>
<tr>
<th>TRUE POSITIVE LIKELIHOOD RATIO</th>
<th>95% CONFIDENCE INTERVAL</th>
<th>TRUE NEGATIVE LIKELIHOOD RATIO</th>
<th>95% CONFIDENCE INTERVAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥10</td>
<td>4.20</td>
<td>2.11–8.35</td>
<td>0.20</td>
</tr>
<tr>
<td>≥15</td>
<td>7.20</td>
<td>4.31–12.00</td>
<td>0.66</td>
</tr>
</tbody>
</table>

C-reactive protein (mg/L)

<table>
<thead>
<tr>
<th>TRUE POSITIVE LIKELIHOOD RATIO</th>
<th>95% CONFIDENCE INTERVAL</th>
<th>TRUE NEGATIVE LIKELIHOOD RATIO</th>
<th>95% CONFIDENCE INTERVAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥10</td>
<td>1.97</td>
<td>1.58–2.45</td>
<td>0.32</td>
</tr>
<tr>
<td>≥20</td>
<td>2.39</td>
<td>1.67–3.41</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Conclusions: Individually, disease history, clinical findings, and laboratory tests are weak. But when combined, they yield high discriminatory power.


Wall thickening (greater than 2 mm), periappendiceal fat stranding, appendiceal wall thickening, and/or an appendicolith (Fig. 30-1). While there remains a concern of ionizing radiation exposure with a CT scan, typical low-dose CT scans result in exposure of 2 to 4 mSv, which is not significantly higher than background radiation (3.1 mSv). Recent trials have also suggested that although low-dose CT scans of 2 mSv do not generate high-resolution images, using these lower resolution images does not affect clinical outcomes. Intravenous contrast is generally preferred in these studies, but it can be avoided in patients with allergies or low estimated glomerular filtration rate (less than 30 mL/minute for 1.73 m²). Several meta-analyses have suggested that CT scan is more sensitive and specific than ultrasound in diagnosing appendicitis.

**Ultrasound.** Ultrasonography has a sensitivity of 0.85 (95% CI 0.79–0.90) and a specificity of 0.90 (95% CI 0.83–0.95). Graded compression ultrasonography is used to identify the anteroposterior diameter of the appendix. An easily compressible appendix <5 mm in diameter generally rules out appendicitis. Features on an ultrasound that suggest appendicitis include a diameter of greater than 6 mm, pain with compression, presence of an appendicolith, increased echogenicity of the fat, and periappendiceal fluid. Ultrasound is cheaper and more readily available than CT scan, and it does not expose patients to ionizing radiation, but it is user-dependent and has limited utility in obese patients. In addition, graded compression is usually

*Figure 30-1. McBurney’s point. 1 = anterior superior iliac spine; 2 = umbilicus; x = McBurney’s point.*
## Table 30-2

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>No. of Studies</th>
<th>No. of Patients CT</th>
<th>No. of Patients US</th>
<th>Sensitivity CT</th>
<th>Sensitivity US</th>
<th>Specificity CT</th>
<th>Specificity US</th>
</tr>
</thead>
<tbody>
<tr>
<td>Terasawa</td>
<td>2004</td>
<td>22</td>
<td>1172</td>
<td>NR</td>
<td>94% (CI: 91%–95%)</td>
<td>86% (CI: 83%–88%)</td>
<td>95% (CI: 93%–96%)</td>
<td>81% (CI: 78%–84%)</td>
</tr>
<tr>
<td>Weston</td>
<td>2005</td>
<td>21</td>
<td>1516</td>
<td>NR</td>
<td>97% (CI: 95%–98%)</td>
<td>87% (CI: 85%–89%)</td>
<td>95% (CI: 93%–96%)</td>
<td>93% (CI: 92%–96%)</td>
</tr>
<tr>
<td>Doria</td>
<td>2006</td>
<td>57</td>
<td>NR</td>
<td>5039</td>
<td>94% (CI: 92%–97%)</td>
<td>88% (CI: 86%–90%)</td>
<td>93% (CI: 92%–95%)</td>
<td>93% (CI: 92%–94%)</td>
</tr>
<tr>
<td>Al-Khayal</td>
<td>2007</td>
<td>25</td>
<td>13697</td>
<td>NR</td>
<td>93% (CI: 92%–95%)</td>
<td>84% (CI: 82%–85%)</td>
<td>91% (CI: 84%–95%)</td>
<td>78% (CI: 67%–86%)</td>
</tr>
<tr>
<td>Van Randen</td>
<td>2008</td>
<td>6</td>
<td>NR</td>
<td>671</td>
<td>91% (CI: 84%–95%)</td>
<td>78% (CI: 67%–86%)</td>
<td>90% (CI: 85%–94%)</td>
<td>83% (CI: 76%–88%)</td>
</tr>
</tbody>
</table>

- **SUMMARY**
  - CT more sensitive than US in five of five meta-analyses.
  - CT more specific than US in four of five meta-analyses.
  - CT has superior positive predictive value in one of two meta-analyses.
  - CT has superior negative predictive value in both meta-analyses.
  - CT is more accurate in the one study reporting results.

*CI = confidence interval; CT = computed tomography; NR = not reported; US = ultrasonography.*
painful for patients with peritonitis. A comparison of the efficacy of ultrasound v. CT scan is found in Table 30-2.

**MRI.** MRI of the abdomen has a sensitivity of 0.95 (95% CI 0.88–0.98) and specificity of 0.92 (95% CI 0.87–0.95) for identification of acute appendicitis. MRI is an expensive test that requires significant expertise to perform and interpret and is usually recommended in patients for whom the risk of ionizing radiation outweighs the relative ease of obtaining a contrast CT scan, i.e., pregnant or pediatric patients.

**Differential Diagnosis**

Causes of acute abdominal pain that are often confused with acute appendicitis include acute mesenteric adenitis, cecal diverticulitis, Meckel’s diverticulitis, acute ileitis, Crohn’s disease, acute pelvic inflammatory disease, torsion of ovarian cyst or graafian follicle, and acute gastroenteritis. Frequently, no organic pathology is identified. Obtaining an antecedent history of a viral infection (mesenteric adenitis or gastroenteritis) and a cervical exam in women (exquisite tenderness with motion in pelvic inflammatory disease) are essential before planning any intervention. Detailed menstrual history can distinguish mittelschmerz (no fever or leukocytosis, mid-menstrual cycle pain) and ectopic pregnancies.

**MANAGEMENT OF APPENDICITIS**

**Uncomplicated Appendicitis**

The preferred approach to manage patients with uncomplicated appendicitis is an appendectomy. Several recent randomized trials and cohort studies have examined the role of nonoperative management of adult patients with appendicitis. In each of these well-designed studies with noninferiority as the endpoint, patients were randomized to either receiving antibiotics or undergoing an appendectomy, which was frequently performed open. A majority of the patients in the nonoperative arm received intravenous antibiotics for a short course followed by a course of a fluoroquinolone and metronidazole, or oral amoxicillin/clavulanic acid. Meta-analysis of the published data found that 26.5% of patients in the nonoperative group required an appendectomy within 1 year. In addition, the rate of adverse events following antibiotics therapy was higher (relative risk [RR] 3.18, 95% CI 1.63–6.21, \( P = 0.0007 \)) and patients who recurred presented more frequently with complicated appendicitis (RR 2.52, 95% CI 1.17–5.43, \( P = 0.02 \)). Currently, conservative management can be offered to informed patients using techniques of shared decision-making, but it is not the standard modality of management of appendicitis, except in patients with significant phobia of surgery. Societal costs and long-term implications of the conservative strategy have not yet been completely evaluated.

**Timing of Surgery.** Emergent surgery is often performed in patients with appendicitis, but studies have evaluated the performance of urgent surgery (waiting less than 12 hours) in a semielective setting after administering antibiotics upon admission. The studies did not reveal any significant difference in outcomes, except for a slightly longer hospital stay in those undergoing urgent surgery. Currently, delaying surgery less than 12 hours is acceptable in patients with short duration of symptoms (less than 48 hours) and in nonperforated, nonangrenous appendicitis.

**Approach of Surgery.** Numerous meta-analyses comparing laparoscopic to open appendectomy have demonstrated relative equivalence of the techniques, with laparoscopic appendectomy resulting in a shorter length of stay (LOS), faster return to work, and lower superficial wound infection rates, especially in obese patients. Open appendectomy results in shorter operative times and lower intra-abdominal infection rates. Costs of the two techniques are relatively similar because of the offset of costs in laparoscopic techniques by shorter LOS. In the United States, laparoscopic appendectomies are increasingly utilized.

**Complicated Appendicitis**

Perforated and gangrenous appendicitis and appendicitis with abscess or phlegmon formation are considered complicated conditions. Patients with perforated appendicitis usually present after 24 hours of onset, although 20% of patients present within 24 hours. Such patients are often acutely ill and dehydrated and require resuscitation. Usually, the perforated abscess is walled off in the right lower quadrant, although retroperitoneal abscesses including psoas abscess, liver abscesses, fistulas, and pylephlebitis (portal vein inflammation) can also occur when left untreated.

Perforated appendicitis can be managed either operatively or nonoperatively. Immediate surgery is necessary in patients that appear septic, but this is usually associated with higher complications, including abscesses and enterocutaneous fistulae due to dense adhesions and inflammation.

The management of long-duration, complicated appendicitis is often staged. Patients are resuscitated and treated with IV antibiotics. Patients with longstanding perforation are better treated with adequate percutaneous image-guided drainage. This strategy is successful in 79% of patients who achieve complete resolution, which occurs more often in lower-grade abscesses, transgluteal drainage, and with CT- (vs. ultrasound-) guided drainage. Operative intervention is performed in patients who fail conservative management and in patients with free intraperitoneal perforation.

**Interval Appendectomy.** The majority of patients with perforated appendicitis (80%) have resolution of their symptoms with drainage and antibiotics. There remains debate about the value of performing an interval appendectomy 6 to 8 weeks after the original inflammatory episode. Proponents of this approach cite the incidence of recurrent appendicitis (7.4%–8.8%) and the presence of appendiceal neoplasms detected on the appendectomy (relevant benign lesions 0.7%, malignant lesions 1.3%). Opponents cite the high incidence of no future events after a median follow-up of 34 months in 91% of patients. Currently, shared decision-making is necessary before proceeding with an interval appendectomy.

**OPERATIVE INTERVENTION**

**Preoperative Preparation**

Once the decision to proceed with surgical intervention is made, patients can be taken to the operating room rather expeditiously. While resuscitative efforts are important in patients who present with significant dehydration or in a compromised host, the majority of patients can be taken to the operating room within a short interval. Placement of a Foley catheter is optional but not necessary while performing an appendectomy. Preoperative antibiotics must be administered at least 30 to 60 minutes prior to skin incision. The choice of antibiotics include cefoxitin, ampicillin/sulbactam, and cefazolin plus metronidazole for
uncomplicated appendicitis. Patients with β-lactam allergies can be given clindamycin in combination with a fluoroquinolone, gentamicin, or aztreonam. Postoperative antibiotics are usually not necessary.

In patients with perforated appendicitis undergoing operative intervention, preoperative antibiotics are necessary to cover gram-negative bacteria and anaerobes. Mono-therapy with piperacillin/tazobactam or combination of cephalosporin with metronidazole are reasonable choices. The duration of postoperative antibiotics is generally less than 4 days once complete source control has been achieved (STOP-IT trial). Patients with incomplete drainage, persistent catheters, complications from surgery, and uncertain resolution of inflammation might need a longer duration of antibiotics.

Operative Technique

Open Appendectomy. An open appendectomy is usually performed under general anesthesia, although regional anesthesia can be used. After wide prep and drape, an incision is usually made on McBurney’s point either in an oblique fashion (McBurney’s incision) or transverse incision (Rocky-Davis incision). A lower midline laparotomy incision is more appropriate for perforated appendicitis with a phlegmon. A muscle-splitting approach can be utilized to access the peritoneum in patients that are well paralyzed. The bed is positioned in Trendelenburg’s with the left side down. The appendix is usually readily identified, but if necessary, it can be found by tracing the anterior taenia (taenia Libera) of the cecum distally. We generally ligate the mesentery early to allow better exposure. If the base of the appendix is viable, ligating the appendix is acceptable. This can be imbricated with a Z-stitch or purse string configuration, or alternatively the mucosa can be fulgurated. In the event of retraction of the appendiceal artery or unexpected bleeding, the incision can be extended medially (Fowler extension). Skin closure is usually performed in a layered fashion, but in cases with significant abscess or contamination, closure by secondary intention or delayed primary closure has been considered. Recent trials have suggested no difference in surgical site infection rates between primary and delayed primary closure. Placement of surgical drains has not been proven to be beneficial in multiple clinical trials for either complicated or uncomplicated appendicitis.

Laparoscopic Appendectomy. Patients undergoing laparoscopic appendectomy are positioned supine with the left arm tucked for better access. Monitors and assistants are positioned appropriately. Access to the peritoneum can be obtained using either the Hasson technique in a periumbilical fashion or with a Verrees or optical trocar in the left upper quadrant 3 cm below the costal margin in the midclavicular line. Five-mm ports are usually placed in the suprapubic and left lower quadrant areas. It is also technically feasible to place the third port in the right upper quadrant. The bed is positioned in Trendelenburg, with the left side down to sweep the bowel away. The appendix is grasped and elevated upwards to identify the window between the mesoappendix and the cecum (Fig. 30-2). Occasionally, it is essential to release the mesenteric attachments of the cecum to mobilize a retrocecal or pelvic appendix to obtain this view. Using a Maryland grasper, the window is created, and the mesoappendix is divided with cautery, clip, or a bipolar energy source. The base of the appendix is divided either with an endoscopic stapler or after placing an endoloop. In the case of a nonviable appendix base, a staple line through the cecum that avoids the ileocecal valve might be sufficient, unless significant inflammation is present. The appendix is retrieved through the midline port in a specimen bag, especially if an appendiceal lesion is suspected. If a periappendiceal phlegmon is encountered or if the operation is being performed for perforated appendicitis, careful sweeping of the bowel with a blunt dissector can release the appendix. It is important to carefully separate adjacent bowel, which can be friable in such settings. Conversion to open surgery should be considered for failure to progress. Typically, once the base of the appendix is identified, it is generally more helpful to divide the stump first. An endoscopic stapler or endoloop can be used for the base, provided the base is viable. Occasionally, an ileocecectomy is necessary when resection of the base of the appendix or cecum is likely to impinge on the ileocecal valve. The mesoappendix is similarly divided with either a stapler with thin leg length staples, a clip, cautery, or energy device.

Novel Techniques

Three novel techniques have been investigated in the performance of an appendectomy: single incision appendectomy, natural orifice transluminal endoscopic surgery (NOTES), and robotic appendectomy. Single incision appendectomy has not been shown to improve outcomes, including cosmetic outcomes, in prospective randomized studies and has been suggested to have a higher incisional hernia rate. NOTES surgery has been shown to have better cosmetic outcome and less postoperative pain in a meta-analysis of NOTES procedures including appendectomies, although only 40 patients were included in the analysis. The risk of luminal contamination and closure of enteral or vaginal mucosa remain suboptimal; for this reason, there has not been widespread dissemination of this technique. Robotic appendectomy allows flexible motions of intraperitoneal instruments and is therefore superior in ergonomics for the surgeon. However, it is extremely expensive and requires larger ports based on most of the current platforms; thus, this technique is also not utilized widely.

Negative Exploration

Upon performing a laparoscopy or laparotomy for suspected appendicitis, if one finds no evidence of appendicitis, a thorough exploration of the peritoneum must be performed to rule out contributing pathology. A normal appendix is often removed to reduce future diagnostic dilemma. Management of incidentally found common conditions is summarized in Table 30-3.

Incidental Appendectomy

The practice of prophylactic appendectomy has been considered during other operations to prevent the future risk of appendicitis. It is routinely performed in children undergoing chemotherapy, compromised hosts with an unclear physical exam, patients with Crohn’s disease with a normal cecum, patients traveling to remote places with no urgent care, and in patients undergoing cytoreductive operations for ovarian malignancies. While there is no evidence clearly evaluating long-term outcomes of patients undergoing incidental appendectomy with an asymptomatic appendix, the risk of adhesions and future complications after an appendectomy has been suggested to be higher than the risk of future appendicitis and increased economic costs. For these reasons, an incidental appendectomy is currently not advocated.
SPECIAL CIRCUMSTANCES

Appendicitis in Children

Almost 1 in 8 children undergo a workup for the diagnosis of appendicitis.\textsuperscript{60,61} Of these, infants and young children are most likely to present with perforated disease (51\%–100\%), while school-age children have lower rates of perforation.\textsuperscript{62,63,64} While most age groups demonstrate the same symptoms previously described in adults, neonates can also present with abdominal distension and lethargy or irritability. The Pediatric Appendicitis Score has components similar to the Alvarado Score and is scored of 10 points, with maximum weight (2 points each) for right lower quadrant tenderness and pain with cough, percussion or hopping. A score of 7 or greater indicates that the patient has a high chance of having appendicitis (78\%–96\% percent).\textsuperscript{65}

In the pediatric population, special considerations must be made to exclude relevant differential diagnoses such as intussusception (currant jelly stools, abdominal mass), gastroenteritis (often no leukocytosis), malrotation (pain out of proportion), pregnancy (ectopic), mesenteric adenitis, torsion of the omentum, and ovarian or testicular torsion.

<table>
<thead>
<tr>
<th>Table 30-3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Management of Intraoperative Findings Mimicking Appendicitis</strong></td>
</tr>
<tr>
<td><strong>Ovarian Torsion</strong></td>
</tr>
<tr>
<td><strong>Crohn’s terminal ileitis</strong></td>
</tr>
<tr>
<td><strong>Meckel’s diverticulitis</strong></td>
</tr>
<tr>
<td><strong>Appendiceal Mass</strong></td>
</tr>
</tbody>
</table>
With regard to the management of children with appendicitis, early appendicitis is treated preferably with a laparoscopic appendectomy, which has better outcomes than open appendectomies in children.66,67 For patients with complicated appendicitis, urgent appendectomy is advocated in the setting of no abscess or mass. Laparoscopic appendectomy appears to retain its benefits in this setting as well.68,69 In the setting of a perforation, antibiotics are continued after surgery for at least 3 days, and preferably 5 days (APSA guidelines).70,71 Management of perforated appendicitis with abscess is similar to adults, although no adverse effects of an early laparoscopic appendectomy have been seen even in this setting.72,73,74

Nonoperative management of appendicitis has also been studied in children.75-77 It may be safe for children with early presentation (less than 48 hours), limited inflammation (WBC less than 18,000/cu.ml), appendicoliths, and no evidence of rupture on imaging.78 Patients are usually administered IV antibiotics until inflammation reduces and then transitioned to oral antibiotics.79 This is usually effective in reducing inflammation (88%–92%), but has a recurrence rate of 22% at 1 year and increased resource utilization.80

Appendicitis in Older Adults
Older adult patients can have diminished inflammation and thus present with perforation or abscess more frequently.81,82

Such patients are at a higher risk for complications because of their premorbid conditions, and it is more prudent to obtain definitive diagnostic imaging prior to taking patients to the operating room. Laparoscopic appendectomy is safe and might allow patients to reduce pain and their hospital stay.83

Appendicitis in Pregnancy
Appendicitis occurs in 1 in 800 to 1 in 1000 pregnancies, mostly in the first and second trimesters. Its incidence is rare in the antepartum state, and it can occur in the postpartum state in geriatric pregnancies (maternal age greater than 35 years).84 While the majority of the clinical features are similar, patients can also present with heartburn, bowel irregularity, flatulence, or a change in bowel habits. The point of maximum tenderness is usually displaced on physical exam. Ultrasonography is the preferred imaging modality, although nonvisualization can occur. Sensitivity can vary from 67% to 100%, and specificity varies from 93% to 96%. An alternative imaging modality is MRI, with a sensitivity of 94% and specificity of 97%.85 While CT can be performed in pregnancy, the risk of fetal irradiation leads many practitioners to avoid it unless other modalities are inconclusive.86 When discussing options with the patient and the patient’s family, it is important to note that the risk of fetal loss is up to 36% if appendiceal perforation occurs.87 Therefore, there remains a lower threshold to operate on such patients, with an acceptable negative exploration rate of as high as 30%. Laparoscopic appendectomies can be safely performed in pregnant patients, although studies suggest a variable but reproducible higher rate of fetal loss (around 7% vs. 3%) than open techniques. Lower intra-abdominal pressures (10–12 mmHg) during insufflation have been suggested to reduce early labor. Nonoperative management has also been proposed for pregnant patients, but treatment failure rates have been reported as high as 25%.

Chronic or Recurrent Appendicitis
Patients with recurrent right lower quadrant abdominal pain not associated with a febrile illness with imaging findings suggestive of an appendicolith or dilated appendix are classified as having chronic appendicitis.88 Patients often report resolution of symptoms with an appendectomy. In the absence of imaging abnormalities, prophylactic appendectomy is not encouraged.45

**OUTCOMES AND POSTOPERATIVE COURSE**

Appendectomy is a relatively safe procedure with an extremely low mortality rate (less than 1%). The commonest adverse events include soft tissue infections, either superficial or deep (including abscesses). Patients with uncomplicated appendicitis do not require further antibiotics after an appendectomy, while patients with perforated appendicitis are treated with 3 to 7 days of antibiotics (4 days from the STOP-IT trial).89 Patients with wound infections can be managed with simple wound opening and packing, and delayed primary closure has not been shown to be beneficial.90 In laparoscopic cases, these are usually the periumbilical ports.91 Patients with deep space abscesses are managed with percutaneous drainage and antibiotics. Fistulas (appendicocutaneous or appendico-vesicular) are managed conservatively as the first step. Bowel obstructions and infertility are infrequent but reported.

**Stump Appendicitis**

An uncommon complication after surgery is the development of appendicitis in an incompletely excised appendiceal stump (greater than 0.5 cm stump length). Optimal management requires reexcision of the appendiceal base, but diagnosis can be difficult and requires careful assessment of the patient’s history, physical exam, and imaging studies.92 Use of the “appendiceal critical view” (appendix placed at 10 o’clock, taenia coli/libera at 3 o’clock, and terminal ileum at 6 o’clock) and identification of where the taeniae coli merge and disappear is paramount to identifying and ligating the base of the appendix during the initial operation (Fig. 30-3). In patients who have had prior appendectomy, a low index of suspicion is important to prevent delay in diagnosis and complications. Prior appendectomy should not be an absolute criterion in ruling out acute appendicitis.

**Appendiceal Neoplasms**

The incidence of appendiceal neoplasms is estimated at around 1% of all appendectomy specimens, although the true incidence of appendiceal neoplasms is not known.93 Neoplasms that occur in the appendix are predominantly gastroenteropancreatic neuroendocrine tumors (or GEP-NETs, previously called carcinoids), mucinous neoplasms, or adenocarcinomas. Almost one-third of the neoplasms of the appendix present with acute appendicitis, while the others are often incidentally detected or are detected after regional spread of disease.94

**Gastroenteropancreatic Neuroendocrine Tumors (GEP-NETs or Carcinoid)**

Appendiceal carcinoid tumors are submucosal rubbery masses that are detected incidentally on the appendix.95 Carcinoid tumors of the appendix are relatively indolent but can develop nodal or hepatic metastases.96 Infrequently, these can be associated with a carcinoid syndrome if there are hepatic metastases (2.9%).100 Upon incidental findings of a suspected carcinoid, the surgeon must evaluate the nodal basin along the ileocolic pedicle and also examine the liver for any signs of metastases. For lesions that are less than 1 cm (95% of all lesions), a negative margin appendectomy is adequate. For tumors 2 cm or larger, a right hemicolectomy is recommended. For lesions 1 to 2 cm in size, there is no consensus on a completion colectomy. A right colectomy is often performed for mesenteric invasion, enlarged
nodes, or positive or unclear margins. Measurement of serum chromogranin A is recommended.

**Goblet Cell Carcinomas**

These lesions were mistakenly called goblet cell carcinoids, implying a rather indolent biology, while goblet cell carcinomas are adenocarcinoid with both adenocarcinoma and neuroendocrine features. Such lesions carry a worse prognosis than carcinoids but slightly better than adenocarcinomas. There is a high risk of peritoneal recurrence in such cases. For incidentally detected lesions, a systematic surveillance of the peritoneum must be performed, and a peritoneal cancer index score must be documented if disease is present. In the absence of metastatic disease, a right hemicolecotomy is generally appropriate, although some advocate for a right colectomy only for tumors 2 cm or larger.

**Lymphomas**

Appendiceal lymphomas are rare (1%–3% of lymphomas, usually non-Hodgkin’s) and difficult to diagnose preoperatively (appendiceal diameter can be 2.5 cm or larger). Management includes an appendectomy in most cases.

**Adenocarcinoma**

Primary adenocarcinoma of the appendix is a rare neoplasm with three major histologic subtypes: mucinous adenocarcinoma, colonic adenocarcinoma, and adenocarcinoid. The most common mode of presentation for appendiceal carcinoma is acute appendicitis. Patients also may present with ascites or a palpable mass, or the neoplasm may be discovered during an operative procedure for an unrelated cause. The recommended treatment for all patients with adenocarcinoma of the appendix is a formal right hemicolecotomy. Appendiceal adenocarcinomas have a propensity for early perforation, although they are not clearly associated with a worsened prognosis. Overall 5-year survival is 55% and varies with stage and grade. Patients with appendiceal adenocarcinoma are at significant risk for both synchronous and metachronous neoplasms, approximately half of which will originate from the gastrointestinal tract.

**Appendiceal Mucoceles and Mucinous Neoplasms of the Appendix**

The term *appendiceal mucocele* broadly describes a mucus-filled appendix that could be secondary to neoplastic or nonneoplastic pathologies (mucosal hyperplasia, simple or retention cysts, mucinous cystadenomas, mucinous cystadenocarcinoma). The most common form of presentation is incidental; however, presentation with appendicitis occurs in a third of cases. On cross-sectional imaging, a low attenuation, round, well encapsulated cystic mass in the right or quadrant is often encountered, and features such as wall irregularity and soft tissue thickening are suggestive of a neoplastic process. It is important to carefully assess for the presence of ascites, peritoneal disease, and scalloping of the liver surface on imaging upon initial evaluation. A reliable diagnosis cannot be established using imaging alone, and it is recommended that surgical excision without capsular disruption is undertaken. The importance of careful handling of a mucocele and the avoidance of rupture cannot be overemphasized because the intraperitoneal spread of neoplastic cells at subsequent development of pseudomyxoma peritonei are nearly certain in cases of adenocarcinoma. When suspecting a mucinous neoplasm of the appendix, it is imperative to systematically examine the peritoneum and document a peritoneal cancer index score if mucin is present. Biopsies to examine the content of epithelial cell, neoplastic cells, and mucin can be helpful.

In cases where a homogeneous cyst without nodularity or signs of dissemination is encountered, laparoscopic excision is acceptable, provided that a stapler is fired across the base of the cecum to avoid a positive margin. The specimen should be placed in a plastic bag and carefully removed through a small incision. In the absence of mesenteric or peritoneal involvement, an appendectomy with concurrent appendiceal lymphadenectomy is sufficient, as the chances of lymph node involvement are quite low. If peritoneal spread is evident upon exploration, it is important to obtain biopsies and document the peritoneal disease burden. An appendectomy is acceptable if the patient has acute appendicitis, but suboptimal debulking is discouraged. In addition, colorectal, ovarian, and endometrial cancers can coexist in the setting of appendiceal mucocles, and careful examination of intra-abdominal structures is important.

When there is discordance between the primary lesion histology and the peritoneum, the peritoneal histology is usually given priority. For instance, if patients had a neoplasm in the appendix but adenocarcinoma in the peritoneum, the patient would be considered as having adenocarcinoma (AJCC M1b) disease. The recent AJCC 8th edition and the PSOGI 2016...
Table 30-4
AJCC 8th edition and the PSOGI 2016 classification consensus of mucinous neoplasia of the appendix

<table>
<thead>
<tr>
<th>LESION</th>
<th>PERITONEAL DISEASE AT DIAGNOSIS</th>
<th>PROGNOSIS</th>
<th>TREATMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-grade appendiceal mucinous neoplasm (LAMN)</td>
<td>Confined to the appendix</td>
<td>Excellent-curative</td>
<td>Negative margin appendectomy, rarely need ileocecectomy</td>
</tr>
<tr>
<td>LAMN</td>
<td>Peri-appendiceal Acellular mucin dissecting through the wall (t4a) or adjacent organs (t4b)</td>
<td>Excellent-low risk of recurrence</td>
<td>Negative margin appendectomy, resection of acellular mucin</td>
</tr>
<tr>
<td>LAMN</td>
<td>Peri-appendiceal Epithelial cells dissecting through the wall (t4a) or adjacent organs (t4b)</td>
<td>Excellent-high risk of recurrence</td>
<td>Negative margin appendectomy, peritoneal surveillance with second look laparoscopy vs. HIPEC</td>
</tr>
<tr>
<td>LAMN</td>
<td>Distant epithelial cells or acellular mucin (M1a) Low grade mucinous carcinoma peritonei</td>
<td>Excellent-high risk of recurrence</td>
<td>Negative margin appendectomy, omentectomy, HIPEC</td>
</tr>
<tr>
<td>High-grade appendiceal mucinous neoplasm (HAMN-rare)</td>
<td>Management is identical to a LAMN with risk stratification as shown above but slightly worse prognosis.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mucinous adenocarcinoma</td>
<td>Confined to the appendix</td>
<td>Very Good</td>
<td>Right hemicolectomy</td>
</tr>
<tr>
<td>Mucinous adenocarcinoma</td>
<td>Peritoneal Dissemination High grade mucinous carcinoma peritonei with or without signet ring cells</td>
<td>Well Differentiated-Very good Moderately differentiated –Good Poorly differentiated/signet ring cell histology: 10 year survival of 10-20%</td>
<td>Cytoreductive surgery and HIPEC, with systemic chemotherapy for high grade histologies</td>
</tr>
<tr>
<td>Adenocarcinoma (non-mucinous, including goblet cell histology)</td>
<td>Management identical to the mucinous histologies, with more extensive use of systemic chemotherapy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serrated Adenoma (rare)</td>
<td>Confined to appendix</td>
<td>Excellent-curative</td>
<td>Appendectomy</td>
</tr>
<tr>
<td>Adenoma (rare)</td>
<td>Confined to appendix</td>
<td>Excellent-curative</td>
<td>Appendectomy</td>
</tr>
</tbody>
</table>


Pseudomyxoma Peritonei Syndrome

Patients with appendiceal mucinous neoplasms develop peritoneal dissemination leading to pseudomyxoma peritonei (PMP) syndrome. This can occur in gastric, ovarian, pancreatic, and colorectal primary tumors as well. Patients with this syndrome can have varied prognosis ranging from curative to palliative. Cytoreductive surgery and hyperthermic intraperitoneal chemotherapy (HIPEC) are considered the standard of care for patients with PMP syndrome from appendiceal primaries.

Early detection and management of limited peritoneal disease is favorable and preferred as opposed to extensive intraperitoneal mucin development. The surgical technique involves parietal and visceral peritoneectomies, and intraperitoneal administration of heated (42°C [108°F]) chemotherapy (usually mitomycin) in the abdomen. Previously considered a morbid surgery, high volume centers and standardized practices have made the morbidity and mortality similar to any major open GI procedure. This technique can also be performed laparoscopically when the disease is detected early and is low volume.

REFERENCES

Entries highlighted in bright blue are key references.

1. Amyand C. Of an inguinal rupture, with a pin in the appendix coeci, incrusted with stone; and some observations on wounds in the guts. Phil Trans. 1735;39:329-342.
4. Sahami S, Kooij IA, Meijer SL, Van den Brink GR, Buskens CJ, Te Velde AA. The link between the appendix and ulcerative


# Liver

**David A. Geller, John A. Goss, Ronald W. Busuttil, and Allan Tsung**

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>History of Liver Surgery</td>
<td>1346</td>
</tr>
<tr>
<td>Liver Anatomy</td>
<td>1346</td>
</tr>
<tr>
<td>Segmental Anatomy</td>
<td>1346</td>
</tr>
<tr>
<td>Hepatic Artery</td>
<td>1348</td>
</tr>
<tr>
<td>Portal Vein</td>
<td>1348</td>
</tr>
<tr>
<td>Hepatic Veins and Inferior Vena Cava</td>
<td>1349</td>
</tr>
<tr>
<td>Bile Duct and Hepatic Ducts</td>
<td>1350</td>
</tr>
<tr>
<td>Neural Innervation and Lymphatic Drainage</td>
<td>1350</td>
</tr>
<tr>
<td>Liver Physiology</td>
<td>1351</td>
</tr>
<tr>
<td>Bilirubin Metabolism</td>
<td>1352</td>
</tr>
<tr>
<td>Formation of Bile</td>
<td>1352</td>
</tr>
<tr>
<td>Drug Metabolism</td>
<td>1352</td>
</tr>
<tr>
<td>Liver Function Tests</td>
<td>1352</td>
</tr>
<tr>
<td>Hepatocellular Injury</td>
<td>1353</td>
</tr>
<tr>
<td>Abnormal Synthetic Function</td>
<td>1353</td>
</tr>
<tr>
<td>Cholestasis</td>
<td>1353</td>
</tr>
<tr>
<td>Jaundice</td>
<td>1353</td>
</tr>
<tr>
<td>Molecular Signaling Pathways in the Liver</td>
<td>1354</td>
</tr>
<tr>
<td>Acute Phase Reaction</td>
<td>1354</td>
</tr>
<tr>
<td>LPS Signaling</td>
<td>1355</td>
</tr>
<tr>
<td>Nitric Oxide</td>
<td>1355</td>
</tr>
<tr>
<td>Heme Oxygenase System</td>
<td>1356</td>
</tr>
<tr>
<td>Toll-Like Receptors</td>
<td>1357</td>
</tr>
<tr>
<td>Radiologic Evaluation of the Liver</td>
<td>1357</td>
</tr>
<tr>
<td>Ultrasound</td>
<td>1357</td>
</tr>
<tr>
<td>Computed Tomography</td>
<td>1358</td>
</tr>
<tr>
<td>Magnetic Resonance Imaging</td>
<td>1359</td>
</tr>
<tr>
<td>Positron Emission Tomography</td>
<td>1360</td>
</tr>
<tr>
<td>Acute Liver Failure</td>
<td>1360</td>
</tr>
<tr>
<td>Etiology</td>
<td>1361</td>
</tr>
<tr>
<td>Clinical Presentation</td>
<td>1361</td>
</tr>
<tr>
<td>Diagnosis and Clinical Management</td>
<td>1361</td>
</tr>
<tr>
<td>Prognosis</td>
<td>1362</td>
</tr>
<tr>
<td>Liver Transplantation</td>
<td>1362</td>
</tr>
<tr>
<td>Emerging Technologies</td>
<td>1362</td>
</tr>
<tr>
<td>Cirrhosis and Portal Hypertension</td>
<td>1362</td>
</tr>
<tr>
<td>Morphologic Classification of Cirrhosis</td>
<td>1362</td>
</tr>
<tr>
<td>Etiology of Cirrhosis</td>
<td>1363</td>
</tr>
<tr>
<td>Clinical Manifestations of Cirrhosis</td>
<td>1364</td>
</tr>
<tr>
<td>Laboratory Findings Associated With Cirrhosis</td>
<td>1364</td>
</tr>
<tr>
<td>Liver Biopsy</td>
<td>1364</td>
</tr>
<tr>
<td>Hepatic Reserve and Assessment of Surgical Risk in the Cirrhotic Patient</td>
<td>1365</td>
</tr>
<tr>
<td>Child-Turcotte-Pugh Score</td>
<td>1365</td>
</tr>
<tr>
<td>Model for End-Stage Liver Disease Scoring System</td>
<td>1365</td>
</tr>
<tr>
<td>Portal Hypertension</td>
<td>1365</td>
</tr>
<tr>
<td>Imaging of the Portal Venous System and Measurement of Portal Venous Pressure</td>
<td>1366</td>
</tr>
<tr>
<td>Etiology and Clinical Features of Portal Hypertension</td>
<td>1366</td>
</tr>
<tr>
<td>Management of Gastroesophageal Varices</td>
<td>1367</td>
</tr>
<tr>
<td>Prevention of Variceal Bleeding</td>
<td>1367</td>
</tr>
<tr>
<td>Management of Acute Variceal Hemorrhage</td>
<td>1367</td>
</tr>
<tr>
<td>Luminal Tamponade</td>
<td>1367</td>
</tr>
<tr>
<td>Transjugular Intrahepatic Portosystemic Shunt</td>
<td>1367</td>
</tr>
<tr>
<td>Balloon-Occluded Retrograde Transvenous Obliteration</td>
<td>1367</td>
</tr>
<tr>
<td>Surgical Shunting</td>
<td>1367</td>
</tr>
<tr>
<td>Nonshunt Surgical Management of Refractory Variceal Bleeding</td>
<td>1368</td>
</tr>
<tr>
<td>Hepatic Transplantation</td>
<td>1368</td>
</tr>
<tr>
<td>Budd-Chiari Syndrome</td>
<td>1368</td>
</tr>
<tr>
<td>Infections of the Liver</td>
<td>1369</td>
</tr>
<tr>
<td>Pyogenic Liver Abscesses</td>
<td>1369</td>
</tr>
<tr>
<td>Amebic Abscesses</td>
<td>1369</td>
</tr>
<tr>
<td>Hydatid Disease</td>
<td>1370</td>
</tr>
<tr>
<td>Ascariasis</td>
<td>1371</td>
</tr>
<tr>
<td>Schistosomiasis</td>
<td>1371</td>
</tr>
<tr>
<td>Viral Hepatitis</td>
<td>1371</td>
</tr>
<tr>
<td>Evaluation of an Incidental Liver Mass</td>
<td>1372</td>
</tr>
<tr>
<td>Hepatic Cysts</td>
<td>1373</td>
</tr>
<tr>
<td>Congenital Cysts</td>
<td>1373</td>
</tr>
<tr>
<td>Biliary Cystadenoma</td>
<td>1373</td>
</tr>
<tr>
<td>Polycystic Liver Disease</td>
<td>1373</td>
</tr>
<tr>
<td>Caroli’s Disease</td>
<td>1374</td>
</tr>
<tr>
<td>Benign Liver Lesions</td>
<td>1374</td>
</tr>
<tr>
<td>Cyst</td>
<td>1374</td>
</tr>
<tr>
<td>Hemangioma</td>
<td>1375</td>
</tr>
<tr>
<td>Adenoma</td>
<td>1375</td>
</tr>
<tr>
<td>Focal Nodular Hyperplasia</td>
<td>1376</td>
</tr>
<tr>
<td>Bile Duct Hamartoma</td>
<td>1376</td>
</tr>
<tr>
<td>Malignant Liver Tumors</td>
<td>1376</td>
</tr>
<tr>
<td>Hepatocellular Carcinoma</td>
<td>1376</td>
</tr>
<tr>
<td>Cholangiocarcinoma</td>
<td>1377</td>
</tr>
<tr>
<td>Gallbladder Cancer</td>
<td>1378</td>
</tr>
<tr>
<td>Metastatic Colorectal Cancer</td>
<td>1378</td>
</tr>
<tr>
<td>Neuroendocrine Tumors</td>
<td>1379</td>
</tr>
<tr>
<td>Other Metastatic Tumors</td>
<td>1379</td>
</tr>
<tr>
<td>Treatment Options for Liver Cancer</td>
<td>1379</td>
</tr>
<tr>
<td>Hepatic Resection</td>
<td>1379</td>
</tr>
<tr>
<td>Liver Transplantation</td>
<td>1380</td>
</tr>
<tr>
<td>Radiofrequency Ablation</td>
<td>1380</td>
</tr>
<tr>
<td>Ethanol Ablation, Cryosurgery, and Microwave Ablation</td>
<td>1380</td>
</tr>
<tr>
<td>Chemoembolization and Hepatic Artery Pump Chemoperfusion</td>
<td>1381</td>
</tr>
<tr>
<td>Yttrium-90 Microspheres</td>
<td>1381</td>
</tr>
<tr>
<td>Stereotactic Radiosurgery and Intensity-Modulated Radiation Therapy</td>
<td>1381</td>
</tr>
<tr>
<td>Downstaging</td>
<td>1381</td>
</tr>
<tr>
<td>Systemic Chemotherapy</td>
<td>1381</td>
</tr>
<tr>
<td>Hepatic Resection Surgical Techniques</td>
<td>1381</td>
</tr>
<tr>
<td>Nomenclature</td>
<td>1381</td>
</tr>
<tr>
<td>Techniques and Devices for Dividing the Hepatic Parenchyma</td>
<td>1382</td>
</tr>
<tr>
<td>Steps in Commonly Performed Hepatic Resections</td>
<td>1383</td>
</tr>
<tr>
<td>Pringle and Ischemic Preconditioning</td>
<td>1385</td>
</tr>
<tr>
<td>Preoperative Portal Vein Embolization</td>
<td>1385</td>
</tr>
<tr>
<td>Staged Hepatectomy, ALPPS, and Repeat Hepatic Resection for Recurrent Liver Cancer</td>
<td>1386</td>
</tr>
<tr>
<td>Laparoscopic Liver Resection</td>
<td>1386</td>
</tr>
</tbody>
</table>
The ancient Greek myth of Prometheus reminds us that the liver is the only organ that regenerates. According to Greek mythology, Zeus was furious with the Titan Prometheus because he gave fire to mortals. In return, Zeus chained Prometheus to Mount Caucasus and sent his giant eagle to eat his liver during the day, only to have it regenerate at night. Although this is folklore, the principles are correct that after hepatic resection, the remnant liver will hypertrophy over weeks to months to regain most of its original liver mass. It is interesting to note that the ancient Greeks seem to have been aware of this fact, because the Greek word for the liver, ἕφαρ, derives from the verb ἥπαομαι, which means “mend” or “repair.” Hence ἕφαρ roughly translates as “repairable.”! The importance of the liver dates back to even biblical times, for the Babylonians (c. 2000 b.c.) considered the liver to be the seat of the soul. There are scattered reports of liver surgery for battlefield injuries, but the first recorded elective hepatic resection was done in 1888 in Germany by Langenbuch. There followed reports of liver resections in the United States (Tiffany, 1890) and Europe (Lucke, 1891), as well as the first recorded elective hepatectomy in 1899.2,3 In 1908, Pringle described in *Annals of Surgery* the “arrest of hepatic hemorrhage due to trauma” by compression of the porta hepatis, a maneuver that now bears his name.4 Possibly due to the potential for massive hemorrhage during liver surgery, very little progress in surgical techniques was recorded for the next half-century. Work by Rex, Cantlie, and others laid the groundwork for experimental and clinical reports in the 1950s by Couinaud, Hjortsjo, Healey, Lortat-Jacob, and Starzl.5,6 These seminal contributions paved the way for the modern era of hepatic resection surgery.

**LIVER ANATOMY**

The liver is the largest organ in the body, weighing approximately 1500 g. It resides in the right upper abdominal cavity beneath the diaphragm and is protected by the rib cage. It is reddish brown and is surrounded by a fibrous sheath known as Glisson’s capsule. The liver is held in place by several ligaments (Fig. 31-1). The round ligament is the remnant of the obliterated umbilical vein and enters the left liver hilum at the front edge of the falciform ligament. The falciform ligament separates the left lateral and left medial segments along the umbilical fissure and anchors the liver to the anterior abdominal wall. Deep in the plane between the caudate lobe and the left lateral segment is the fibrous ligamentum venosum (Arantius’ ligament), which is the obliterated ductus venosus and is covered by the plate of Arantius. The left and right triangular ligaments secure the two sides of the liver to the diaphragm. Extending from the triangular ligaments anteriorly on the liver are the coronary ligaments. The right coronary ligament also extends from the right undersurface of the liver to the peritoneum overlying the right kidney, thereby anchoring the liver to the right retroperitoneum. These ligaments (round, falciform, triangular, and coronary) can be divided in a bloodless plane to fully mobilize the liver to facilitate hepatic resection. Centrally and just to the left of the gallbladder fossa, the liver attaches via the hepatoduodenal and the gastrohepatic ligaments (Fig. 31-2). The hepatoduodenal ligament is known as the porta hepatitis and contains the common bile duct, the hepatic artery, and the portal vein. From the right side and deep (dorsal) to the porta hepatitis is the foramen of Winslow, also known as the epiploic foramen (see Fig. 31-2). This passage connects directly to the lesser sac and allows complete vascular inflow control to the liver when the hepatoduodenal ligament is clamped using the Pringle maneuver.

**Segmental Anatomy**

The liver is grossly separated into the right and left lobes by the plane from the gallbladder fossa to the inferior vena cava (IVC), known as Cantlie’s line.7 The right lobe typically accounts for 60% to 70% of the liver mass, with the left lobe (and caudate lobe) making up the remainder. The caudate lobe lies to the left and anterior of the IVC and contains three subsegments:
the Spiegel lobe, the paracaval portion, and the caudate process.\(^7\) The falciform ligament does not separate the right and left lobes, but rather it divides the left lateral segment from the left medial segment. The left lateral and left medial segments also are referred to as sections as defined in the Brisbane 2000 terminology, which is outlined later in the section titled “Hepatic Resection.” A significant advance in our understanding of liver anatomy came from the cast work studies of the French surgeon and anatomist Couinaud in the early 1950s. Couinaud divided the liver into eight segments, numbering them in a clockwise direction beginning with the caudate lobe as segment I.\(^6\) Segments II and III comprise the left lateral segment, and segment IV is the left medial segment (Fig. 31-3). Thus, the left lobe is made up of the left lateral segment (Couinaud’s segments II and III) and the left medial segment (segment IV). Segment IV can be subdivided into segment IVA and segment IVB. Segment IVA is cephalad and just below the diaphragm, spanning from segment VIII to the falciform ligament adjacent to segment II. Segment IVB is caudad and adjacent to the gallbladder fossa. Many anatomy textbooks also refer to segment IV as the quadrate lobe. Quadrate lobe is an outdated term, and the preferred term is segment IV or left medial segment. Most surgeons still refer to segment I as the caudate lobe, rather than segment I. The right lobe is comprised of segments V, VI, VII, and VIII, with segments V and VIII making up the right anterior lobe and segments VI and VII making up the right posterior lobe.

Figure 31-1. Hepatic ligaments suspending the liver to the diaphragm and anterior abdominal wall.

Figure 31-2. In situ liver hilar anatomy with hepatoduodenal and gastrohepatic ligaments. Foramen of Winslow is depicted.

Figure 31-3. Couinaud’s liver segments (I through VIII) numbered in a clockwise manner. The left lobe includes segments II to IV, the right lobe includes segments V to VIII, and the caudate lobe is segment I. IVC = inferior vena cava.
Additional functional anatomy was highlighted by Bismuth based on the distribution of the hepatic veins. The three hepatic veins run in corresponding scissura (fissures) and divide the liver into four sectors. The right hepatic vein runs along the right scissura and separates the right posterolateral sector from the right anterolateral sector. The main scissura contains the middle hepatic vein and separates the right and left livers. The left scissura contains the course of the left hepatic vein and separates the left posterior and left anterior sectors.

**Hepatic Artery**

The liver has a dual blood supply consisting of the hepatic artery and the portal vein. The hepatic artery delivers approximately 25% of the blood supply, and the portal vein approximately 75%. The hepatic artery arises from the celiac axis (trunk), which gives off the left gastric, splenic, and common hepatic arteries (Fig. 31-4). The common hepatic artery then divides into the gastroduodenal artery and the hepatic artery proper. The right gastric artery typically originates off of the hepatic artery proper, but this is variable. The hepatic artery proper divides into the right and left hepatic arteries. This “classic” or standard arterial anatomy is present in only approximately 76% of cases, with the remaining 24% having variable anatomy. It is critical to understand the arterial (and biliary) anatomic variants to avoid surgical complications when operating on the liver, gallbladder, pancreas, or adjacent organs.

The most common hepatic arterial variants are shown in Fig. 31-5. Approximately 10% to 15% of the time there is a replaced or accessory right hepatic artery arising from the superior mesenteric artery (SMA). When there is a replacement or accessory right hepatic artery, it travels posterior to the portal vein and then takes up a right lateral position before diving into the liver parenchyma. This can be recognized visually on a preoperative computed tomography (CT) or magnetic resonance imaging (MRI) scan and confirmed by palpation in the hilum where a separate right posterior pulsation is felt distinct from that of the hepatic artery proper that lies anteriorly in the hepatoduodenal ligament to the left of the common bile duct. In approximately 3% to 10% of cases, there exists a replacement (or accessory) left hepatic artery coming off of the left gastric artery and running obliquely in the gastrohepatic ligament anterior to the caudate lobe before entering the hilar plate at the base of the umbilical fissure. Other less common variants (approximately 1–2% each) are the presence of both replaced right and replaced left hepatic arteries, as well as a completely replaced common hepatic artery coming off the SMA (see Fig. 31-5). Although not well demonstrated in the illustration, the clue for a completely replaced common hepatic artery coming off the SMA is the presence of a strong arterial pulsation to the right of and posterior to the common bile duct, rather than the left side and anterior, in the porta hepatis. Another important point is that the right hepatic artery passes deep and posterior to the common bile duct approximately 88% of the time but crosses anterior to the common bile duct in approximately 12% of cases. The cystic artery feeding the gallbladder usually arises from the right hepatic artery in Calot’s triangle.

**Portal Vein**

The portal vein is formed by the confluence of the splenic vein and the superior mesenteric vein. The inferior mesenteric vein usually drains into the splenic vein upstream from the confluence (Fig. 31-6). The main portal vein traverses the porta hepatitis before dividing into the left and right portal vein branches. The left portal vein typically branches from the main portal vein outside of the liver with a sharp bend to the left and consists of the transverse portion followed by a 90° turn at the base of the umbilical fissure to become the umbilical portion before entering the liver parenchyma (Fig. 31-7). The left portal vein then divides to give off the segment II and III branches to the left lateral segment, as well as the segment IV branches that supply the left medial segment. The left portal vein also provides the dominant inflow branch to the caudate lobe (although branches can arise from the main and right portal veins also), usually close to the bend between the transverse and umbilical portions. The division of the right portal vein is usually higher in the hilum and may be close to (or inside) the liver parenchyma at the hilar plate. Twenty percent to 35% of individuals have aberrant portal venous anatomy, with portal vein trifurcation or an aberrant branch from the left portal vein supplying the right anterior lobe being the most frequent.

The portal vein drains the splanchnic blood from the stomach, pancreas, spleen, small intestine, and majority of the colon to the liver before returning to the systemic circulation. The portal vein pressure in an individual with normal physiology is low at 3 to 5 mmHg. The portal vein is valveless, however, and in the setting of portal hypertension, the pressure can be...
quite high (20 to 30 mmHg). This results in decompression of the systemic circulation through portocaval anastomoses, most commonly via the coronary (left gastric) vein, which produces esophageal and gastric varices with a propensity for major hemorrhage. Another branch of the main portal vein is the superior pancreaticoduodenal vein (which comes off low in an anterior lateral position and is divided during pancreaticoduodenectomy). Closer to the liver, the main portal vein typically gives off a short branch (posterior lateral) to the caudate process on the right side. It is important to identify this branch and ligate it during hilar dissection for anatomic right hemihepatectomy to avoid avulsion.

**Hepatic Veins and Inferior Vena Cava**

There are three hepatic veins (right, middle, and left) that pass obliquely through the liver to drain the blood to the suprahepatic IVC and eventually the right atrium (Fig. 31-8). The right hepatic vein drains segments V through VIII; the middle hepatic vein drains segment IV as well as segments V and VIII; and the left hepatic vein drains segments II and III. The caudate lobe

---

**Figure 31-5.** Common hepatic artery anatomic variants. SMA = superior mesenteric artery.

**Figure 31-6.** Portal vein anatomy. The portal vein is formed by the confluence of the splenic and superior mesenteric veins. The inferior mesenteric vein drains into the splenic vein. The coronary (left gastric) vein drains into the portal vein in the vicinity of the confluence. v. = vein.

**Figure 31-7.** Anatomy of the left portal vein (LPV). Note the transverse and umbilical portions of the LPV.
is unique because its venous drainage feeds directly into the IVC. In addition, the liver usually has a few small, variable short hepatic veins that directly enter the IVC from the undersurface of the liver. The left and middle hepatic veins form a common trunk approximately 95% of the time before entering the IVC, whereas the right hepatic vein inserts separately (in an oblique orientation) into the IVC. There is a large inferior accessory right hepatic vein in 15% to 20% of cases that runs in the hepaticocaval ligament. This can be a source of torrential bleeding if control of it is lost during right hepatectomy. The hepatic vein branches bisect the portal branches inside the liver parenchyma (i.e., the right hepatic vein runs between the right anterior and posterior portal veins; the middle hepatic vein passes between the right anterior and left portal vein; and the left hepatic vein crosses between the segment II and III branches of the left portal vein).

Bile Duct and Hepatic Ducts
Within the hepatoduodenal ligament, the common bile duct lies anteriorly and to the right. It gives off the cystic duct to the gallbladder and becomes the common hepatic duct before dividing into the right and left hepatic ducts. In general, the hepatic ducts follow the arterial branching pattern inside the liver. The right anterior hepatic duct usually enters the liver above the hilar plate, whereas the right posterior duct dives behind the right portal vein and can be found on the surface of the caudate process before entering the liver. The left hepatic duct typically has a longer extrahepatic course before giving off segmental branches behind the left portal vein at the base of the umbilical fissure. Considerable variation exists, and in 30% to 40% of cases, there is a nonstandard hepatic duct confluence with accessory or aberrant ducts (Fig. 31-9). The cystic duct itself also has a variable pattern of drainage into the common bile duct. This can lead to potential injury or postoperative bile leakage during cholecystectomy or hepatic resection, and the surgeon needs to expect these variants. The gallbladder sits adherent to hepatic segments IVB (left lobe) and V (right lobe).

Neural Innervation and Lymphatic Drainage
The parasympathetic innervation of the liver comes from the left vagus, which gives off the anterior hepatic branch, and the right vagus, which gives off the posterior hepatic branch. The sympathetic innervation involves the greater thoracic splanchnic nerves and the celiac ganglia, although the function of these nerves is poorly understood. The denervated liver after hepatic transplantation seems to function with normal capacity. A common source of referred pain to the right shoulder and scapula as well as the right side or back is the right phrenic

Figure 31-8. Confluence of the three hepatic veins (HVs) and the inferior vena cava (IVC). Note that the middle and left HVs drain into a common trunk before entering the IVC. a. = artery; v. = vein. (Adapted with permission from Cameron JL: Atlas of Surgery. Vol. I, Gallbladder and Biliary Tract, the Liver, Portasystemic Shunts, the Pancreas. Toronto: BC Decker; 1990.)
nerve, which is stimulated by tumors that stretch Glisson’s capsule or by diaphragmatic irritation.

Lymph is produced within the liver and drains via the perisinusoidal space of Disse and periportal clefts of Mall to larger lymphatics that drain to the hilar cystic duct lymph node (Calot’s triangle node), as well as the common bile duct, hepatic artery, and retropancreatic and celiac lymph nodes. This is particularly important for resection of hilar cholangiocarcinoma, which has a high incidence of lymph node metastases. The hepatic lymph also drains cephalad to the cardiophrenic lymph nodes, and the latter can be pathologically identified on a staging CT or MRI scan.

LIVER PHYSIOLOGY

The liver is the largest gland in the body and has an extraordinary spectrum of functions. These include processes such as storage, metabolism, production, and secretion. One crucial role is the processing of absorbed nutrients through the metabolism of glucose, lipids, and proteins. The liver maintains glucose concentrations in a normal range over both short and long periods by performing several important roles in carbohydrate metabolism. In the fasting state, the liver ensures a sufficient supply of glucose to the central nervous system. The liver can produce glucose by breaking down glycogen through glycogenolysis and by de novo synthesis of glucose through gluconeogenesis from noncarbohydrate precursors such as lactate, amino acids, and glycerol. In the postprandial state, excess circulating glucose is removed by glycogen synthesis or glycolysis and lipogenesis. The liver also plays a central role in lipid metabolism through the formation of bile and the production of cholesterol and fatty acids. Protein metabolism occurs in the liver through amino acid deamination, resulting in the production of ammonia as well as the production of a variety of amino acids. In addition to

Figure 31-9. Main variations of hepatic duct confluence. As described by Couinaud in 1957, the bifurcation of the hepatic ducts has a variable pattern in approximately 40% of cases. CHD = common hepatic duct; lh = left hepatic; R = right; ra = right anterior; rp = right posterior. (Reproduced with permission from Blumgart LH, Fong Y: Surgery of the Liver and Biliary Tract, 3rd ed, Vol. I. London: Elsevier; 2000.)
metabolism, the liver also is responsible for the synthesis of most circulating plasma proteins. Among these proteins are albumin, factors of the coagulation and fibrinolytic systems, and compounds of the complement cascade. Furthermore, the detoxification of many substances through drug metabolism occurs in the liver, as do immunologic responses through the many immune cells found in its reticuloendothelial system.

**Bilirubin Metabolism**

Bilirubin is the breakdown product of normal heme catabolism. Bilirubin is bound to albumin in the circulation and sent to the liver. In the liver, it is conjugated to glucuronic acid to form bilirubin digluconuride in a reaction catalyzed by the enzyme glucuronyl transferase, making it water soluble. This glucuronide is then excreted into the bile canaliculi. A small amount dissolves in the blood and is then excreted in the urine. The majority of conjugated bilirubin is excreted in the intestine as waste because the intestinal mucosa is relatively impermeable to conjugated bilirubin. However, it is permeable to unconjugated bilirubin and urobilinogens, a series of bilirubin derivatives formed by the action of bacteria. Thus, some of the bilirubin and urobilinogens are reabsorbed in the portal circulation; they are again excreted by the liver or enter the circulation and are excreted in the urine.

**Formation of Bile**

Bile is a complex fluid containing organic and inorganic substances dissolved in an alkaline solution that flows from the liver through the biliary system and into the small intestine. The main components of bile are water, electrolytes, and a variety of organic molecules including bile pigments, bile salts, phospholipids (e.g., lecithin), and cholesterol. The two fundamental roles of bile are to aid in the digestion and absorption of lipids and lipid-soluble vitamins and to eliminate waste products (bilirubin and cholesterol) through secretion into bile and elimination in feces. Bile is produced by hepatocytes and secreted through the biliary system. In between meals, bile is stored in the gallbladder and concentrated through the absorption of water and electrolytes. Upon entry of food into the duodenum, bile is released from the gallbladder to aid in digestion. The human liver can produce about 1 L of bile daily.

Bile salts, in conjunction with phospholipids, are responsible for the digestion and absorption of lipids in the small intestine. Bile salts are sodium and potassium salts of bile acids conjugated to amino acids. The bile acids are derivatives of cholesterol synthesized in hepatocytes. Cholesterol, ingested from the diet or derived from hepatic synthesis, is converted into the bile acids cholic acid and chenodeoxycholic acid. These bile acids are conjugated to either glycine or taurine before secretion into the biliary system. Bacteria in the intestine can remove glycine and taurine from bile salts. They can also convert some of the primary bile acids into secondary bile acids by removing a hydroxyl group, producing deoxycholic acid from cholic acid and lithocholic acid from chenodeoxycholic acid.

Bile salts secreted into the intestine are efficiently reabsorbed and reused. Approximately 90% to 95% of the bile salts are absorbed from the small intestine at the terminal ileum. The remaining 5% to 10% enter the colon and are converted to the secondary salts, deoxycholic acid and lithocholic acid. The mixture of primary and secondary bile salts and bile acids is absorbed primarily by active transport in the terminal ileum. The absorbed bile salts are transported back to the liver in the portal vein and reexcreted in the bile. Those lost in the stool are replaced by synthesis in the liver. The continuous process of secretion of bile salts in the bile, their passage through the intestine, and their subsequent return to the liver is termed the enterohepatic circulation.

**Drug Metabolism**

The liver plays an important role in providing mechanisms for ridding the body of foreign molecules (xenobiotics) that are absorbed from the environment. In most cases, a drug is relatively lipophilic to ensure good absorption. The liver participates in the elimination of these lipid-soluble drugs by transforming them into more readily excreted hydrophilic products. There are two main reactions important for drug metabolism. Phase 1 reactions include oxidation, reduction, and hydrolysis of molecules. These result in metabolites that are more hydrophilic than the original chemicals. The cytochrome P450 system is a family of hemoproteins important for oxidative reactions involving drugs and toxic substances. Phase 2 reactions, also known as conjugation reactions, are synthetic reactions that involve addition of subgroups to the drug molecule. These subgroups include glucuronate, acetate, glutathione, glycine, sulfate, and methyl groups. These drug reactions occur mainly in the smooth endoplasmic reticulum of hepatocytes.

Many factors can affect drug metabolism in the liver. When the rate of metabolism of a drug is increased (i.e., enzyme induction), the duration of the drug action will decrease. However, when the metabolism of a drug is decreased (i.e., enzyme inhibition), the drug will circulate for a longer period of time. It is important to note that some drugs may be converted to active products by metabolism in the liver. An example is acetaminophen when taken in larger doses. Normally, acetaminophen is conjugated by the liver to harmless glucuronide and sulfate metabolites that are water soluble and eliminated in the urine. During an overdose, the normal metabolic pathways are overwhelmed, and some of the drug is converted to a reactive and toxic intermediate by the cytochrome P450 system. Glutathione normally reacts with this intermediate, leading to the production and subsequent excretion of a harmless product. However, as glutathione stores are diminished, the reactive intermediate cannot be detoxified and it combines with lipid membranes of hepatocytes, which results in cellular necrosis. Thus, treatment of acetaminophen overdoses consists of replenishing glutathione stores by supplementing with sulfhydryl compounds such as acetylcysteine.

**Liver Function Tests**

Liver function tests is a term frequently used to refer to measurement of the levels of a group of serum markers for evaluation of liver dysfunction. Most commonly, levels of aspartate transaminase (AST), alanine transaminase (ALT), alkaline phosphatase (AP), γ-glutamyltranspeptidase (GGT), and bilirubin are included in this panel. This term is a misnomer, however, because most of these tests measure not liver function but rather cell damage. More accurate measurement of the liver’s synthetic function is provided by serum albumin levels and prothrombin time (PT). Although measuring liver enzyme levels is important in the assessment of a patient’s liver disease, these test results can be nonspecific. Thus, evaluation of patients with suspected liver disease should always involve careful interpretation of abnormalities in these liver test results in the context of a thorough history and physical examination. The approach to
evaluating abnormal laboratory values also can be simplified by 
categorizing the type of abnormality that predominates (hepa-
tocellular damage, abnormal synthetic function, or cholestasis).

**Hepatocellular Injury**

Hepatocellular injury of the liver is usually indicated by abnor-
malities in levels of the liver aminotransferases AST and ALT. 
These enzymes participate in gluconeogenesis by catalyzing 
the transfer of amino groups from aspartic acid or alanine 
to ketoglutaric acid to produce oxaloacetic acid and pyruvic 
acid, respectively (these enzymes are also referred to as *serum 
glutamic-oxaloacetic transaminase [SGOT]* and *serum glutamic-
pyruvic transaminase [SGPT]*). AST is found in liver, cardiac 
muscle, skeletal muscle, kidney, brain, pancreas, lungs, and red 
blood cells and thus is less specific for disorders of the liver. 
ALT is predominately found in the liver and thus is more specific 
for liver disease. Hepatocellular injury is the trigger for release of 
these enzymes into the circulation. Common causes of elevated 
aminotransferase levels include viral hepatitis, alcohol abuse, 
medications, genetic disorders (Wilson’s disease, hemochroma-
tosis, α1-antitrypsin deficiency), and autoimmune diseases.

The extent of serum aminotransferase elevations can sug-
gest certain etiologies of the liver injury. However, the levels of 
the enzymes in these tests correlate poorly with the severity of 
hepatocellular necrosis because they may not be significantly 
elevated in conditions of hepatic fibrosis or cirrhosis. In alco-
holic liver disease, an AST to ALT ratio of >2:1 is common. 
Mild elevations of transaminase levels can be found in nonal-
coholic fatty liver disease, chronic viral infection, or medica-
tion-induced injury. Moderate increases in the levels of these 
enzymes are common in acute viral hepatitis. In conditions of 
ischemic insults, toxin ingestions (i.e., acetaminophen), and ful-
miant hepatitis, AST and ALT levels can be elevated to the 
thousands.

**Abnormal Synthetic Function**

Albumin synthesis is an important function of the liver and thus 
can be measured to evaluate the liver’s synthetic function. The 
liver produces approximately 10 g of albumin per day. How-
ever, albumin levels are dependent on a number of factors such 
as nutritional status, renal dysfunction, protein-losing enteropa-
thies, and hormonal disturbances. In addition, level of albumin 
is not a marker of acute hepatic dysfunction due to albumin’s 
long half-life of 15 to 20 days.

Most clotting factors (except factor VIII) are synthesized 
exclusively in the liver, and thus their levels can also be used as a 
measure of hepatic synthetic function. Measurements of the pro-
thrombin time (PT) and international normalized ratio (INR) 
are some of the best tests of hepatic synthetic function. PT measures 
the rate of conversion of prothrombin to thrombin. To standard-
ize the reporting of PT and avoid interlaboratory variability, the 
INR was developed. The INR is the ratio of the patient’s PT 
to the mean control PT. Because vitamin K is involved in the 
γ-carboxylation of factors used to measure PT (factors II, VII, 
IX, and X), values also may be prolonged in other conditions 
such as vitamin K deficiency and warfarin therapy.

**Cholestasis**

Cholestasis is a condition in which bile flow from the liver to the 
duodenum is impaired. Disturbances in bile flow may be due to 
intrahepatic causes (hepatocellular dysfunction) or extrahepatic 
causes (biliary tree obstruction). Cholestasis often results in the 
release of certain enzymes and thus can be detected by measur-
ing the serum levels of bilirubin, AP, and GGT. Bilirubin is a 
breakdown product of hemoglobin metabolism. Unconjugated 
bilirubin is insoluble and thus is transported to the liver bound 
to albumin. In the liver, it is conjugated to allow excretion in 
bile. Measured total bilirubin levels can be normal or high in 
patients with significant liver disease because of the liver’s abil-
ity to conjugate significant amounts of bilirubin. Thus, to help 
aid in the diagnosis of hyperbilirubinemia, fractionation of total 
bilirubin is usually performed to distinguish between conjugated 
(direct) and unconjugated (indirect) bilirubin. *Indirect bilirubin* 
is a term frequently used to refer to unconjugated bilirubin in 
the circulation because the addition of another chemical is ne-
cessary to differentiate this fraction from the whole. Normally, 
>90% of serum bilirubin is unconjugated. The testing process 
for conjugated bilirubin, in contrast, is direct without the addi-
tion of other agents. The direct bilirubin test measures the lev-
els of conjugated bilirubin and δ bilirubin (conjugated bilirubin 
bound to albumin).

The patterns of elevation of the different fractions of bili-
rubin provide important diagnostic clues as to the cause of chole-
lostasis. In general, an elevated indirect bilirubin level suggests 
intrahepatic cholestasis, and an elevated direct bilirubin level 
suggests extrahepatic obstruction. Mechanisms that can result 
in increases in unconjugated bilirubin levels include increased 
bilirubin production (hemolytic disorders and resorption of 
hematomas) or defects (inherited or acquired) in hepatic uptake 
or conjugation. The rate-limiting step in bilirubin metabolism 
is the excretion of bilirubin from hepatocytes, so conjugated 
hyperbilirubinemia can be seen in inherited or acquired dis-
orders of intrahepatic excretion or extrahepatic obstruction. 
Conjugated bilirubin that cannot be excreted accumulates in 
hepatocytes, which results in its secretion into the circulation. 
Because conjugated bilirubin is water soluble, it can be found 
in the urine of patients with jaundice.

AP is an enzyme with a wide tissue distribution but is 
found primarily in the liver and bones. In the liver, it is expressed 
by the bile duct epithelium. In conditions of biliary obstruction, 
levels rise as a result of increased synthesis and release into 
the serum. Because the half-life of serum AP is approximately 
7 days, it may take several days for levels to normalize even 
after resolution of the biliary obstruction.

GGT is another enzyme found in hepatocytes and released 
from the bile duct epithelium. Elevation of GGT is an early 
marker and also a sensitive test for hepatobiliary disease. Like 
AP elevation, however, it is nonspecific and can be produced by 
a variety of disorders in the absence of liver disease. Increased 
levels of GGT can be induced by certain medications, alcohol 
abuse, pancreatic disease, myocardial infarction, renal failure, 
and obstructive pulmonary disease. For this reason, elevated 
GGT levels are often interpreted in conjunction with other 
enzyme abnormalities. For example, a raised GGT level with 
increased AP level supports a liver source.

**Jaundice**

Jaundice refers to the yellowish staining of the skin, sclera, and 
mucous membranes with the pigment bilirubin. Hyperbilirubi-
nemia usually is detectable as jaundice when blood levels rise 
above 2.5 to 3 mg/dL. Jaundice can be caused by a wide range 
of benign and malignant disorders. However, when present, it 
may indicate a serious condition, and thus knowledge of the 
differential diagnosis of jaundice and a systematic approach to
the workup of the patient is necessary. Workup of a patient with jaundice is simplified by organizing the possible causes of the disorder into groups based on the location of bilirubin metabolism. As mentioned previously, bilirubin metabolism can take place in three phases: prehepatic, intrahepatic, and posthepatic. The prehepatic phase includes the production of bilirubin from the breakdown of heme products and its transport to the liver. The majority of the heme results from red blood cell metabolism and the rest from other heme-containing organic compounds such as myoglobin and cytochromes. In the liver, the insoluble unconjugated bilirubin is then conjugated to glucuronic acid to allow for solubility in bile and excretion. The posthepatic phase of bilirubin metabolism consists of excretion of soluble bilirubin through the biliary system into the duodenum. Dysfunction in any of these phases can lead to jaundice.\(^\text{10}\)

**Prehepatic.** Jaundice as a result of elevated levels of unconjugated bilirubin occurs from faulty prehepatic metabolism and usually arises from conditions that interfere with proper conjugation of bilirubin in the hepatocyte. Insufficient conjugation is often seen in processes that result in excessive heme metabolism. Subsequently, the conjugation system is overwhelmed, which results in unconjugated hyperbilirubinemia. Causes of hemolysis include inherited and acquired hemolytic anemias. Inherited hemolytic anemias include genetic disorders of the red blood cell membrane (hereditary spherocytosis and elliptocytosis), enzyme defects (glucose-6-phosphate dehydrogenase deficiency), and defects in hemoglobin structure (sickle cell anemia and thalassemias). Hemolytic anemias can also be acquired, and these can be further divided into those with immune-mediated and those with non-immune-mediated causes. Immune-mediated hemolytic anemias result in a positive finding on a direct Coombs test and have a variety of autoimmune and drug-induced causes. In contrast, direct Coombs test results are negative in nonimmune hemolytic anemias. The causes in this latter category are varied and include drugs and toxins that directly damage red blood cells, mechanical trauma (heart valves), microangiopathy, and infections. Prehepatic dysfunction of bilirubin metabolism also can result from failure in the transport of unconjugated bilirubin to the liver by albumin in any condition that leads to plasma protein loss. A poor nutritional state or excess protein loss as seen in burn patients can lead to elevated levels of unconjugated bilirubin in the circulation and jaundice.

**Intrahepatic.** Intrahepatic causes of jaundice involve the intracellular mechanisms for conjugation and excretion of bile from the hepatocyte. The enzymatic processes in hepatocytes can be affected by any condition that impairs hepatic blood flow and subsequent function of the liver (ischemic or hypoxic events). Furthermore, there are multiple inherited disorders of enzyme metabolism that can result in either unconjugated or conjugated hyperbilirubinemia. Gilbert’s syndrome is a genetic variant characterized by diminished activity of the enzyme glucuronyltransferase, which results in decreased conjugation of bilirubin to glucuronide. It is a benign condition that affects approximately 4% to 7% of the population. Typically, the disease results in transient mild increases in unconjugated bilirubin levels and jaundice during episodes of fasting, stress, or illness. These episodes are self-limited and usually do not require further treatment. Another inherited disorder of bilirubin conjugation is Crigler-Najjar syndrome. It is a rare disease found in neonates and can result in neurotoxic sequelae from bilirubin encephalopathy.

In addition to defects in conjugation, disorders in bilirubin excretion in hepatocytes can also lead to jaundice. Rotor’s syndrome and Dubin-Johnson syndrome are two uncommon genetic disorders that disrupt secretion of conjugated bilirubin from the hepatocyte into the bile and result in conjugated hyperbilirubinemia. There are also multiple acquired conditions that result in inflammation and intrahepatic cholestasis by affecting hepatocyte mechanisms for conjugation and excretion of bile. Viruses, alcohol abuse, sepsis, and autoimmune disorders all can result in inflammation in the liver with subsequent disruption of bilirubin transport in the liver. In addition, jaundice can also occur from the cytotoxic effects of many medications, including acetaminophen, oral contraceptives, and anabolic steroids.

**Posthepatic.** Posthepatic causes of jaundice are usually the result of intrinsic or extrinsic obstruction of the biliary duct system that prevents the flow of bile into the duodenum. There is a wide spectrum of pathologies that may present with obstructive jaundice. Intrinsic obstruction can occur from biliary diseases, including cholelithiasis, cholecystocholangitis, benign and malignant biliary strictures, cholangiocarcinoma, cholangitis, and disorders of the papilla of Vater. Extrinsic compression of the biliary tree is commonly due to pancreatic disorders. Patients with pancreatitis, pseudocysts, and malignancies can present with jaundice due to external compression of the biliary system. Finally, with the growing armamentarium of endoscopic tools and minimally invasive surgical approaches, surgical complications are becoming more frequent causes of extrahepatic cholestasis. Misadventures with surgical clips, retained stones, and inadvertent ischemic insults to the biliary system can result in obstructive jaundice recognized at any time from immediately postoperatively to many years later.

**MOLECULAR SIGNALING PATHWAYS IN THE LIVER**

**Acute Phase Reaction**

The liver is the site of synthesis of acute phase proteins that consists of a group of plasma proteins that are rapidly released in response to inflammatory conditions elsewhere in the body. The synthesis of these proteins in the liver is mediated by a number of inflammatory mediators. Cytokines such as tumor necrosis factor alpha (TNF-\(\alpha\)), interferon-\(\gamma\) (IFN-\(\gamma\)), interleukin-1 (IL-1), interleukin-6 (IL-6), and interleukin-8 (IL-8) are released by inflammatory cells into the circulation at sites of injury and modulate the acute phase response. In response to these cytokines, the liver increases synthesis and release of a wide variety of proteins including ceruloplasmin, complement factors, C-reactive protein (CRP), \(\alpha\)-dimeric protein, alpha 1-antitrypsin, and serum amyloid A. There are proteins, such as serum albumin and transferrin, that also decrease (negative acute phase proteins) in response to inflammation.

The acute phase response of the liver can be initiated in response to infection, trauma, or malignancy. The purpose of the release of these proteins from the liver is to contain infectious processes, prevent further tissue damage and to begin reparative and regeneration processes to restore body homeostasis. For example, products of the complement pathways can attach to microbes to allow for phagocytosis and act as chemoattractants to the areas of inflammation. CRP is an important acute phase protein that is also involved in the clearance of microorganisms by binding to their membranes and functioning as an opsonin to
facilitate phagocytosis. Other proteins such as α1-antitrypsin are protease inhibitors and restrict the protease activity of enzymes of inflammatory cells. Thus, the secretion of acute phase proteins from the liver during the acute phase response is an early defense measure against harmful stimuli prior to the full activation of the immune response.11

**LPS Signaling**

The liver is a complex organ with an important function in immune surveillance and clearance of bacteria and their products. This function is facilitated by the fact that the liver receives all the drainage of the gastrointestinal tract via the portal blood flow, making it the last barrier preventing bacteria and their toxins from reaching the systemic circulation. The importance of preventing bacteria and their products from reaching the systemic blood stream is evident in patients who are infected with gram-negative bacteria. Gram-negative bacteria infection produces an acute inflammatory reaction that can lead to septic shock and multiple organ failure. The complications of gram-negative sepsis are initiated by endotoxin (lipopolysaccharide [LPS]). LPS is a glycolipid constituent of gram-negative bacteria outer membranes composed of a hydrophilic polysaccharide portion and a hydrophobic domain called lipid A. The lipid A structure is the LPS component responsible for the biological effects of LPS. Mere nanogram amounts of LPS injected into humans are needed to result in the manifestations of septic shock. The profound effects of LPS are caused not only by the direct effect of LPS itself but also by the activation of LPS-sensitive cells, resulting in the excessive release of cytokines and other inflammatory mediators.

Since sepsis from gram-negative bacterial infection continues to be a major cause of morbidity and mortality, significant efforts have been made to identify the molecules involved in LPS-binding and signaling. Lipopolysaccharide binding protein (LBP), CD14, MD2, and Toll-like receptors have all been identified as important mediators in the pathway of LPS stimulation. LBP is an acute-phase protein synthesized by hepatocytes that binds the lipid A moiety of LPS and forms a soluble LBP-LPS complex. This LBP-LPS complex then interacts with CD14, a receptor identified as important in LPS recognition, resulting in the release of inflammatory cytokines and mediators. Studies have shown that although LBP is important, it is not required for LPS to interact with CD14; however, its presence markedly decreases the concentration of LPS necessary for cellular activation. This may be important especially at the low concentrations of LPS found under physiological conditions. CD14 exists in two forms, membrane (mCD14) and soluble (sCD14) form. The interaction of LPS with membrane CD14 or soluble CD14 is important in host clearance of LPS. This interaction is also responsible for the toxic effects of LPS seen in the liver and systemic circulation after the release of inflammatory cytokines and mediators. While membrane CD14 is a membrane protein found on the surface of myeloid lineage and mediates the activation of these cells by LPS, soluble CD14 is found in the serum and enables responses to LPS by cells that do not express CD14. In addition to its important role in the release of LBP as an acute phase reactant during LPS-mediated inflammatory insults, the liver is also one of the major sources of soluble CD14 into the circulation.

The binding of the LBP-LPS complex to CD14 is not enough to transduce an intracellular LPS signal. Membrane CD14 is a glycosyl phosphatidylinositol-anchored protein without a membrane-spanning domain. Thus, signaling further downstream of LPS requires additional elements. In studies using chemically modified, radioiodinated LPS capable of cross-linking to nearby proteins, LPS has been shown to cross-link specifically to two other molecules, TLR4 and MD-2. TLR4 is a member of the family of proteins called Toll-like receptors and has been identified as the transmembrane coreceptor to CD14. TLR4 was originally identified as the molecular sensor for bacterial LPS when studies demonstrated that mutations in the tlr4 gene were responsible for defective LPS signaling in mutant mice. Thus, initiation of LPS signal cascade requires the interaction of LPS directly with the heteromeric receptor complex of CD14, TLR4, and MD-2. Activation of this complex senses the presence of bacterial LPS at the cell surface and then transmits a signal into the cytoplasm through two distinct pathways. One pathway is dependent upon an adaptor known as myeloid differentiation factor 88 (MyD88). The other pathway is MyD88-independent and relies on an adaptor known as Toll/IL-1 receptor domain-containing adaptor-inducing IFNβ (TRIF).

The liver is the main organ involved in the clearance of LPS from the bloodstream and so plays a critical role in the identification and processing of LPS. Kupffer cells are the resident macrophages of the liver and have been shown to participate in LPS clearance. Studies have demonstrated that the majority of radiolabelled LPS injected intravenously is quickly cleared from the circulation and found in the liver, primarily localized to the Kupffer cells. Kupffer cells also contribute to the inflammatory cascade by producing cytokines in response to LPS. Interestingly, hepatocytes, the parenchymal cells of the liver, also have all the components required for LPS recognition and signaling and can participate in the response to LPS and process LPS for clearance.

Although the liver is essential in the host response to gram-negative bacteria infection by contributing to LPS clearance and to the LPS-induced inflammatory reaction, evidence reveals that LPS may actually have a reciprocal role in the pathogenesis of liver disorders. The relationship between LPS and liver disease is not a novel concept. Early studies have shown the correlation between the presence or absence of gut-derived LPS and the development of liver injury. Attempts to eliminate gut-derived LPS have had protective effects in various animal models of liver injury, including alcohol-induced liver disease. Other studies have shown the synergism between LPS and hepatotoxins in worsening liver injury.

In summary, the liver is essential in the clearance of LPS, but it can also contribute to the negative systemic effects seen in gram-negative sepsis by excessive activation of the LPS signaling pathway. In addition, there is evidence that this signaling pathway may participate in the pathogenesis of a variety of liver disease. An understanding and characterization of the LPS pathway within the liver is an important step to understanding the molecular basis for the lethal effect of LPS during sepsis and liver disorders.12,13

**Nitric Oxide**

Nitric oxide (NO) is a diffusible, free radical gas that was first identified in 1980 as endothelium-derived relaxing factor. Its physiologic and pathophysiologic importance was first discovered in the cardiovascular system with its vital role as a vasodilator. However, its mediation in a variety of other diverse biological activities has since been discovered. In the liver, the
The role of NO in inflammatory states of the liver is complex and is at times conflicting. Under physiologic conditions, NO is important in maintaining hepatic perfusion. However, in inflammatory conditions, such as ischemia/reperfusion (I/R), nitric oxide can play either a protective or harmful role depending on the enzymatic source (inducible versus endothelial nitric oxide synthase) and the type of ischemia reperfusion (cold vs. warm). It appears that the low level of constitutively expressed eNOS-derived NO is primarily beneficial in models of I/R injury with vasodilatation and subsequent improvement in hepatic microcirculation as the proposed mechanism of protection. Interestingly, activation of iNOS in similar models suggests a potentially harmful role for iNOS. Nitric oxide, through its reaction with reactive nitrogen and oxygen intermediates generated in the course of reperfusion injury, can contribute to much of the hepatocellular damage depending on the intracellular ratio of these intermediates to nitric oxide. The production of iNOS and NO are also closely tied to multiple other inflammatory mediators in the liver, and activation of these downstream signals may explain some of the detrimental effects of NO in I/R injury of the liver. Thus, given its diverse biological effects as a signaling molecule, it is not surprising that NO plays both protective and potentially harmful role in the setting of hepatic I/R injury. The final effect of NO varies in different liver diseases and depends on the overall hepatic environment. The potential use of NO pharmacologic manipulation to treat hepatic disease will require careful balance of the risks and benefits of this simple, yet extremely complicated, molecule.14,15

Heme Oxygenase System

The heme oxygenase (HO) is the rate-limiting enzyme in the degradation of heme to yield biliverdin, carbon monoxide (CO), and free iron. The HO system, which is activated in response to multiple cellular stresses, has been shown to be an endogenous cytoprotectant in a variety of inflammatory conditions. There are currently three heme oxygenase isozymes identified. HO-1 is the inducible form of heme oxygenase, while HO-2 and HO-3 are constitutively expressed. The function of HO in heme degradation is essential due to the potentially toxic effects of heme. An excess of heme can cause cellular damage from oxidative stress due to its production of reactive oxygen species. Thus, the HO system is an important defense mechanism against free heme-mediated oxidative stress.

HO-1 has been shown to be induced in a variety of organs during diverse conditions such as hypoxia, endotoxemia, ischemia/reperfusion (I/R), hyperthermia, and radiation. It is thought that HO-1 is involved in maintaining redox homeostasis during cellular stress. In the liver, HO-1 is thought to normally modulate hepatic microvasculature tone through its generation of CO and, like nitric oxide, its activation of guanylyl cyclase. This important role is demonstrated in animal models of portal hypertension where inhibition of HO-1 exacerbates hypertension. Since HO-1 is induced as a protective mechanism in response to various stimuli, targeted induction of HO-1 has been studied as a therapeutic strategy for protection against inflammatory processes. HO-1 overexpression exerts hepatoprotective effects in models of I/R injury, hemorrhagic shock and resuscitation, acetaminophen-induced hepatonecrosis, and sepsis-mediated liver injury.

Although HO-1 has been shown to provide protective effects in a variety of inflammatory states, the specific mechanisms by which HO-1 mediates its protective effects are remains to be fully elucidated. Originally thought to be only potentially toxic waste, the byproducts generated during heme catabolism now appear to play important roles against cellular stress. The well-known hazardous effects of high doses of CO are attributable to its ability to bind hemoglobin and myoglobin, preventing the release of oxygen to tissues. However, only recently have the physiological and beneficial roles of CO been identified. CO is produced in injured tissues via induction of HO-1 and contributes to the attenuation of proinflammatory processes. Similar to NO, CO plays an important role in maintaining the microcirculation through its activation of soluble guanylyl cyclase and subsequent elevation of intracellular cyclic 3’5’-guanosine monophosphate (cGMP). The signaling activities of cGMP lead to smooth muscle relaxation and inhibition and platelet aggregation. In addition, CO has also been shown to inhibit proinflammatory cytokines (TNF-α, IL-1) and chemokines while simultaneously inducing anti-inflammatory cytokines (IL-10). Exogenous low-dose CO has been shown to protect the liver from I/R injury and endotoxemia.

Biliverdin and bilirubin are other metabolites of heme that are also recognized as possible mediators of HO-1’s protective
The cytosolic enzyme biliverdin reductase catalyzes the reduction of biliverdin to bilirubin. Both biliverdin and bilirubin have important endogenous antioxidant properties. Free iron, the third byproduct of heme oxidation is known to be cytotoxic by catalyzing the production of hydroxyl radicals. However, HO-1 induction is associated with increased levels of ferritin, the free iron sequestering protein. Thus, the increase in ferritin with the subsequent decrease in intracellular concentrations of free iron results in a net antioxidant effect. Importantly, both bilirubin and ferritin have been shown to protect against liver injury in a variety of I/R models.

In summary, HO-1 is upregulated and protective in multiple conditions of hepatic stress. Until recently, the degradation products of the HO system were thought to only be potentially toxic waste. It now appears that CO, biliverdin/bilirubin, ferritin are important in the maintenance of cellular redox homeostasis and may play a role in the mechanism of hepatoprotection in disease. Studies involving induction of HO-1 expression and use of its metabolic products hold therapeutic promises for novel protective agents against disorders of hepatic inflammation.\(^\text{16,17}\)

**Toll-Like Receptors**

The liver is a central regulator of the systemic immune response following acute insults to the body. Not only does it play a crucial role in modulating the systemic inflammatory response to infection or injury, it is also subject to injury and dysfunction from these same processes. Recent advances in the study of mechanisms for the activation of the innate immune system have pointed to the Toll-like receptors (TLRs) as a common pathway for immune recognition of microbial invasion and tissue injury. By recognizing either microbial products or endogenous molecules released from damaged sites, the TLR system is capable of alerting the host of danger by activating the innate immune system. Initially, this is manifested by the production of inflammatory mediators and the rapid uptake of invading microbes and their products. When excessive, this inflammatory response can contribute to organ damage and dysfunction.

To date, 13 TLRs have been described in mice, and 11 in humans. TLRs are a family of proteins that are mammalian homologues to the *Drosophila* Toll, a protein that functions in development and immunity. The cytoplasmic portion of Toll-like receptors is similar to that of the IL-1 receptor (IL-1R) family and is called the Toll/IL-1 receptor (TIR) domain. Unlike the IL-1 receptor extracellular portion that consists of an immunoglobulin-like domain, the Toll-like receptors have leucine-rich repeats in their extracellular portion. The TLR receptors have many structural similarities both extracellularly and intracellularly, but they differ from each other in ligand specificities, expression patterns, and with some variability in the signaling pathways they activate.

The TLR receptors were initially identified as components of the innate immune system that acted as a front-line defense mechanism against infections. Their recognition of patterns on pathogens, such as microbial peptidoglycans, lipopolysaccharide, lipoteichoic acids, bacterial DNA, and single-stranded RNA, resulted in the activation of an inflammatory response meant for controlling the invading organisms. In situations of noninfectious inflammation such as seen in trauma, clinicians have long recognized similar activation of the same inflammatory pathways and systemic manifestations. This observation, among others, led to the hypothesis that the immune system is designed to recognize any threats, whether from pathogens or tissue damage, that may lead to disruption of homeostasis. In conditions of sterile inflammation, the activation of immune cells is through the release of endogenous danger molecules, normal cell constituents released by damage or dying cells or components of the extracellular matrix, released by the action of proteases at the site of tissue damage. Recent observations show that both microbial products and endogenous danger molecules can be recognized through the TLR system.

Perhaps more than any of the other TLR family members, TLR4 sits at the interface of microbial and sterile inflammation. Whereas the role of TLR4 in the recognition of lipopolysaccharide (LPS) is well established, only recently has it become apparent that TLR4 also participates in the recognition of endogenous danger molecules. In vivo evidence for TLR4-mediated danger signaling comes from studies of acute tissue injury in hemorrhagic shock, trauma, and I/R models. In each case, TLR4-mutant animals exhibited reduced injury or inflammation compared to wild-type controls. In efforts to identify the ligands responsible for TLR4-dependent signaling in noninfectious insults, multiple molecules have been suggested. These include heat shock proteins, fibrinogen, hyaluronic acid, heparan sulfate, and high mobility group box-1. Although a central role for TLR4 in recognizing tissue injury is building, studies are beginning to suggest that other TLR family members may also participate in the recognition of endogenous molecules released by tissue injury. The very recent realization that certain TLR family members also respond to endogenous molecules released from stressed or damaged tissues points to a molecular basis for a shared mechanism of innate immune activation by infection and injury.\(^\text{18-20}\)

**RADIOLOGIC EVALUATION OF THE LIVER**

**Ultrasound**

Abdominal ultrasound is a commonly applied imaging modality used to evaluate abdominal symptoms. Ultrasound technology is based on the pulse-echo principle. The ultrasound transducer converts electrical energy to high-frequency sound energy that is transmitted into tissue. Although some of the ultrasound waves are transmitted through the tissue, some are reflected back, and the ultrasound image is produced when the ultrasound receiver detects those reflected waves. This real-time gray-scale (B-mode) imaging is augmented by Doppler flow imaging. Doppler ultrasound not only can detect the presence of blood vessels but also can determine the direction and velocity of blood flow. Ultrasonography is a useful initial imaging test of the liver because it is inexpensive, widely available, involves no radiation exposure, and is well tolerated by patients. It is excellent for diagnosing biliary pathology and focal liver lesions. In addition, liver injury can be evaluated in trauma patients using the focused abdominal sonography for trauma examination. Limitations of ultrasound include incomplete imaging of the liver, most often at the dome or beneath ribs on the surface, and incomplete visualization of lesion boundaries. Moreover, obesity and overlying bowel gas also can interfere with image quality. Thus, ultrasonographically detected masses usually require further evaluation by other imaging modalities due to the lower sensitivity and specificity of ultrasound compared with CT and MRI.

The advent of contrast-enhanced ultrasound has improved the ability of this modality to differentiate among benign and malignant lesions. The injection of gas microbubble agents can
increase the sensitivity and specificity of ultrasound in detecting and diagnosing liver lesions. Microbubbles are <10 μm and, when given intravenously, allow for more effective echo enhancement. Contrast-enhanced ultrasound imaging of the liver improves delineation of liver lesions through identification of dynamic enhancement patterns and the vascular morphology of the lesion. In addition, some agents exhibit a late liver-specific phase in which the bubbles are taken up by cells in the reticuloendothelial system and accumulate in normal liver parenchyma after the vascular enhancement has faded.

The use of intraoperative ultrasound of the liver has rapidly expanded over the years with the increasing number and complexity of hepatic resections being performed. It has the ability to provide the surgeon with real-time accurate information useful for surgical planning. Intraoperative ultrasound is considered the gold standard for detecting liver lesions, and studies have shown that it can identify 20% to 30% more lesions than other preoperative imaging modalities. Importantly, it has been shown to influence surgical management in almost 50% of planned liver resections for malignancies. Applications for intraoperative ultrasound of the liver include tumor staging, visualization of intrahepatic vascular structures (Fig. 31-10), and guidance of resection plane by assessment of the relationship of a mass to the vessels. In addition, biopsy of lesions and ablation of tumors can be guided by intraoperative ultrasound.

Ultrasound elastography, also referred to as transient elastography, can be used to assess the degree of fibrosis or cirrhosis in the liver. Low-frequency vibrations transmitted through the liver induce an elastic shear wave that is detected by pulse-echo ultrasonography as the wave propagates through the liver. The velocity of the wave correlates with the stiffness of the organ—the wave travels faster through fibrotic or cirrhotic tissues. Ultrasound elastography has been found in large cohorts of individuals to have a sensitivity of 87% and a specificity of 91% for the diagnosis of cirrhosis when compared with liver biopsy. Unlike liver biopsy, ultrasound elastography is noninvasive and can be repeated often without additional risk to the patient. Furthermore, this rapid test can acquire information from a larger area of the tissue relative to needle biopsy, providing a better understanding of the entire hepatic parenchyma and reducing sampling error.

**Computed Tomography**

CT produces a digitally processed cross-sectional image of the body from a large series of X-ray images. The introduction of helical (spiral) CT has improved the imaging capabilities of this technique compared to earlier conventional axial CT by combining a continuous patient-table motion with continuous rotation of the CT gantry and allowing rapid acquisition of a volume of data within a single breath hold. With the recent advent of multidetector row CT scanners, high-resolution images can be obtained in submillimeter section thickness within a short scan time, with virtually no penalty in increased radiation dose. Together, these technologic advances have led to reduced motion artifacts due to variations in inspiration, facilitated optimal contrast delivery, and allowed for the capability to generate high-resolution reformations in any desired plane. In a single examination, modern-day CT scans provide detailed morphologic information on the number, size, distribution, and vascularity of liver lesions, all of which are vital in guiding the clinical management and therapeutic plan.

Contrast medium is routinely used in CT evaluation of the liver because of the similar densities of most pathologic liver masses and normal hepatic parenchyma. A CT scan with a dual- or triple-phase bolus of intravenous contrast agent is performed to achieve the greatest enhancement of contrast between normal and pathologic tissues. Ideally, contrast media should be selectively delivered to either the tumor or the liver, but not both. Radiologists use the dual blood supply of the liver and the hemodynamics of hepatic tumors to achieve this goal. The liver is unique in that it has a dual blood supply. As previously noted, the portal vein supplies approximately 75% of the blood...
flow and the hepatic artery the remaining 25%. However, many liver tumors receive the majority of their blood supply from the hepatic artery. After injection of the contrast agent, the rapid scan time of helical CT allows for CT sections through the liver in both the arterial dominant phase (20 to 30 seconds after the beginning of contrast delivery) and venous or portal dominant phase (60 to 70 seconds after contrast injection) (Fig. 31-11). Thus, many hepatic tumors that derive the majority of their blood supply from the hepatic artery as well as other hypervascular lesions are well delineated in the arterial phase. On the other hand, the portal phase provides optimal enhancement of the normal liver parenchyma because the majority of its blood supply is derived from the portal vein. This allows for detection of hypovascular lesions because they will appear hypoattenuated in relation to the brighter normal liver parenchyma. Furthermore, the arterial and portal phase images allow for noninvasive mapping of the hepatic arterial and venous anatomy, information that is crucial in the preoperative planning for patients undergoing liver surgery.

CT cholangiography has emerged as a new imaging modality for biliary disease. This technique usually involves the use of contrast agents, which are excreted by hepatocytes into the bile ducts. CT cholangiography, therefore, provides information on hepatocyte function and bile flow, in addition to high-resolution depiction of the biliary tree. It has compared well to endoscopic retrograde cholangiopancreatography (ERCP) in identifying obstructive biliary disease. One of the major advantages of CT cholangiography over other imaging modalities is the ability to depict small nondilated peripheral biliary radicals. This technique may be useful in the context of live liver donation or complex biliary surgery to aid in the preoperative depiction of biliary anatomy. It also may be applicable in the postoperative setting for the detection of biliary leakage or obstruction. A limitation of CT cholangiography is that the biliary tree may not be well visualized in patients with excessively dilated bile ducts or in those with hyperbilirubinemia, as bilirubin excretion is impaired in these cases.

**Magnetic Resonance Imaging**

MRI is a technique that produces images based on magnetic fields and radio waves. The MRI scanner creates a powerful magnetic field that aligns the hydrogen atoms in the body, and radio waves are used to alter the alignment of this magnetization. Different tissues absorb and release radio wave energy at different rates, and this information is used to construct an image of the body. Most tissues can be differentiated by differences in their characteristic T1 and T2 relaxation times. T1 is a measure of how quickly a tissue can become magnetized, and T2 measures how quickly it loses its magnetization. As with CT technology, advances in MRI now provide the opportunity to perform single-breath T1-weighted imaging and respiration-triggered T2-weighted imaging. The development of breath-hold imaging techniques has eliminated many of the motion artifacts that previously limited the sensitivity and application of MRI for imaging of the liver. Compared with CT scanning, the major advantages of MRI pertain to higher soft tissue contrast resolution and excellent depiction of fluid-containing structures, while obviating the need for ionizing radiation.

As with the iodinated contrast media use in CT scanning, multiple contrast agents have been developed for MRI to increase the difference in signal intensity between normal liver and pathologic lesions. Various gadolinium-based compounds have been used as MRI contrast agents that behave in a manner very similar to iodine in CT. Liver-specific MRI contrast agents also have been developed that rely either on excretion by Kupffer cells, such as ferumoxide (Feridex, Advanced Magnetics, Cambridge, MA), or on secretion in bile by hepatocytes, including gadoxetate (Eovist or Primovist, Bayer-Schering, Berlin, Germany). These agents combine the information obtained during a standard MRI with additional functional data, which in turn yields improved detection and characterization of lesions within the liver.

Just as ultrasound elastography is useful in the diagnosis of hepatic fibrosis and cirrhosis, magnetic resonance (MR) elastography appears promising as an imaging modality in
reducing the need for liver biopsy. In this technique, a vibration device is used to induce a shear wave in the liver. A modified MRI machine detects the shear wave, then generates a color-coded image that depicts the wave velocity, and hence stiffness, throughout the organ. Although preliminary studies have shown that MR elastography can detect cirrhosis with a high degree of accuracy, the clinical utility of this modality, especially given its relatively high cost, remains to be determined.

Magnetic resonance cholangiopancreatography (MRCP) enables rapid, noninvasive depiction of both the biliary tree and the pancreatic duct without the use of ionizing radiation or intravenous contrast media. One of the most common clinical indications for MRCP is biliary obstruction. MRCP provides visualization of dilated bile ducts, and the high spatial and contrast resolution often enables accurate assessment of the level of occlusion in the biliary tree. MRCP also can be enhanced with liver-specific MRI contrast agents that are actively secreted into the bile, but the clinical indications for such studies are still a matter of intensive investigation.

**Positron Emission Tomography**

Positron emission tomography (PET) is a nuclear medicine test that produces images of metabolic activity in tissues by detecting gamma rays emitted by a radioisotope incorporated into a metabolically active molecule. Fluorodeoxyglucose (FDG) is the most common metabolic molecule used in PET imaging. Although traditional imaging such as CT, ultrasound, and MRI provide anatomic information, PET offers functional imaging of tissues with high metabolic activity, including most types of metastatic tumors. PET imaging increasingly is used as a tool in the diagnostic evaluation of a patient with potentially resectable metastatic tumors. PET imaging increasingly is used as a tool in the diagnostic evaluation of a patient with potentially resectable metastatic tumors. PET imaging increasingly is used as a tool in the diagnostic evaluation of a patient with potentially resectable metastatic tumors. PET imaging increasingly is used as a tool in the diagnostic evaluation of a patient with potentially resectable metastatic tumors.

More than 20% of patients with colorectal cancer initially present with hepatic metastasis, and a large percentage of patients undergoing resection for their primary colorectal cancer eventually experience disease recurrence in the liver. The role of FDG-PET/CT in colorectal cancers lies predominantly in tumor staging and follow-up, particularly in the detection of occult intrahepatic metastases or extrahepatic disease. Although hepatic resection of colorectal metastases provides survival rates nearing 50%, the presence of extrahepatic disease is a poor prognosticator and usually precludes aggressive surgical intervention. Thus, accurate information regarding the extent of the disease is necessary for management of patients with colorectal metastases. PET/CT has also been shown to be more accurate than contrast-enhanced CT in tumor surveillance after radiofrequency ablation. The sensitivity of FDG-PET, however, is lowered by neoadjuvant chemotherapy, most likely secondary to reduced metabolic activity within the tumor.

Although the role of PET/CT in the clinical management of liver metastases has been well-established, its utility in the diagnostic workup of primary liver tumors is still debated. In hepatocellular carcinoma (HCC), FDG uptake correlates with the degree of differentiation—high-grade HCC lesions have increased FDG uptake compared to low-grade HCCs. As a result, the overall sensitivity of FDG-PET/CT in the detection of HCCs is reported to be only 50% to 65%, rendering this modality insufficient when used alone in the diagnosis of primary HCCs. For this reason, dual-tracer PET has been introduced to improve sensitivity in detecting all HCC. This modality combines the use of FDG, which accumulates in poorly differentiated tumors, with 11C-acetate, a tracer preferentially accumulated by well-differentiated HCC lesions. Although the clinical benefits of dual-tracer PET/CT have yet to be fully established, this combined modality has the potential to become a valuable tool in the diagnosis and staging of HCC. In cholangiocarcinoma tumors, FDG avidity depends on the morphologic characteristics and location of the lesion. Therefore, FDG-PET and FDG-PET/CT have not been shown to be highly beneficial in the diagnosis of primary cholangiocarcinoma, but they may be beneficial in the detection of regional and distal metastases, which can affect clinical decision making and patient management.

**ACUTE LIVER FAILURE**

Acute liver failure (ALF) occurs when the rate and extent of hepatocyte death exceeds the liver’s regenerative capabilities. It was initially described as a specific disease entity in the 1950s.

---

**Figure 31-12.** Computed tomography (CT)–positron emission tomography (PET) scans before and after resection of liver metastasis from colorectal cancer in a 54-year-old patient. CT scan shows large 10-cm right lobe liver metastasis (left panel), and PET scan findings are strongly positive (middle panel). Two years after right hepatectomy, the patient has no evidence of recurrence and significant hypertrophy of the left lobe (right panel).
It also has been referred to as fulminant hepatic failure. ALF is a rare disorder affecting approximately 2000 patients annually in the United States. ALF is defined by the development of hepatic encephalopathy occurring within 26 weeks of severe liver injury in a patient without a history of previous liver disease or portal hypertension. The manifestations of ALF may include cerebral edema, hemodynamic instability, increased susceptibility to bacterial and fungal infections, renal failure, coagulopathy, and metabolic disturbances. Even with current medical care, ALF can progress rapidly to hepatic coma and death. The most common cause of death is intracranial hypertension due to cerebral edema, followed by sepsis and multisystem organ failure. The causes of ALF, which are the most important variables in determining outcome, are numerous and can include viral infection as well as drug overdose, reaction, and toxicity. It has been determined that the etiologic factor leading to ALF varies according to geographic location. Before the introduction of orthotopic liver transplantation (OLT), the chance for survival was <20%. Currently, most series report 5-year survival rates of >70% for affected patients.

**Etiology**

Differences in etiology, management, and patient outcomes have been described for various regions of the globe. In the East and developing portions of the world, the most common causes of ALF are viral infections, primarily hepatitis B, A, and E. In these areas, there are a relatively small number of drug-induced cases. In contrast, 65% of cases of ALF in the West are thought to be due to drugs and toxins, with acetaminophen (paracetamol) being the most common etiologic agent in the United States, Australia, United Kingdom, and most of Europe. In France and Spain, where acetaminophen sales are restricted, the rate of acetaminophen-induced ALF is quite low. Acetaminophen-induced ALF is also uncommon in South America. The U.S. Acute Liver Failure Study Group identified several other causes of ALF, including autoimmune hepatitis, hypoperfusion of the liver (in cardiomyopathy or cardiogenic shock), pregnancy-related conditions, and Wilson’s disease. Even with exhaustive efforts to identify a cause, approximately 20% of all cases of ALF remain indeterminate in origin.

**Clinical Presentation**

In a multicenter study involving 17 tertiary care centers and 308 patients in the United States, 73% of all patients with ALF were female, with a median age of 38 years. The most common ethnic group affected was whites (74%), followed by Hispanics (9%) and African Americans (3%). Patients were ill for a median of 6 days before the onset of encephalopathy and had a median of 2 days between the onset of jaundice and the development of encephalopathy. Hepatic coma grade at presentation was approximately equally distributed across grades I to IV. Eighty-four percent of the patients in the study were referred from outside hospitals, 40% had a serum creatinine level exceeding 2.0 mg/dL, and 14% had an arterial pH of <7.30. In addition, 44% of the patients acquired a culture-proven infection.

**Diagnosis and Clinical Management**

When the medical history is obtained, it is important to address the possibility of exposure to viral infections, medications, and other possible toxins. The possibility of previous liver disease needs to be explored. The physical examination must assess and document the patient’s mental status as well as attempt to identify findings of chronic liver disease. The initial laboratory examination must evaluate the severity of the ALF as well as attempt to identify the cause (Table 31-1). A liver biopsy should be performed if certain disease entities such as autoimmune hepatitis or lymphoma are a possibility. Because of the associated coagulopathy, if a liver biopsy is needed, it is usually safest to obtain the tissue via a transjugular approach. Patients with ALF should be admitted to the hospital and monitored frequently. Due to the rapidity with which this disease process may progress, a liver transplant center should be contacted, and the affected patient should be transferred to the center early in the evaluation period.

If acetaminophen overdose is suspected to have occurred within a few hours of presentation, administration of activated charcoal may be useful to reduce the volume of acetaminophen present in the gastrointestinal (GI) tract. N-acetylcysteine (NAC), the clinically effective antidote for acetaminophen overdose, should be administered as early as possible to any patient with suspected acetaminophen-associated ALF. NAC also should be administered to patients with ALF of unclear etiology because replenishing glutathione may be beneficial in this patient population as well. NAC can be administered either orally (140 mg/kg initial dose, followed by 70 mg/kg every 4 hours × 17 doses) or via the intravenous route (loading dose of 150 mg/kg, followed by a maintenance dose of 50 mg/kg every 4 hours × 12 doses). For patients who are suspected of having drug-induced hepatotoxicity, it is important to obtain details regarding all prescription and nonprescription drugs, herbs, and dietary supplements that may have been taken in the previous year. Most instances of drug-induced hepatotoxicity occur in the first 6 months after drug initiation. Any suspected offending agent must be discontinued, and an attempt should be made to administer only essential medications.

The majority of patients with ALF need to be monitored in the intensive care unit (ICU) setting, and specific attention needs to be given to fluid management, ulcer prophylaxis, hemodynamic monitoring, electrolyte management, and surveillance for and treatment of infection. Surveillance cultures should be performed to identify bacterial and fungal infections as early as possible. Serum phosphorus levels need to be monitored.

**Table 31-1**

<table>
<thead>
<tr>
<th>Acute liver failure laboratory evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete blood count</td>
</tr>
<tr>
<td>Complete metabolic panel</td>
</tr>
<tr>
<td>Amylase and lipase levels</td>
</tr>
<tr>
<td>Liver function tests</td>
</tr>
<tr>
<td>Prothrombin time/international normalized ratio</td>
</tr>
<tr>
<td>Factor V level</td>
</tr>
<tr>
<td>Factor VII level</td>
</tr>
<tr>
<td>Arterial blood gas concentrations</td>
</tr>
<tr>
<td>Arterial serum ammonia level</td>
</tr>
<tr>
<td>ABO typing</td>
</tr>
<tr>
<td>Acute hepatitis panel</td>
</tr>
<tr>
<td>Autoimmune marker levels</td>
</tr>
<tr>
<td>Ceruloplasmin level</td>
</tr>
<tr>
<td>Toxicology screening</td>
</tr>
<tr>
<td>Acetaminophen level</td>
</tr>
<tr>
<td>HIV screening</td>
</tr>
<tr>
<td>Pregnancy test (females)</td>
</tr>
</tbody>
</table>

HIV = human immunodeficiency virus.
Hypophosphatemia, a sign of hepatic regeneration, may indicate a higher likelihood of spontaneous recovery and needs to be corrected via intravenous (IV) administration of phosphate. Sedation should be avoided, and the head of the bed should be elevated at least 30°. Neurologic examinations should be performed frequently. Intracranial pressure monitoring is reserved for patients in whom a neurologic examination is no longer reliable. CT scans of the head should be performed to rule out mass lesion or hemorrhage, but they provide only limited information regarding increased intracranial pressure. The administration of blood products for thrombocytopenia and prolonged PT is recommended only in the setting of hemorrhage or before invasive procedures. Acute renal failure is a frequent complication in patients with ALF, and efforts should be made to protect renal function by maintaining sufficient perfusion and avoiding nephrotoxic medications. Should renal replacement therapy become necessary, continuous venovenous hemodialysis should be used rather than intermittent hemodialysis because continuous venovenous hemodialysis provides better hemodynamic and intracranial pressure stability. The most severely affected patients have a poor prognosis with medical management alone and require liver transplantation. Identifying these patients early in the clinical course is important both to maximize the time available to obtain a donor liver allograft for those in need and to avoid transplant in those who will recover without it.

**Prognosis**

Accurate identification of ALF patients who will recover spontaneously is important because of the severe shortage of donor liver allografts and the potential complications of lifelong non-specific immunosuppression. The most widely applied prognostic scoring system is the King’s College Hospital ALF criteria. This scoring system has separate criteria predicting a poor medical management outcome for acetaminophen-related and non–acetaminophen-related forms of ALF (Table 31-2). Many other prognostic models exist such as the Acute Physiology and Chronic Health Evaluation II (APACHE II) score, the Clichy criteria, and actin-free Gc-globulin serum concentration. Overall, prognostic scoring systems have proven to have acceptable specificity but low sensitivity in determining patient outcome and therefore should not replace the judgment of an experienced clinician.

**Liver Transplantation**

Despite advances in medical management, OLT remains the only definitive therapy for patients unable to regenerate sufficient hepatocyte mass in a timely manner. The advent of OLT has coincided with a rise in overall ALF survival rates from approximately 20% in the pretransplantation era to >70% at the present time. One-year posttransplant survival for patients with ALF has been reported to be as high as 80% to 90%. Although these improvements in survival rates are impressive, it should be noted that 10% of patients still die while awaiting OLT, which confirms that the potential for improved patient outcome still has not been realized because of the ongoing liver allograft shortage.

**Emerging Technologies**

As mentioned earlier, patient survival could be improved if additional time could be gained for the patient while awaiting liver replacement or hepatocyte regeneration. The development of a support device to replace the acutely failing liver has been a highly sought after (and elusive) goal. Several systems have been tested without definitive evidence of efficacy. Transient improvement in hepatic encephalopathy has been observed in several trials, but improvement in hepatocyte function and long-term benefit have not been realized. Liver support trials are difficult to perform due to access to liver replacement, the rarity of affected patients, and the heterogeneous causes and varying levels of disease severity. Therefore, additional data are necessary, and liver support systems should be used only as part of an approved clinical trial. Focus has now shifted toward xenotransplantation, organ engineering, and cell transplantation. In the meantime, living donation and auxiliary liver transplantation can help overcome the organ shortage.

### CIRRHOSIS AND PORTAL HYPERTENSION

Cirrhosis, the final sequela of chronic hepatic insult, is characterized by the presence of fibrous septa throughout the liver subdividing the parenchyma into hepatocellular nodules (Fig. 31-13). Cirrhosis is the consequence of sustained wound healing in response to chronic liver injury. Approximately 40% of cirrhotic patients are asymptomatic, but progressive deterioration leading to the need for liver transplantation or death is typical after the development of end-stage liver disease (ESLD). The complications of ESLD include progressive hyperbilirubinemia, malnutrition, decreased synthetic function of the liver, coagulopathy, portal hypertension (i.e., ascites and variceal bleeding), hepatic encephalopathy, and life-limiting fatigue. ESLD carries a 5-year mortality of 50%, with 70% of deaths due to liver failure. In the United States, cirrhosis accounts for 30,000 deaths per year and is the most common nonneoplastic cause of death among patients with hepatobiliary and digestive diseases. An additional 10,000 to 12,000 deaths occur annually due to HCC, the most rapidly increasing neoplasm in the United States.

### Morphologic Classification of Cirrhosis

Morphologically, cirrhosis can be described as micronodular, macronodular, or mixed. Micronodular cirrhosis is characterized
by thick regular septa, small uniform regenerative nodules, and involvement of virtually every hepatic lobule. Macronodular cirrhosis frequently has septa and regenerative nodules of varying sizes. The regenerative nodules consist of irregularly sized hepatocytes with large nuclei and cell plates of varying thickness. Mixed cirrhosis is present when regeneration is occurring in a micronodular liver and over time converts to a macronodular pattern. This morphologic categorization is limited, and cirrhosis is a dynamic process in which nodule size varies over time. The three patterns correlate poorly with etiology, and the same pattern can result from a variety of disease processes. Conversely, a single disease process can demonstrate several morphologic patterns. Irrespective of etiology and morphologic pattern, the cirrhotic liver frequently demonstrates right hepatic lobe atrophy, caudate lobe and left lateral segment hypertrophy, recanalization of the umbilical vein, a nodular surface contour, dilatation of the portal vein, gastroesophageal varices, and splenomegaly on radiographic evaluation.

**Etiology of Cirrhosis**

Cirrhosis can result from a wide range of disease processes, including viral, autoimmune, drug-induced, alcohol-induced, nonalcoholic fatty liver disease, and metabolic diseases (Table 31-3). In the diagnosis of alcoholic liver disease, documentation of chronic alcohol abuse is imperative. Liver biopsy will reveal the typical findings of alcoholic hepatitis, including hepatocyte necrosis, Mallory bodies, neutrophil infiltration, and perivenular inflammation.

Nonalcoholic fatty liver disease (NAFLD) covers a wide spectrum of disorders including simple fatty liver, nonalcoholic steatohepatitis (NASH), fibrosis/cirrhosis and NASH-associated hepatocellular carcinoma.47 NAFLD is now the most common chronic liver disease worldwide48 and NASH is a progressive form of NAFLD characterized by steatosis with hepatocellular injury and chronic inflammation.49 NASH affects 3% to 5% of the population, and approximately 1 in 10 NASH patients will progress to cirrhosis, thereby placing them at risk for the well-described consequences of cirrhosis, including hepatocellular carcinoma.50 Although hepatocellular carcinoma arises less frequently in patients with NASH compared to other liver diseases (e.g., hepatitis C viral infection), the overall higher prevalence and more rapidly increasing incidence of NASH relative to other chronic liver diseases mean that the majority of HCC will arise in the setting of NAFLD in the near future.51,52

Patients with nonalcoholic steatohepatitis (NASH) often endorse a history of diabetes mellitus or metabolic syndrome. The diagnosis of NASH requires the demonstration of steatohepatitis on biopsy, the lack of a history of significant alcohol consumption, and exclusion of other causes of hepatic steatosis. Although cryptogenic cirrhosis, or cirrhosis without an apparent cause, accounted for a third of all cases in the past, this proportion has declined over time as it becomes increasingly apparent that many of such patients may actually have unrecognized NASH.

Due to the high prevalence of fatty liver disease, many patients considered for hepatic surgery will have background hepatic steatosis or steatohepatitis. In addition, chemotherapy treatment (e.g., irinotecan) for colorectal cancer liver metastases induces steatosis and steatohepatitis in the nontumor-bearing liver.53 This has important implications as fatty liver disease can increase morbidity after liver resection.54,55 Thus, understanding the deleterious effects of steatosis and steatohepatitis...
is crucial for the multidisciplinary care of patients undergoing liver surgery.

Chronic hepatitis C infection is the most common cause of chronic liver disease and the most frequent indication for liver transplantation in the United States. The identification of chronic hepatitis C infection is facilitated by serologic assays that detect antibody to hepatitis C and molecular assays that quantify hepatitis C viral RNA. Chronic hepatitis B, on the other hand, can be diagnosed based on the detection of hepatitis B surface antigen (HBsAg) more than 4 to 6 months after initial infection. Additional tests of hepatitis B viral replication, such as the hepatitis B e antigen (HBeAg) and hepatitis B viral DNA, can be used to confirm ongoing infection and to guide appropriate antiviral therapy.

Autoimmune causes of cirrhosis include primary biliary cirrhosis, primary sclerosing cholangitis, and autoimmune hepatitis. Patients with primary biliary cirrhosis may be asymptomatic or may present with a history of fatigue, pruritus, and skin hyperpigmentation that is not related to jaundice. Anti-mitochondrial antibodies will test positive in the vast majority of cases. Affected patients also may have marked elevations in serum cholesterol, while hyperbilirubinemia is seen late in the course of the disease. Primary sclerosing cholangitis is a chronic cholestatic disease of the liver associated with ulcerative colitis and Crohn’s disease. The clinical presentation can include pruritus, steatorrhea, fat-soluble vitamin deficiencies, and metabolic bone disease. The diagnosis is often established by imaging of the biliary tree, which reveals a characteristic picture of diffuse, multifocal strictures with focal dilation of the bile ducts resulting in a beaded appearance. Complications are common and can include biliary strictures, cholangitis, cholelithiasis, and cholangiocarcinoma. Autoimmune hepatitis is often accompanied by an elevation in serum globulins, particularly gamma globulins. Liver biopsy will show nonspecific changes such as a portal mononuclear cell infiltrate with the characteristic presence of plasma cells. Many patients with autoimmune hepatitis will respond to treatment with prednisone with or without azathioprine.

Hereditary hemochromatosis is the most common metabolic disorder causing cirrhosis. This entity should be suspected if the patient’s clinical presentation includes skin hyperpigmentation, diabetes mellitus, pseudogout, cardiomyopathy, or a family history of cirrhosis. Elevated plasma ferritin and increased iron saturation levels suggest the presence of iron overload, but these findings also can be seen in other diseases of the liver. Confirmatory testing can be achieved by means of genetic testing, liver biopsy, or by assessing the response to phlebotomy. Other uncommon metabolic disorders leading to cirrhosis include Wilson’s disease and α1-antitrypsin deficiency.

Clinical Manifestations of Cirrhosis

The clinical history associated with cirrhosis can include fatigue, anorexia, weight loss, jaundice, abdominal pain, peripheral edema, ascites, GI bleeding, and hepatic encephalopathy. On physical examination, a number of findings have been described in patients with cirrhosis. Spider angiomata and palmar erythema are believed to be caused by alterations in sex hormone metabolism. Finger clubbing may be a consequence of hypoalbuminemia, while the pathogenesis of white nail beds and Dupuytren’s contractures are less well understood. Males may develop features of feminization such as gynecomastia, loss of chest and axillary hair, and testicular atrophy. Spleenomegaly is common, whereas the cirrhotic liver itself may be enlarged, normal sized, or small. Ascites and pleural effusion can be seen with fluid accumulation. Portal hypertension can manifest as caput medusae and/or the presence of the Cruveilhier-Baumgarten murmur, a venous hum that can be auscultated in the epigastrium resulting from collaterals between the portal system and the remnant of the umbilical vein. Jaundice usually does not appear until the bilirubin rises above 2 to 3 mg/dL. Asterixis can be detected in patients with hepatic encephalopathy. Other manifestations include fetor hepaticus, as well as features suggestive of malnutrition such as weakness, weight loss, and temporal muscle wasting.

Although fat stores and muscle mass are reduced, resting energy expenditure is increased. Muscle cramps occur frequently in the cirrhotic patient and are felt to correlate with ascites, low mean arterial pressure, and plasma renin activity. Abdominal hernias are common with ascites and should be electively repaired only in patients with well-compensated cirrhosis; otherwise, the hernia should be repaired at the time of or after hepatic transplantation. HCC can occur in all forms of cirrhosis, and every cirrhotic patient should undergo screening for the development of HCC every 6 months via imaging and measurement of a serum α-fetoprotein (AFP) level. Cirrhosis is associated with increased cardiac output and heart rate as well as decreased systemic vascular resistance and blood pressure. Patients with cirrhosis are more prone to infections due to impaired phagocytic activity of the reticuloendothelial system. Bacterial infections, often of intestinal origin, are common and must be suspected in a patient with unexplained pyrexia or clinical deterioration. Spontaneous bacterial peritonitis also is seen in cases of cirrhosis with ascites. Intrinsic drug metabolism is reduced in the cirrhotic liver, and this fact needs to be recognized when prescribing medications.

Laboratory Findings Associated With Cirrhosis

Laboratory findings vary in the cirrhotic patient depending on the degree of compensation; however, in general, a number of trends are seen. The cirrhotic patient usually has a mild normocytic normochromic anemia. The white blood cell and platelet counts are reduced, and the bone marrow is macronormoblastic. The PT is prolonged and does not respond to vitamin K therapy, and the serum albumin level is depressed. Urobilinogen is present and urinary sodium excretion is diminished in the presence of ascites. The serum levels of bilirubin, transaminases, and alkaline phosphatase may all be elevated. However, normal liver function test results do not eliminate the possibility of cirrhosis.

Liver Biopsy

The diagnosis of cirrhosis can be made in many cases from a constellation of clinical features, laboratory values, and radiographic findings. Histopathologic examination of liver tissue is occasionally needed to confirm the diagnosis of cirrhosis and in determining disease etiology, activity, and progression. Liver biopsy can be performed via a percutaneous, transjugular, or laparoscopic approach. If needed, ultrasound or CT guidance can be helpful in obtaining an adequate sample and avoiding other viscera.

Various serologic markers of hepatic fibrosis are currently being investigated to help predict the presence of cirrhosis without the need for liver biopsy. However, no currently available marker is sufficiently accurate for clinical use. Ultrasound elastography, which measures the stiffness of the liver by inducing
an elastic shear wave that propagates through the tissue, shows promise as a noninvasive test in identifying patients with advanced fibrosis and cirrhosis. (See earlier section, “Radiologic Evaluation of the Liver.”)

**Hepatic Reserve and Assessment of Surgical Risk in the Cirrhotic Patient**

Assessing the hepatic reserve of the cirrhotic patient is important because cirrhosis and portal hypertension can have a negative impact on the outcome of nontransplant surgical procedures. Patients with liver disease undergoing surgery are at increased risk for surgical and anesthesia-related complications. The actual risk depends on the type of anesthetic used, the specific surgical procedure performed, and the severity of liver disease. Previous studies have demonstrated that emergency operations, cardiac surgery, hepatic resections, and abdominal surgery, particularly cholecystectomy, gastric resection, and colectomy, generate the highest operative risk among cirrhotic patients. Additionally, preoperative patient characteristics such as anemia, ascites, encephalopathy, malnutrition, hypoalbuminemia, hypoxemia, infection, jaundice, portal hypertension, and prolonged PT have also been associated with inferior outcomes after surgery. Nontransplant surgical procedures are contraindicated in patients with acute fulminant hepatitis and those with severe decompensated chronic hepatitis.

A number of laboratory tests have been used to assess hepatic reserve in patients with cirrhosis. Tests of indocyanine green, sorbitol, and galactose elimination capacity as well as the carbon-13 galactose breath test and carbon-13 aminopyrine breath test have all been disappointing clinically due to their dependence on flow to the liver as well as the unavailability and complexity of the tests. The monoethylglycinexylidide (MEGX) test, which measures MEGX formation after the administration of lidocaine, has been shown to be approximately 80% sensitive and specific in diagnosing cirrhosis. However, this test loses both sensitivity and specificity as the serum bilirubin level rises and interferes with the fluorescent readout system.

**Child-Turcotte-Pugh Score**

The Child-Turcotte-Pugh (CTP) score was originally developed to evaluate the risk of portocaval shunt procedures performed for portal hypertension and subsequently has been shown to be useful in predicting surgical risks of other intra-abdominal operations on cirrhotic patients (Table 31-4). Numerous studies have demonstrated overall surgical mortality rates of 10% for patients with class A cirrhosis, 30% for those with class B cirrhosis, and 75% to 80% for those with class C cirrhosis. The CTP score is derived from five variables as shown in Table 31-4. The problems with the CTP score are the presence of subjective variables (encephalopathy and ascites), its narrow range (5 to 15 points), and the equal weighting given to each variable. Multiple retrospective studies have demonstrated that perioperative mortality and morbidity rates correlate well with the CTP score, and for over 30 years, this measure had been used as the principal predictor of operative risk.

**Model for End-Stage Liver Disease Scoring System**

The Model for End-Stage Liver Disease (MELD) is a linear regression model based on three objective laboratory values (INR, bilirubin level, and creatinine level). It was originally developed as a tool to predict mortality after transjugular intrahepatic portosystemic shunt (TIPS) but has been validated and used as the sole method of liver transplant allocation in the United States since 2002. The MELD formula is as follows:

\[
\text{MELD Score} = 9.57 \ln(\text{SCR}) + 3.78 \ln(\text{Tbil}) + 11.2 \ln(\text{INR}) + 6.43
\]

where Ln represents natural logarithm, SCr is serum creatinine level (in milligrams per deciliter), and Tbil is serum bilirubin level (in milligrams per deciliter).

A number of studies have examined the relative values of MELD and CTP scores in predicting postoperative mortality in cirrhotic patients undergoing nontransplant surgical procedures. Northup and colleagues demonstrated that MELD score was the only statistically significant predictor of 30-day mortality. In this study, mortality increased by approximately 1% for each MELD point up to a score of 20 and by 2% for each MELD point above 20. A comparison of the MELD model with the CTP classification showed good correlation between the two measures in predicting mortality, especially in the setting of emergency surgery. In these studies, the relative risk of mortality increased by 14% for each 1-point increase in MELD score. As a result, it has been proposed that patients with a MELD score below 10 can safely undergo elective surgery, those with MELD between 10 and 15 may undergo surgery with caution, while those with MELD scores in excess of 15 should not be subjected to elective surgical procedures.

**Portal Hypertension**

The portal venous system contributes approximately 75% of the blood and 72% of the oxygen supplied to the liver. In the average adult, 1000 to 1500 mL/min of portal venous blood is supplied to the liver. However, this amount can be significantly increased in the cirrhotic patient. The portal venous system is without valves and drains blood from the spleen, pancreas, gallbladder, and abdominal portion of the alimentary tract into the liver. Tributaries of the portal vein communicate with veins draining directly into the systemic circulation. These collateral communications occur at the gastroesophageal junction, anal canal, falciform ligament, splenic venous bed and left renal vein, and retroperitoneum (Fig. 31-14). The normal portal venous pressure is 5 to 10 mmHg, and at this pressure, very little blood is shunted from the portal venous system into the systemic circulation. As portal venous pressure increases, however, the
collateral communications with the systemic circulation dilate, and a large amount of blood may be shunted around the liver and into the systemic circulation.

**Imaging of the Portal Venous System and Measurement of Portal Venous Pressure**

The patency of the portal vein and the nature of the collateral circulation should be established. An understanding of portal vein patency and anatomy is crucial before undertaking porto-systemic shunts, hepatic resection, or hepatic transplantation. The simplest initial investigation is abdominal ultrasonography. A large portal vein suggests portal hypertension but is not diagnostic. Doppler ultrasound is capable of outlining the anatomy of the portal vein, excluding the presence of thrombosis, and identifying the direction of portal venous blood flow. Doppler ultrasound also is useful in evaluating blood flow through surgical shunts and TIPS. Abdominal CT and magnetic resonance angiography both are capable of revealing portal vein anatomy as well as patency. Visceral angiography and portal venography are reserved for cases that cannot be evaluated satisfactorily by noninvasive methods and require further clarification of portal patency or anatomy.

The most accurate method of determining portal hypertension is hepatic venography. The most commonly used procedure involves placing a balloon catheter directly into the hepatic vein and measuring the free hepatic venous pressure (FHVP) with the balloon deflated and the wedged hepatic venous pressure (WHVP) with the balloon inflated to occlude the hepatic vein. The hepatic venous pressure gradient (HVPG) is then calculated by subtracting the free from the wedged venous pressure (HVPG = WHVP – FHVP). The HVPG represents the pressure in the hepatic sinusoids and portal vein and is a measure of portal venous pressure. Clinically significant portal hypertension is evident when HVPG exceeds 10 mmHg.

**Etiology and Clinical Features of Portal Hypertension**

The causes of portal hypertension can be divided into three major groups: presinusoidal, sinusoidal, and postsinusoidal. Although multiple disease processes can result in portal hypertension (Table 31-5), in the United States, the most common cause of portal hypertension is usually an intrahepatic one, namely, cirrhosis. The most significant clinical finding associated with portal hypertension is the development of gastro-esophageal varices, which are mainly supplied by the anterior branch of the left gastric (coronary) vein.

<table>
<thead>
<tr>
<th>Table 31-5: Etiology of portal hypertension</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Presinusoidal</strong></td>
</tr>
<tr>
<td>Sinistral/extrahepatic</td>
</tr>
<tr>
<td>Splenic vein thrombosis</td>
</tr>
<tr>
<td>Splenomegaly</td>
</tr>
<tr>
<td>Splenic arteriovenous fistula</td>
</tr>
<tr>
<td><strong>Intrahepatic</strong></td>
</tr>
<tr>
<td>Schistosomiasis</td>
</tr>
<tr>
<td>Congenital hepatic fibrosis</td>
</tr>
<tr>
<td>Nodular regenerative hyperplasia</td>
</tr>
<tr>
<td>Idiopathic portal fibrosis</td>
</tr>
<tr>
<td>Myeloproliferative disorder</td>
</tr>
<tr>
<td>Sarcoïd</td>
</tr>
<tr>
<td>Graft-versus-host disease</td>
</tr>
<tr>
<td><strong>Sinusoidal</strong></td>
</tr>
<tr>
<td><strong>Intrahepatic</strong></td>
</tr>
<tr>
<td>Cirrhosis</td>
</tr>
<tr>
<td>Viral infection</td>
</tr>
<tr>
<td>Alcohol abuse</td>
</tr>
<tr>
<td>Primary biliary cirrhosis</td>
</tr>
<tr>
<td>Autoimmune hepatitis</td>
</tr>
<tr>
<td>Primary sclerosing cholangitis</td>
</tr>
<tr>
<td>Metabolic abnormality</td>
</tr>
<tr>
<td><strong>Postsinusoidal</strong></td>
</tr>
<tr>
<td><strong>Intrahepatic</strong></td>
</tr>
<tr>
<td>Vascular occlusive disease</td>
</tr>
<tr>
<td><strong>Posthepatic</strong></td>
</tr>
<tr>
<td>Budd-Chiari syndrome</td>
</tr>
<tr>
<td>Congestive heart failure</td>
</tr>
<tr>
<td>Inferior vena caval web</td>
</tr>
<tr>
<td>Constrictive vena caval web</td>
</tr>
</tbody>
</table>

Figure 31-14. Intra-abdominal venous flow pathways leading to engorged veins (varices) from portal hypertension. 1, Coronary vein; 2, superior hemorrhoidal veins; 3, paraumbilical veins; 4, Retzius’ veins; 5, veins of Sappey; A, portal vein; B, splenic vein; C, superior mesenteric vein; D, inferior mesenteric vein; E, inferior vena cava; F, superior vena cava; G, hepatic veins; a, esophageal veins; a1, aygos system; b, vasa brevia; c, middle and inferior hemorrhoidal veins; d, intestinal; e, epigastric veins.
Portal hypertension also results in splenomegaly with enlarged, tortuous, and even aneurysmal splenic vessels. Splenomegaly is frequently associated with functional hypersplenism, causing leukopenia, thrombocytopenia, and anemia. Ascites occurs in the setting of severe portal hypertension in combination with hepatocytic dysfunction. The umbilical vein may recanalize and dilate, leading to visible collaterals on the abdominal wall. Anorectal varices are present in approximately 45% of cirrhotic patients and must be distinguished from hemorrhoids, which do not communicate with the portal system and are not present at increased incidence in patients with portal hypertension. Large spontaneous venous shunts may form between the portal venous system and the left renal vein or the IVC, but these shunts are ineffective in reducing portal venous pressures and preventing bleeding from gastroesophageal varices.

**Management of Gastroesophageal Varices**

The most significant manifestation and the leading cause of morbidity and mortality related to portal hypertension is variceal bleeding. Approximately 30% of patients with compensated cirrhosis and 60% of patients with decompensated cirrhosis have esophageal varices. One third of all patients with varices will experience variceal bleeding. Each episode of bleeding is associated with a 20% to 30% risk of mortality. If left untreated, 70% of patients who survive the initial bleed will experience recurrent variceal hemorrhage within 2 years of the index hemorrhage.

**Prevention of Variceal Bleeding**

Current measures aimed at preventing variceal bleeding include the administration of nonselective β-blockers and prophylactic endoscopic surveillance with variceal band ligation. Meta-analyses have demonstrated that nonselective β-blockers such as propranolol and nadolol reduce the index variceal bleed by approximately 45% and decrease bleeding mortality by 50%. However, approximately 20% of patients do not respond to β-blockade, and another 20% cannot tolerate β-blockade due to medication side effects. Endoscopic surveillance with prophylactic variceal band ligation has been associated with a lower incidence of a first variceal bleed.

Variceal band ligation is recommended for patients with medium to large varices, performed every 1 to 2 weeks until obliteration, followed by esophagogastroduodenoscopy (EGD) 1 to 3 months later and surveillance EGD every 6 months to monitor for recurrence of varices.

**Management of Acute Variceal Hemorrhage**

Patients with acute variceal hemorrhage should be admitted to an ICU for resuscitation and management. Blood resuscitation should be performed carefully to reach a hemoglobin level of approximately 8 g/dL. Overzealous replacement of blood products and administration of saline can lead to both rebleeding and increased mortality. Administration of fresh frozen plasma and platelets can be considered in patients with severe coagulopathy. Use of recombinant factor VIIa has not been shown to be more beneficial than standard therapy and therefore is not recommended at this time. Cirrhotic patients with variceal bleeding have a high risk of developing bacterial infections, which are associated with increased risks of rebleeding and mortality. Spontaneous bacterial peritonitis accounts for approximately half of these infections, with urinary tract infections and pneumonias comprising the remainder. The use of short-term prophylactic antibiotics (e.g., ceftriaxone 1 g intravenously) has been shown both to decrease the rate of bacterial infections and to increase survival.

Vasoactive medications decrease blood flow to the gastroesophageal varices and can be initiated as soon as the diagnosis of variceal bleeding is made. Although vasopressin is the most potent available vasoconstrictor, its use is limited by its systemic vasoconstrictive effects that can produce hypertension, myocardial ischemia, arrhythmias, ischemic abdominal pain, and limb gangrene. Octreotide, a somatostatin analog, has the advantage that it can be administered for 5 days or longer, and it is currently the preferred pharmacologic agent for initial management of acute variceal bleeding. In addition to pharmacologic therapy, endoscopy with variceal band ligation should be carried out as soon as possible. This combination of pharmacologic and endoscopic therapy has been shown both to improve the initial control of bleeding and to increase the 5-day hemostasis rate.

**Luminal Tamponade**

When medical and endoscopic measures fail to control variceal hemorrhage, balloon tamponade using a Sengstaken-Blakemore tube will control refractory bleeding in up to 90% of patients. However, its application is limited due to the potential for complications, which include aspiration, airway obstruction, and esophageal perforation due to overinflation or pressure necrosis. Therefore, the use of a Sengstaken-Blakemore tube should not exceed 36 hours to avoid tissue necrosis, and this treatment modality should only be considered a temporary bridge to more definitive measures of variceal hemorrhage control.

**Transjugular Intrahepatic Portosystemic Shunt**

The TIPS procedure involves implantation of a metallic stent between an intrahepatic branch of the portal vein and a hepatic vein radicle. The needle track is dilated until a portal pressure gradient of ≤12 mmHg is achieved. TIPS can be performed in 95% of patients by an experienced interventional radiologist, controls variceal bleeding in >90% of cases refractory to medical treatment, and should not affect subsequent hepatic transplantation. Possible complications include bleeding either intra-abdominally or via the biliary tree, infections, renal failure, decreased hepatic function, and hepatic encephalopathy, which occur in 25% to 30% of patients after the TIPS procedure. A high rate of thrombosis is seen and can be attributed to intimal hyperplasia of the metallic stent. Frequent follow-up with repeated interventions such as dilation or restenting often are needed to maintain TIPS patency.

**Balloon-Occluded Retrograde Transvenous Obliteration**

The balloon-occluded retrograde transvenous obliteration (BRTO) procedure has been used for the specific management of bleeding gastric varices in patients with spontaneous gastrorenal or splenorenal shunts shown on contrast-enhanced cross-sectional imaging. Using a transjugular or transfemoral approach, a balloon-occlusion catheter is directed through the left renal vein into the spontaneous shunt, which is then obliterated with the use of a sclerosing agent. BRTO effectively controls hemorrhage from gastric varices and preserves portal flow to the liver, thereby reducing the risk of hepatic encephalopathy relative to TIPS. The occlusion of spontaneous shunts, however, can theoretically exacerbate portal hypertension, precipitate hemorrhage from esophageal varices, and exacerbate the accumulation of ascites.

**Surgical Shunting**

The need for surgical shunts has been reduced since the introduction of the TIPS procedure and hepatic transplantation. At this time, the recommendation is that surgical shunts be considered only in patients who have MELD scores of <15, who are...
not candidates for hepatic transplantation, or who have limited access to TIPS therapy and the necessary follow-up. The aim of the surgical shunt is to reduce portal venous pressure, maintain total hepatic and portal blood flow, and avoid the high incidence of complicating hepatic encephalopathy. Patient survival is determined by hepatic reserve.

The portacaval shunt, as first described by Eck in 1877, either joins the portal vein to the IVC in an end-to-side fashion and completely disrupts portal vein flow to the liver, or joins it in a side-to-side fashion and thereby maintains partial portal venous flow to the liver. Currently this shunt is rarely performed due to the high incidence of hepatic encephalopathy and decreased liver function resulting from the reduction of portal perfusion. The Eck fistula also makes subsequent hepatic transplantation much more technically difficult.

The mesocaval shunt uses an 8- or 10-mm polytetrafluoroethylene (PTFE) graft to connect the superior mesenteric vein to the IVC. The mesocaval shunt is technically easier to perform and can be easily ligated during subsequent hepatic transplantation. The smaller caliber of the shunt avoids the deleterious effects of portal blood flow deprivation on hepatic function. Small-diameter portosystemic shunts have been reported to reduce the incidence of encephalopathy but at the expense of increased risks of shunt thrombosis and rebleeding.

The surgical shunt currently used most often is the distal splenorenal or Warren shunt (Fig. 31-15). This shunt is technically the most difficult to perform. It requires division of the gastroesophageal collaterals and allows venous drainage of the stomach and lower esophagus through the short gastrospenic veins into the spleen, and ultimately decompresses the left upper quadrant by allowing the splenic vein to drain directly into the left renal vein via an end-to-side splenic to left renal vein anastomosis. This shunt has the advantages of being associated with a lower rate of hepatic encephalopathy and decompensation and not interfering with subsequent liver transplantation.

Nonshunt Surgical Management of Refractory Variceal Bleeding

In the patient with extrahepatic portal vein thrombosis and refractory variceal bleeding, the Sugiura procedure may be considered. The Sugiura procedure consists of extensive devascularization of the stomach and distal esophagus along with transaction of the esophagus, splenectomy, truncal vagotomy, and pyloroplasty. As with performance of surgical shunts, patient survival is dependent on hepatic reserve at the time of the surgical procedure. Experience in Western countries is somewhat limited, and a number of modifications have been made to the original Sugiura procedure over time.

Hepatic Transplantation

Patients with cirrhosis, portal hypertension, and variceal bleeding usually die as a result of hepatic failure and not acute blood loss. Therefore, hepatic transplantation must be considered in the patient with ESLD because it represents the patient’s only chance for definitive therapy and long-term survival. Hepatic transplantation also can be considered for the patient with variceal bleeding refractory to all other forms of management. Survival after hepatic transplantation is not affected adversely by the previous performance of endoscopic variceal band ligation, TIPS, or splenorenal or mesocaval shunts. Previous creation of an Eck fistula, however, does make hepatic transplantation much more technically difficult, and therefore this procedure should be avoided in the transplant candidate. In addition to saving the patient’s life, hepatic transplantation reverses most of the hemodynamic and humoral changes associated with cirrhosis.

Budd-Chiari Syndrome

Budd-Chiari syndrome (BCS) is an uncommon congestive hepatopathy characterized by the obstruction of hepatic venous outflow. The incidence of BCS is 1 in 100,000 of the general population worldwide. Patients may present with acute signs and symptoms of abdominal pain, ascites, and hepatomegaly or more chronic symptoms related to long-standing portal hypertension. BCS is defined as primary when the obstructive process involves an endoluminal venous thrombosis. BCS is considered as a secondary process when the veins are compressed or invaded by a neighboring lesion originating outside the vein.

A thorough evaluation demonstrates one or more thrombotic risk factors in approximately 75% to 90% of patients with primary BCS. Primary myeloproliferative disorders, such as...
essential thrombocythemia or polycythemia rubra, account for approximately 35% to 50% of the primary cases of BCS. All known inherited thrombophilias also have been implicated in the development of BCS. Activated protein C resistance, generally related to factor V Leiden mutation, is present in approximately 25% of patients. Anticardiolipin antibodies, hyperhomocysteinemia, and oral contraceptive use all have been shown to be risk factors for BCS.

Clinically significant BCS is usually the result of obstruction of two or more of the major hepatic veins. The obstruction results in hepatomegaly, liver congestion, and right upper quadrant pain. In addition, liver perfusion via the portal vein may be decreased, and 70% of affected patients have noninflammatory centrilobular necrosis on biopsy. Although ALF is rare, most patients will go on to develop chronic portal hypertension and ascites. Caudate lobe hypertrophy occurs in approximately 50% of cases and is due to the fact that the caudate lobe has direct venous drainage into the IVC. This caudate lobe hypertrophy, in turn, can result in further obstruction of the IVC.

Abdominal ultrasonography is the initial investigation of choice and can demonstrate the absence of hepatic vein flow, spider web hepatic veins, and collateral hepatic veins. CT or MRI of the abdomen also is capable of demonstrating hepatic vein thrombosis and evaluating the IVC but is limited in that it cannot show direction of blood flow. The definitive radiographic study to evaluate BCS is hepatic venography to determine the presence and extent of hepatic vein thrombus as well as measure IVC pressures.

Initial treatment consists of diagnosing and medically managing the underlying disease process and preventing extension of the hepatic vein thrombosis through systemic anticoagulation. The BCS-associated portal hypertension and ascites can be medically managed in a manner similar to that in most cirrhotic patients. Radiologic and surgical intervention should be reserved for patients whose condition is nonresponsive to medical therapy. Percutaneous angioplasty and TIPS, in combination with thrombolytic therapy, are the preferred strategies to restore the outflow of blood from the liver. Thrombolytic therapy alone may be attempted for acute thrombosis. Surgical shunting, namely with the side-to-side portacaval shunt, essentially turns the portal vein into a hepatic outflow tract. Most patients with a portacaval shunt show improvement in hepatic function and fibrosis at 1 year without significant hepatic encephalopathy. However, the enthusiasm for this procedure has been curbed due to the relatively high rate of operative mortality and shunt dysfunction. Patients with progressive BCS and manifestations of ESLD will ultimately require hepatic transplantation.

INFECTIONS OF THE LIVER

The liver contains the largest portion of the reticuloendothelial system in the human body and is therefore able to handle the continuous low-level exposure to enteric bacteria that it receives through the portal venous system. Due to the high level of reticuloendothelial cells in the liver, nonviral infections are unusual.

Pyogenic Liver Abscesses

Pyogenic liver abscesses are the most common liver abscesses seen in the United States. They may be single or multiple and are more frequently found in the right lobe of the liver. The abscess cavities are variable in size and, when multiple, may coalesce to give a honeycomb appearance. Approximately 40% of abscesses are monomicrobial, an additional 40% are polymicrobial, and 20% are culture-negative. The most common infecting agents are gram-negative bacteria. *Escherichia coli* is found in two thirds of cases, and other common organisms include *Streptococcus faecalis*, *Klebsiella*, and *Proteus vulgaris*. Anaerobic organisms such as *Bacteroides fragilis* also are seen frequently. In patients with endocarditis and infected indwelling catheters, *Staphylococcus* and *Streptococcus* species are more commonly found.

In the past, pyogenic liver abscesses often resulted from infections of the intestinal tract such as acute appendicitis and diverticulitis, which then spread to the liver via the portal circulation. With improved imaging modalities and earlier diagnosis of these intra-abdominal infections, this particular etiology of pyogenic liver abscesses has become less common. Pyogenic liver abscesses also occur as a result of impaired biliary drainage, subacute bacterial endocarditis, infected indwelling catheters, dental work, or the direct extension of infections such as diverticulitis or Crohn’s disease into the liver. There appears to be an increasing incidence due to infection by opportunistic organisms among immunosuppressed patients, including transplant and chemotherapy recipients as well as patients with acquired immunodeficiency syndrome (AIDS).

Patients commonly present with right upper quadrant pain and fever. Jaundice occurs in up to one-third of affected patients. A thorough history and physical examination are usually helpful in identifying the underlying cause of the liver abscess. Leukocytosis, an elevated sedimentation rate, and an elevated AP level are the most common laboratory findings. Significant abnormalities in the results of the remaining liver function tests are unusual. Blood cultures will only reveal the causative organism in approximately 50% of cases. Ultrasound examination of the liver reveals pyogenic abscesses as round or oval hypoechoic lesions with well-defined borders and a variable number of internal echoes. CT scan is highly sensitive in the localization of pyogenic liver abscesses, which appear hypodense with peripheral enhancement and may contain air-fluid levels indicating gas-producing infectious organisms (Fig. 31-16). MRI of the abdomen can also detect pyogenic abscesses with a high level of sensitivity but plays a limited role because of its inability to be used for image-guided diagnosis and therapy.

The current cornerstones of treatment include correction of the underlying cause and IV antibiotic therapy. Empiric antibiotic therapy should cover gram-negative and anaerobic organisms; percutaneous needle aspiration and culture of the aspirate may be useful in guiding subsequent antibiotic therapy. IV antibiotic therapy should be continued for at least 8 weeks and can be expected to be effective in 80% to 90% of patients. Placement of a percutaneous drainage catheter is beneficial only for a minority of patients, as most pyogenic abscesses are quite viscous and catheter drainage is often ineffective.

Surgical drainage either via the laparoscopic or open approach may become necessary if initial therapies fail. Anatomic surgical resection can be performed in patients with recalcitrant abscesses. It must be kept in mind throughout the evaluation and treatment of the presumed pyogenic abscess that a necrotic hepatic malignancy must not be mistaken for a hepatic abscess. Therefore, early diagnosis and progression to surgical resection should be advocated for patients who do not respond to initial antibiotic therapy.

Amecic Abscesses

*Entamoeba histolytica* is a parasite that is endemic worldwide, infecting approximately 10% of the world’s population.
Amebiasis is most common in subtropical climates, especially in areas with poor sanitation. *E. histolytica* exists as cysts in a vegetative form that are capable of surviving outside the human body. The cystic form passes through the stomach and small bowel unharmed and then transforms into a trophozoite in the colon. Here it invades the colonic mucosa forming typical flask-shaped ulcers, enters the portal venous system, and is carried to the liver. Occasionally, the trophozoite will pass through the hepatic sinusoid and into the systemic circulation, which results in lung and brain abscesses.

Amebics multiply and block small intrahepatic portal radicles with consequent focal infarction of hepatocytes. They contain a proteolytic enzyme that also destroys liver parenchyma. The abscesses formed are variable in size and can be single or multiple. The amebic abscess is most commonly located in the superior-anterior aspect of the right lobe of the liver near the diaphragm and has a necrotic central portion that contains a thick, reddish brown, pus-like material. This material has been likened to anchovy paste or chocolate sauce. Amebic abscesses are the most common type of liver abscesses worldwide.

Amebiasis should be considered in patients who have traveled to an endemic area and present with right upper quadrant pain, fever, hepatomegaly, and hepatic abscess. Leukocytosis is common, whereas elevated transaminase levels and jaundice are unusual. The most common biochemical abnormality is a mildly elevated AP level. Even though this disease process is secondary to a colonic infection, the presence of diarrhea is unusual. Most patients have a positive fluorescent antibody test for *E. histolytica*, and test results can remain positive for some time after a clinical cure. This serologic test has a high sensitivity, and therefore amebiasis is unlikely if the test results are negative.

Ultrasound and CT scanning of the abdomen are both very sensitive but nonspecific for the detection of amebic abscesses. Amebic abscesses usually appear on CT as well-defined low-density round lesions that have enhancement of the wall, somewhat ragged in appearance with a peripheral zone of edema. The central cavity may have septations as well as fluid levels. CT scanning is also useful in the detection of extrahepatic involvement.

Hydatid Disease

Hydatid disease is due to infection by the tapeworm *Echinococcus granulosus* in its larval or cyst stage. The tapeworm lives in canids, which are infected by eating the viscera of sheep that contain hydatid cysts. Scolecis, contained in the cysts, adhere to the small intestine of dogs and become adult taenias, which attach to the intestinal wall. Each worm sheds approximately 500 ova into the bowel. The infected ova-containing feces of dogs contaminate grass and farmland, and the ova are ingested by intermediate hosts such as sheep, cattle, pigs, and humans. The ova have chitinious envelopes that are dissolved by gastric juice. The liberated ovum then burrows through the intestinal mucosa and is carried by the portal vein to the liver, where it develops into an adult cyst. Most cysts are caught in the hepatic sinusoids, and therefore 70% of hydatid cysts form in the liver. A few ova pass through the liver and are held up in the pulmonary capillary bed or enter the systemic circulation, forming cysts in the lung, spleen, brain, or bones.

Hydatid disease is most common in sheep-raising areas, where dogs have access to infected offal. These include South Australia, New Zealand, Africa, Greece, Spain, and the Middle East. Hydatid cysts commonly involve the right lobe of the liver, usually the anterior-inferior or posterior-inferior segments. The uncomplicated cyst may be silent and found only incidentally or at autopsy. Occasionally, the affected patient presents with symptoms such as dull right upper quadrant pain or abdominal distention. Cysts may become secondarily infected, involve other organs, or even rupture, which leads to an allergic or anaphylactic reaction.

The diagnosis of hydatid disease is based on the findings of an enzyme-linked immunosorbent assay (ELISA) for echinococcal antigens, and results are positive in approximately 85% of infected patients. The ELISA results may be negative in an infected patient if the cyst has not leaked or does not contain scolecis or if the parasite is no longer viable. Eosinophilia is seen in approximately 30% of infected patients. Ultrasonography and CT scanning of the abdomen are both quite sensitive for detecting hydatid cysts. The appearance of the cysts on images depends on the stage of cyst development. Typically, hydatid cysts are well-defined hypodense lesions with a distinct wall. Ring-like calcifications of the pericysts are present in 20% to 30% of cases. As healing occurs, the entire cyst calcifies densely, and a lesion with this appearance is usually dead or inactive. Daughter cysts generally occur in a peripheral location within the main cyst and are typically slightly hypodense compared with the mother cyst. MRI of the abdomen may be useful to evaluate the pericyst, cyst matrix, and daughter cyst characteristics.

Unless the cysts are small or the patient is not a suitable candidate for surgical resection, the treatment of hydatid disease...
is surgically based because of the high risk of secondary infection and rupture. Medical treatment with albendazole relies on drug diffusion through the cyst membrane. The concentration of drug achieved in the cyst is uncertain but is better than that of mebendazole, and albendazole can be used as initial treatment for small, asymptomatic cysts. For most cysts, surgical resection involving laparoscopic or open complete cyst removal with instillation of a sclerosing agent is preferred and usually is curative. If complete cystectomy is not possible, then formal anatomic liver resection can be undertaken. During surgical resection, caution must be exercised to avoid rupture of the cyst with release of protoscolices into the peritoneal cavity. Peritoneal contamination can result in an acute anaphylactic reaction or peritoneal implantation of sclerites with daughter cyst formation and inevitable recurrence.

_Echinococcus multilocularis_ occurs in the Northern Hemisphere and can infect the liver in a fashion similar to that described earlier, although the cysts are multilocular. Infection of the lung also is common (alveolar echinococcosis). Canine species such as wolves, foxes, and dogs ingest infected viscera of an intermediate host (e.g., rodents, moose) and become infected; humans become infected incidentally by ingesting contaminated food or water. Treatment consists of albendazole; however, infection in the lung produces a more generalized granulomatous reaction, can present in a manner similar to that of a malignancy, and often requires resection.

**Ascariasis**

Ascaris infection is particularly common in the Far East, India, and South Africa. Ova of the roundworm *Ascaris lumbricoides* arrive in the liver by retrograde locomotion in the bile ducts from the GI tract. The adult worm is 10 to 20 cm long and may lodge in the common bile duct, causing partial bile duct obstruction and secondary cholangitic abscesses. The _Ascaris_ may serve as a nidus for the development of intrahepatic gallstones. The clinical presentation in an affected patient may include biliary colic, acute cholecystitis, acute pancreatitis, or hepatic abscesses. Plain abdominal radiographs, abdominal ultrasound, and ERCP can demonstrate the _Ascaris_ as linear filling defects within the bile ducts. Occasionally, worms can be seen moving into and out of the biliary tree from the duodenum. Treatment consists of the administration of piperazine citrate, mebendazole, or albendazole in combination with endoscopic extraction of the worms. Surgical intervention may become necessary if the _Ascaris_ cannot be removed via ERCP.

**Schistosomiasis**

Schistosomiasis affects >200 million people in 74 countries. Hepatic schistosomiasis occurs when emboli of the ova in the intestines reach the liver via the mesenteric venous system. Eggs excreted in the feces hatch in water to release free-swimming embryos, which enter snails and develop into fork-tailed cercariae. They then reenter human skin during contact with infected water. They burrow down to the capillary bed and enter the bloodstream, leading to widespread hematogenous dissemination. Those entering the intrahepatic portal system grow rapidly, resulting in a granulomatous reaction. The degree of consequent portal fibrosis is related to the adult worm load.

Schistosomiasis has three stages of clinical symptomatology: the first includes itching after the entry of cercariae through the skin; the second includes fever, urticaria, and eosinophilia; and the third involves hepatic fibrosis followed by presinusoidal portal hypertension. During this third phase, the liver shrinks, the spleen enlarges, and the patient may develop complications of portal hypertension while hepatic function is maintained. Active infection is detected by stool examination. Serologic tests indicate past exposure but do not provide information regarding the timing of infection. A negative serologic test result excludes the presence of schistosomal infection. Serum levels of transaminases are usually normal, but the AP level may be mildly elevated. A decreased serum albumin level is usually the result of frequent GI bleeds and malnutrition.

Medical treatment of schistosomiasis includes education on hygiene and the avoidance of infected water. Treatment with praziquantel 40 to 75 mg/kg as a single dose is the treatment of choice for all forms of schistosomiasis and produces few side effects. GI bleeding usually is controlled by endoscopic variceal ligation. However, in a patient with refractory GI portal hypertensive bleeding, distal splenorenal shunt or gastric devascularization and splenectomy may be considered.

**Viral Hepatitis**

The role of the surgeon in the management of viral hepatitis is somewhat limited. However, the disease entities of hepatitis A, B, and C need to be kept in mind during any evaluation for liver disease. Hepatitis A usually results in an acute self-limited illness and only rarely leads to fulminant hepatic failure. Patients can present with fatigue, malaise, nausea, vomiting, anorexia, fever, and right upper quadrant abdominal pain. The most common physical findings are jaundice and hepatomegaly. Because the disease is self-limited, the treatment is usually supportive. Patients who develop fulminant infection require aggressive therapy and should be transferred to a center capable of performing liver transplantation.

Hepatitis B and C, on the other hand, can both lead to chronic liver disease, cirrhosis, and HCC. The prevalence of chronic hepatitis B infection in the U.S. population is estimated to be 0.27%, but hepatitis B remains a major burden in resource-limited countries, accounting for 30% of cirrhosis and 53% of HCC cases. The ultimate goal of treatment for chronic hepatitis B is viral suppression, thereby preventing the development of clinical outcomes such as cirrhosis, liver failure, and HCC. The cornerstone of current antiviral therapy includes pegylated interferon and nucleoside analogs such as tenofovir or entecavir. These agents have been proven to reduce complications of cirrhosis and HCC and perhaps reverse previous damage to the liver. Interferon therapy produces various side effects including fatigue, flu-like symptoms, mood changes, bone marrow suppression, and stimulation of autoimmunity. On the other hand, the nucleoside analogs are generally well-tolerated by patients. Compared with lamivudine, the nucleoside analogs are less likely to produce resistance and are more likely to be clinically effective. Despite its high rate of viral resistance, lamivudine may be the preferred treatment in some countries because of its relatively low cost.

Acute hepatitis C viral (HCV) infection typically develops 2 to 26 weeks after exposure to the virus, and presenting symptoms can include jaundice, nausea, dark urine, and right upper quadrant abdominal pain. Diagnosis is confirmed by testing for the presence of HCV RNA and anti-HCV antibodies in the serum; viral RNA is first detectable in the serum by polymerase chain reaction (PCR) within days to weeks following the exposure, whereas antibodies will not appear until 2 to 6 months after. If the diagnosis of acute hepatitis is established and the virus is not spontaneously cleared within 12 weeks, patients should generally be treated with pegylated interferon
monotherapy. High rates of sustained viral response, in excess of 80%, have been achieved with the early treatment of acute HCV infection. Unfortunately, patients with acute HCV infection are typically asymptomatic, and the vast majority of cases go undetected. Untreated, most of these patients will eventually develop chronic infection.

Chronic HCV infection often follows a progressive course over many years and can ultimately result in cirrhosis, HCC, and the need for liver transplantation. Cirrhosis secondary to hepatitis C remains the leading indication for liver transplantation in the United States, Europe, and Japan. The decision to treat a patient with chronic HCV infection is complex and involves consideration of multiple factors including the natural history of the disease, stage of fibrosis, and the efficacy and adverse effects related to the treatment regimen. Patients with genotype 1 are now treated with triple-agent therapy including pegylated interferon, ribavirin, and a protease inhibitor. The protease inhibitors telaprevir and boceprevir were recently approved for the treatment of chronic HCV genotype 1 infection, and their addition to the treatment regimen has been shown to increase the rate of sustained viral response from 40% to 70%. These protease inhibitors do not exhibit significant antiviral activity against other HCV genotypes, and therefore genotypes 2, 3, and 4 are treated with interferon and ribavirin alone. Genotypes 2 and 3 are generally more responsive to treatment than genotypes 1 and 4, but the therapeutic response also is dependent on other factors such as baseline viral load, ethnicity, and the patient’s genetic background and compliance with the regimen.

**EVALUATION OF AN INCIDENTAL LIVER MASS**

A liver mass often is identified incidentally during a radiologic imaging procedure performed for another indication. For example, a liver mass may be discovered during evaluation for gallbladder disease or kidney stones. In addition, with advances in imaging technology, previously undetected lesions not infrequently are now identified. Although many of these lesions are benign and will require no further treatment, the concern for malignancy requires a thorough evaluation. Thus, an orderly approach should be taken to the workup of an incidental liver lesion to minimize unnecessary testing.

The evaluation of an incidental liver mass begins with a history and physical examination (Fig. 31-17). The patient should be asked about abdominal pain, weight loss, previous

---

**Figure 31-17.** Algorithm for diagnostic workup of an incidental liver lesion. The evaluation includes history and physical examination, blood work, imaging studies, and liver biopsy (if needed). AFP = α-fetoprotein; BUN = blood urea nitrogen; CA 19-9 = cancer antigen 19-9; CEA = carcinoembryonic antigen; creat = creatinine; CT = computed tomography; EGD = esophagogastroduodenoscopy; glu = glucose; Gyn = gynecologic; HTN = hypertension; MRI = magnetic resonance imaging; OCP = oral contraceptive pill; PAP = Papanicolaou; US = ultrasound.
liver disease, cirrhosis, alcohol use, viral hepatitis, blood transfusions, tattoos, oral contraceptive use (in women), and personal or family history of cancer. On physical examination, jaundice, scleral icterus, hepatomegaly, splenomegaly, palpable mass, or stigmata of portal hypertension should be noted. After completion of the history and physical examination, blood work should be performed, including complete blood count; platelet count; measurement of levels of electrolytes, blood urea nitrogen, creatinine, glucose, and albumin; liver function tests; serum ammonia level; coagulation studies; hepatitis screen; and measurement of levels of the tumor markers carcinoembryonic antigen, AFP, and cancer antigen 19-9.

The differential diagnosis for an incidental liver mass includes cysts, benign solid lesions, and primary or metastatic cancers (Table 31-6). Ultrasound or CT is commonly performed to evaluate respiratory or abdominal symptoms, and these imaging studies usually lead to the discovery of an incidental liver lesion. Further radiologic evaluation by dual- or triple-phase CT scan or MRI is often necessary to fully characterize the extent and nature of the lesion. Different types of liver masses have distinct appearances and patterns of contrast enhancement on these studies, facilitating the clinician in the diagnosis and formulation of the treatment plan. The use of liver-specific contrast agents in MRI provides information on hepatocyte function in combination with the structural data obtained during a standard MRI, and yields improved detection and characterization of liver lesions. CT cholangiography or MRCP can be obtained, particularly when visualization of the biliary tract is desired. These modalities may be useful in the depiction of benign or malignant strictures resulting in biliary obstruction. In the evaluation of metastases to the liver from a variety of primary cancers, FDG-PET/CT has emerged as an indispensable tool in disease staging and in the follow-up after treatment. The techniques available for imaging of liver lesions are described in detail earlier in the section “Radiologic Evaluation of the Liver.”

Liver biopsy is indicated when biochemical analysis and diagnostic imaging fail to lead to a definitive diagnosis. Percutaneous liver biopsy with ultrasound or CT guidance is the simplest, fastest, and most commonly performed approach to obtain hepatic tissue for histologic examination. Absolute contraindications to percutaneous liver biopsy include significant coagulopathy (as in patients with decompensated cirrhosis), biliary dilatation, and suspicion for hemangioma or echinococcal cyst. Obesity and the presence of ascites are relative contraindications that can present a challenge to the percutaneous approach. In these patients, laparoscopic liver biopsy can be considered. The laparoscopic technique is likely to have a higher diagnostic yield in cirrhotic patients with ascites and/or coagulopathy, in whom bleeding risk is excessive by the percutaneous route. Laparoscopy also offers the opportunity to stage the extent of disease in patients with various intra-abdominal malignancies.

### HEPATIC CYSTS

#### Congenital Cysts

The majority of hepatic cysts are asymptomatic. Hepatic cysts are usually identified incidentally and can occur at any time throughout life. The most common benign lesion found in the liver is the congenital or simple cyst. The exact prevalence of simple hepatic cysts in the U.S. population is not known, but the female to male ratio is approximately 4:1, and the prevalence is approximately 2.8% to 3.6%.

Simple cysts are the result of excluded hyperplastic bile duct rests. Simple cysts usually are identified in hepatic imaging studies as thin-walled, homogeneous, fluid-filled structures with few to no septations. The cyst epithelium is cuboidal and secretes a clear nonbilious serous fluid. With the exception of large cysts, simple cysts are usually asymptomatic. Large simple cysts may cause abdominal pain, epigastric fullness, and early satiety. Occasionally the affected patient presents with an abdominal mass.

Asymptomatic simple cysts are best managed conservatively. The preferred treatment for symptomatic cysts is ultrasound- or CT-guided percutaneous cyst aspiration followed by sclerotherapy. This approach is approximately 90% effective in controlling symptoms and ablatting the cyst cavity. If percutaneous treatment is unavailable or ineffective, treatment may include either laparoscopic or open surgical cyst fenestration. The laparoscopic approach is being used more frequently and is 90% effective. The excised cyst wall is sent for pathologic analysis to exclude the presence of carcinoma, and the remaining cyst wall must be carefully inspected for evidence of neoplastic change. If such change is present, complete resection is required, either by enucleation or formal hepatic resection.

#### Biliary Cystadenoma

Biliary cystadenomas are slow-growing, unusual, benign lesions that most commonly present as large lesions in the right lobe of the liver. Although these lesions are usually benign, they can undergo malignant transformation. Patients with biliary cystadenomas commonly present with abdominal pain. An abdominal mass occasionally can be identified on physical examination. In contrast to simple cysts, biliary cystadenomas have walls that appear thicker with soft tissue nodules and septations that usually enhance. The protein content of the fluid can be variable and can affect the radiographic images on CT and MRI. Surgical resection is the preferred mode of treatment.

#### Polycystic Liver Disease

Adult polycystic liver disease (PCLD) occurs as an autosomal dominant disease and usually presents in the third decade of life. Approximately 44% to 76% of affected families are found to have mutations of *PKD1*, and approximately 75% have mutations of *PKD2*. The prevalence and number of hepatic cysts
Speciﬁc Considerations

PART II

Liver cysts are the most frequently encountered liver lesion overall and are described in detail in the section “Hepatic Cysts.” Cystic lesions of the liver can arise primarily (congenital, single or multiple, cell of origin (hepatocellular, cholangiocellular, or mesenchymal), and benign or malignant. Benign liver cysts are cysts, hemangiomas, focal nodular hyperplasia (FNH), and hepatocellular adenomas (see Table 31-6). Many of these lesions have typical features in imaging studies that help conﬁrm the diagnosis.

Cyst

Hepatic cysts are the most frequently encountered liver lesion and are described in detail in the section “Hepatic Cysts.” Cystic lesions of the liver can arise primarily (congenital) or secondarily from trauma (seroma or biloma), infection (pyogenic or parasitic), or neoplastic disease. Congenital cysts are usually simple cysts containing thin serous fluid and are
reported to occur in 5% to 14% of the population, with higher prevalence in women. In most cases, congenital cysts are differentiated from secondary cysts (infectious or neoplastic origin) in that they have a well-defined thin wall and no solid component and are filled with homogeneous, clear fluid. For benign solid liver lesions, the differential diagnosis includes hemangioma, adenoma, FNH, and bile duct hamartoma.

**Hemangioma**

Hemangiomas (also referred to as hemangiomata) are the most common solid benign masses that occur in the liver. They consist of large endothelial-lined vascular spaces and represent congenital vascular lesions that contain fibrous tissue and small blood vessels that eventually grow. They are predominantly seen in women and occur in 2% to 20% of the population. They can range from small (≤1 cm) to giant cavernous hemangiomas (10 to 25 cm). Most hemangiomas are discovered incidentally with little clinical consequence. However, large lesions can cause symptoms as a result of compression of adjacent organs or intermittent thrombosis, which in turn results in further expansion of the lesion. Spontaneous rupture (bleeding) is rare, but surgical resection can be considered if the patient is symptomatic. Resection can be accomplished by enucleation or formal hepatic resection, depending on the location and involvement of intrahepatic vascular structures and hepatic ducts.

The majority of hemangiomas can be diagnosed by liver imaging studies. On biphasic contrast CT scan, large hemangiomas show asymmetrical nodular peripheral enhancement that is isodense with large vessels and exhibit progressive centripetal enhancement fill-in over time (Fig. 31-18). On MRI, hemangiomas are hypointense on T1-weighted images and hyperintense on T2-weighted images. With gadolinium enhancement, hemangiomas show a pattern of peripheral nodular enhancement similar to that seen on contrast CT scans. Caution should be exercised in ordering a liver biopsy if the suspected diagnosis is hemangioma because of the risk of bleeding from the biopsy site, especially if the lesion is at the edge of the liver.

**Adenoma**

Hepatic adenomas are benign solid neoplasms of the liver. They are most commonly seen in premenopausal women older than 30 years of age and are typically solitary, although multiple adenomas also can occur. Prior or current use of estrogens (oral contraceptives) is a clear risk factor for development of liver adenomas, although they can occur even in the absence of oral contraceptive use. On gross examination, they appear soft and encapsulated and are tan to light brown. Histologically, adenomas lack bile duct glands and Kupffer cells, have no true lobules, and contain hepatocytes that appear congested or vacuolated due to glycogen deposition. On CT scan, adenomas usually have

---

**Figure 31-18.** Computed tomographic scans showing classic appearance of benign liver lesions. Focal nodular hyperplasia (FNH) is hypervascular on arterial phase, isodense to liver on venous phase, and has a central scar (upper panels). Adenoma is hypovascular (lower left panel). Hemangioma shows asymmetrical peripheral enhancement (lower right panel).
sharply defined borders and can be confused with metastatic tumors. With venous phase contrast, they can look hypodense or isodense in comparison with background liver, whereas on arterial phase contrast, subtle hypervascular enhancement often is seen (see Fig. 31-18). On MRI scans, adenomas are hyperintense on T1-weighted images and enhance early after gadolinium injection. With the use of liver-specific MRI contrast agents such as gadoxetate (Eovist or Primovist, Bayer-Schering, Berlin, Germany), hepatic adenomas can be better distinguished from FNH by their enhancement characteristics during the hepaticobiliary phase of imaging. The new MRI contrast agent, gadobenate dimeglumine (MultiHance, Bracco Diagnostics, Milan, Italy), is eliminated through both renal and biliary excretion. Therefore, liver lesions that contain hepatocytes with intact biliary excretion mechanism will take up this contrast agent and be easily distinguished from lesions that do not. This contrast agent has improved our ability to differentiate hepatic adenoma from FNH with a high degree of accuracy.

Hepatic adenomas carry a significant risk of spontaneous rupture with intraperitoneal bleeding. The clinical presentation may be abdominal pain, and in 10% to 25% of cases, hepatic adenomas present with spontaneous intraperitoneal hemorrhage. Hepatic adenomas also have a risk of malignant transformation to a well-differentiated HCC. Therefore, it usually is recommended that large hepatic adenomas (>4–5 cm) be surgically resected.

**Focal Nodular Hyperplasia**

FNH is a solid, benign lesion of the liver believed to be a hyperplastic response to an anomalous artery. Similar to adenomas, they are more common in women of childbearing age, although the link to oral contraceptive use is not as clear as with adenomas. A good-quality biphasic CT scan usually is diagnostic of FNH, on which such lesions appear well circumscribed with a typical central scar (see Fig. 31-18). They show intense homogeneous enhancement on arterial phase contrast images and are often isodense or invisible compared with background liver on the venous phase. On MRI scans, FNH lesions are hypointense on T1-weighted images and isointense to hyperintense on T2-weighted images. After gadolinium administration, lesions are hyperintense but become isointense on delayed images. The fibrous septa extending from the central scar are also more readily seen with MRI. Unlike adenomas, FNH lesions usually do not rupture spontaneously and have no significant risk of malignant transformation. Therefore, the management of FNH is usually reassurance and prospective observation irrespective of size. Surgical resection can be recommended, however, when patients are symptomatic or when hepatic adenoma or HCC cannot be definitively excluded. Oral contraceptive or estrogen use should be stopped when either FNH or adenoma is diagnosed.

**Bile Duct Hamartoma**

Bile duct hamartomas are typically small liver lesions, 2 to 4 mm in size, visualized on the surface of the liver at laparotomy. They are firm, smooth, and whitish yellow in appearance. They can be difficult to differentiate from small metastatic lesions, and excisional biopsy often is required to establish the diagnosis.

**MALIGNANT LIVER TUMORS**

Malignant tumors in the liver can be classified as primary (cancers that originate in the liver) or metastatic (cancers that spread to the liver from an extrahepatic primary site) (see Table 31-6). Primary cancers in the liver that originate from hepatocytes are known as hepatocellular carcinomas (HCCs or hepatomas), whereas cancers arising in the bile ducts are known as cholangiocarcinomas.

In the United States, approximately 150,000 new cases of colorectal cancer are diagnosed each year, and the majority of patients (approximately 60%) will develop hepatic metastases over their lifetime. Hence, the most common tumor seen in the liver is metastatic colorectal cancer. This compares with approximately 30,000 new cases of HCC diagnosed annually in the United States. Interestingly, in a Western series of 1000 consecutive new liver cancer patients seen at a university medical center, 47% had HCC, 17% had colorectal cancer metastases, 11% had cholangiocarcinomas, 7% had neuroendocrine metastases, and 18% had other tumors.77 Although these figures do not reflect the incidence or prevalence of these liver cancers, they are indicative of referral patterns in a tertiary academic medical center with a large liver transplantation team and active hepatology clinic.

**Hepatocellular Carcinoma**

HCC is the fifth most common malignancy worldwide, with an estimated 750,000 new cases diagnosed annually. Because of its high fatality, it is the third most common cause of cancer death worldwide.78 Major risk factors are viral hepatitis (B or C), alcoholic cirrhosis, hemochromatosis, and NASH. In Asia, the risk is as high as 35 to 117 per 100,000 persons per year, whereas in the United States, the risk is only 7 per 100,000 persons per year.79 Although cirrhosis is not present in all cases, it has been estimated to be present 70% to 90% of the time. In a person with cirrhosis, the annual conversion rate to HCC is 2% to 6%.79 In patients with chronic HCV infection, cirrhosis usually is present before the HCC develops; however, in cases of hepatitis B virus infection, HCC tumors can occur before the onset of cirrhosis. HCCs are typically hypervascular with blood supplied predominantly from the hepatic artery. Thus, the lesion often appears hypervascular during the arterial phase of CT studies (Fig. 31-19) and relatively hypodense during the delayed phases due to early washout of the contrast medium by the arterial blood. MRI imaging also is effective in characterizing HCC. HCC is variable on T1-weighted images and usually hyperintense on T2-weighted images. As with contrast CT, HCC enhances in the arterial phase after gadolinium injection because of its hypervascularity and becomes hypointense in the delayed phases due to contrast washout. HCC has a tendency to invade the portal vein, and the presence of an enhancing portal vein thrombus is highly suggestive of HCC.

The treatment of HCC is complex and is best managed by a multidisciplinary liver transplant team. A complete algorithm for the evaluation and management of HCC is shown in Fig. 31-20. For patients without cirrhosis who develop HCC, resection is the treatment of choice. For patients with Child’s class A cirrhosis with preserved liver function and no portal hypertension, resection also is considered. If resection is not possible because of poor liver function and the HCC meets transplant criteria (discussed later), liver transplantation is the treatment of choice.104,101 The Barcelona-Clinic Liver Cancer Group has refined its HCC management strategy and has developed the American Association for the Study of Liver Diseases Practice Guidelines.82 Management guidelines vary slightly in Asia, Europe, the
United States, and other countries based in part on availability of organ donors for liver transplantation. Living donor liver transplantation also is an alternative for patients with HCC awaiting transplantation to avoid dropout as a candidate for cadaveric donor liver transplantation due to tumor progression. Specific treatment options are described in the next section.

**Cholangiocarcinoma**

Cholangiocarcinoma, or bile duct cancer, is the second most common primary malignancy of the liver. Cholangiocarcinoma is an adenocarcinoma of the bile ducts; it forms in the biliary epithelial cells and can be subclassified into peripheral (intrahepatic) bile duct cancer and central (extrahepatic) bile duct cancer. Extrahepatic bile duct cancer can be located distally or proximally. When proximal, it is referred to as a hilar cholangiocarcinoma (Klatskin’s tumor). Hilar cholangiocarcinoma originates in the wall of the bile duct at the hepatic duct confluence and usually presents with obstructive jaundice rather than an actual liver mass. In contrast, a peripheral (or intrahepatic) cholangiocarcinoma represents a tumor mass within a hepatic lobe or at the periphery of the liver. A biopsy specimen from the cholangiocarcinoma will show adenocarcinoma, but pathologists are often unable to differentiate metastatic adenocarcinoma to the liver from primary bile duct adenocarcinoma. Therefore, a search for a primary site should be undertaken in cases in which an incidentally discovered liver lesion is proven to be an adenocarcinoma on biopsy.

Hilar cholangiocarcinoma is difficult to diagnose and typically presents as a stricture of the proximal hepatic duct causing painless jaundice. It preferentially grows along the length of the bile ducts, often involving the periductal lymphatics with frequent lymph node metastases. Surgical resection offers the only chance for cure of cholangiocarcinoma. The location and extent of tumor dictate the operative approach. In one series of 225 patients with hilar cholangiocarcinoma, 65 (29%) were deemed to have unresectable tumors by initial imaging. Of the remaining 160 patients who underwent exploratory surgery with curative intent, 80 (50%) were found to have inoperable tumors. Histologically negative margins, concomitant hepatic resection, and well-differentiated tumor histology were associated with improved outcome after resection. In another series of 61 patients undergoing surgical exploration for hilar cholangiocarcinoma, the 5-year actuarial survival rates for an R0 or R1 resection were 45% and 26%, respectively. In a large series reported by Nagino and colleagues, 132 patients with hilar cholangiocarcinoma underwent extended heptectomy with resection of the caudate lobe and extrahepatic bile duct, and/or portal vein resection (n = 63) after portal vein embolization. The 3- and 5-year survival rates were 41.7% and 26.8%, respectively.

In the absence of associated primary sclerosing cholangitis (PSC), surgical resection is the treatment of choice for hilar cholangiocarcinoma. However, approximately 10% of patients with cholangiocarcinoma have PSC. Furthermore, cholangiocarcinoma in the setting of PSC is frequently multicentric and often is associated with underlying liver disease, with eventual cirrhosis and portal hypertension. As a result, experience has shown that resection of cholangiocarcinoma in patients with PSC yields dismal results. This led transplant centers to consider OLT for patients with hilar cholangiocarcinoma. The initial results of transplantation were disappointing, however, with high recurrence and overall 3-year survival rates of <30%.

Because the growth of hilar cholangiocarcinoma indicates that this disease spreads in a locoregional manner, a rationale for the use of neoadjuvant chemoradiation was developed by the transplant team at the University of Nebraska in the late 1980s. This was adapted in 1993 by the transplant team at the Mayo Clinic, which led to the current Mayo Clinic protocol. The pretransplant Mayo protocol consists of external-beam radiation therapy plus a protracted course of intravenous 5-fluorouracil followed by iridium-192 brachytherapy. Patients then undergo an abdominal exploration with staging. If findings are negative, patients are given capecitabine for 2 of every 3 weeks until transplantation. Even after restaging with CT/MRI and endoscopic ultrasonography, approximately 15% to 20% of patients will have positive findings for tumor on abdominal
The 5-year survival rate for those undergoing transplantation for cholangiocarcinoma at the Mayo Clinic is approximately 70% and compares favorably with the rate for resection. Current eligibility criteria for this Mayo Clinic protocol include unresectable hilar cholangiocarcinoma or hilar cholangiocarcinoma with PSC. The tumor must have a radial dimension of ≤3 cm with no intrahepatic or extrahepatic metastases, and the patient must not have undergone prior radiation therapy or transperitoneal biopsy. Many centers have adopted similar protocols with comparable results.

Peripheral, or intrahepatic, cholangiocarcinoma is less common than hilar cholangiocarcinoma. In a series of 53 patients at Memorial Sloan-Kettering Cancer Center who underwent surgical exploration for a diagnosis of intrahepatic cholangiocarcinoma, 33 (62%) were found to have resectable tumors. Actuarial 3-year survival for patients undergoing resection was 55%. Factors predictive of poor survival included vascular invasion, histologically positive margins, and multiple tumors. In a large series in Taiwan, 373 patients with peripheral cholangiocarcinoma underwent surgical treatment from 1977 to 2001. Absence of mucobilia, nonpapillary tumor type, tumor of advanced stage, nonhepatectomy, and lack of postoperative chemotherapy were five independent prognostic factors that adversely affected overall survival. Liver transplantation has been performed for peripheral cholangiocarcinoma; however, currently all but one center in the United States have eschewed this approach because of organ shortages and relatively high recurrence rates.

**Gallbladder Cancer**

Gallbladder cancer is a rare aggressive tumor with a very poor prognosis. Over 90% of patients have associated cholelithiasis. In one study examining the mode of presentation over a 10-year period from 1990 to 2000 in 44 patients diagnosed with gallbladder cancer, the diagnosis was found to be made preoperatively in 57%, intraoperatively in 11%, and incidentally after cholecystectomy in 32%. Surgical approaches can be classified into (a) reoperation for an incidental finding of gallbladder cancer after cholecystectomy, and (b) radical resection in patients with advanced disease. The results are dismal for radical resection in patients with advanced disease and positive hilar lymph nodes. For incidental gallbladder cancer beyond stage T1, reoperation with central liver resection, hilar lymphadenectomy, and evaluation of cystic duct stump is most commonly performed. The role of formal lobectomy or extended lobectomy as well as common bile duct resection is more controversial. In a single-center study of 23 patients undergoing attempted curative treatment by surgical resection, survival was 85% at 1 year, 63% at 2 years, and 55% at 3 years. In a multicenter study encompassing 115 patients with incidentally discovered gallbladder cancer who underwent re-resection, residual disease in the liver was identified in 46% of patients (0% of those with stage T1 disease, 10% of those with T2 tumors, and 36% of those with T3 disease). T stage also was associated with the risk of metastasis to locoregional lymph nodes (lymph node metastasis for T1 of 13%; for T2, 31%; and for T3, 46%). In another study, a German registry of incidental gallbladder cancer identified 439 patients. Patients with tumors staged as T2 or T3 after cholecystectomy had better survival if they underwent reoperation than if they were managed with observation. Hence, reoperation should be considered for all patients who have T2 or T3 tumors or for whom the accuracy of staging is in question.

**Metastatic Colorectal Cancer**

Over 50% to 60% of patients diagnosed with colorectal cancer will develop hepatic metastases during their lifetime. Resection for hepatic metastases has been a routine part of treatment for
colorectal cancer since the publication of a large single-center experience demonstrating its safety and efficacy. Predictors of poor outcome in that study included node-positive primary, disease-free interval <12 months, more than one tumor, tumor size >5 cm, and carcinoembryonic antigen level >200 ng/mL. Traditional teaching suggested that hepatic resection for metastatic colorectal cancer to the liver, if technically feasible, should be performed only for fewer than four metastases. However, later studies challenged this paradigm. In a series of 235 patients who underwent hepatic resection for metastatic colorectal cancer, the 10-year survival rate of patients with four or more nodules was 29%, nearly comparable to the 32% survival rate of patients with only a solitary tumor metastasis. In the Memorial Sloan-Kettering Cancer Center series of 98 patients with four or more colorectal hepatic metastases who underwent resection between 1998 and 2002, the 5-year actuarial survival was 33%. Furthermore, improved chemotherapeutic regimens and surgical techniques have produced aggressive strategies for the management of this disease. Many groups now consider volume of future liver remnant and the health of the background liver, and not actual tumor number, as the primary determinants in selection for an operative approach. Hence, resectability is no longer defined by what is actually removed, but indications for hepatic resection now center on what will remain after resection.

Use of neo-adjuvant chemotherapy, portal vein embolization, two-stage hepatectomy, simultaneous ablation, and resection of extrahepatic tumors in select patients have increased the number of patients eligible for a surgical approach.

**Neuroendocrine Tumors**

Hepatic metastases from neuroendocrine tumors have a protracted natural history and commonly are associated with debilitating endocrinopathies. Several groups have advocated an aggressive surgical approach of debulking surgery, both to control symptoms and to extend survival. In a series of 170 patients undergoing resection of hepatic metastases from neuroendocrine tumors between 1977 and 1998 at the Mayo Clinic, overall survival was 61% and 35% at 5 and 10 years, respectively. There was no difference in survival between patients with carcinoid tumors and those with islet cell tumors. Major hepatectomy was performed in 91 patients (54%), and recurrence rate was 84% at 5 years. Belghiti’s group has described a two-stage strategy used in 41 patients with a primary neuroendocrine tumor and synchronous bilobar liver metastases. In the first stage, the primary tumor is resected and limited resection of metastases in the left hemiliver, combined with right portal vein ligation, is performed. After 8 weeks of hypertrophy, a right hepatectomy or extended right hepatectomy (also referred to as a right trisectionectomy; resection of Couinaud’s segments IV, V, VI, VII, and VIII of the liver) is performed. In patients treated using this strategy, the 2-, 5-, and 8-year Kaplan-Meier overall survival rates were 94%, 94%, and 79%, respectively, and disease-free survival rates were 85%, 50%, and 26%, respectively. Because systemic therapy has had little success in the treatment of advanced tumors, a broader approach using multimodal therapy has been used to increase survival and improve hormone-related symptoms. These therapies include radiofrequency or microwave ablation and intra-arterial therapy with chemoembolization or radioembolization (yttrium-90). Some centers perform liver transplantation for selected patients (carcinoid histology; primary tumor removed with curative resection; primary tumor drained by portal system; ≤50% hepatic parenchyma involved; good response or stable disease for at least 6 months during pretransplantation period; and age 55 years or younger), although this is not routine.

**Other Metastatic Tumors**

Nearly every cancer has the propensity to metastasize to the liver. Historically, enthusiasm was low for resecting metastases other than those from a colorectal cancer primary. This was due in part to the recognition that many other primary cancers (such as breast cancer) represent a systemic disease when liver metastases are present. However, more recent studies have shown acceptable 5-year survival rates in the 20% to 40% range for resection of hepatic metastases from breast, renal, and other GI tumors. In a large study of hepatic resection for non-colorectal, nonendocrine liver metastases in 1452 patients, negative prognostic factors were nonbreast origin, age >60 years, disease-free interval of <12 months, need for major hepatectomy, performance of R2 resection, and presence of extrahepatic metastases.

**TREATMENT OPTIONS FOR LIVER CANCER**

In general, the major treatment options for liver cancer can be categorized as shown in Table 31-7. The decision making for any given patient is complex and is best managed by a multidisciplinary liver tumor board. The treatments listed in Table 31-7 are not mutually exclusive; the important point is to select the most appropriate initial treatment after a complete evaluation. In general, surveillance imaging (CT or MRI) is performed every 3 to 4 months during the first year after diagnosis to observe for response, progression, or recurrence. The treatment plan is individualized and modified according to the response of the patient.

**Hepatic Resection**

For primary liver cancers or hepatic metastases, hepatic resection is the gold standard and treatment of choice. Although there are anecdotal reports of long-term survival after ablation and other regional liver therapies, liver resection remains the only real option for cure. For HCC in the setting of cirrhosis, liver transplantation also offers the potential for long-term survival, albeit with the consequences of immunosuppression. Hepatic resection...
also has been advocated for HCC in select patients with cirrhosis before secondary liver transplantation, although not without some controversy.\textsuperscript{117} Many large series of patients undergoing major hepatectomy now report mortality rates of <5%.\textsuperscript{118-121} Previously, a 1-cm tumor margin was considered desirable; however, recent studies have reported comparable survival rates with smaller margins.\textsuperscript{122-124} Technical innovation in liver surgery and a better understanding of perioperative care have even allowed surgeons to perform resections in cases with IVC involvement with extracorporeal liver surgery.\textsuperscript{125} The technical aspects of anatomic hepatic lobectomies are described later.

**Liver Transplantation**

The rationale supporting OLT for HCC includes the fact that most HCCs (>80%) arise in the setting of cirrhosis.\textsuperscript{126} The cirrhotic liver often does not have enough reserve to tolerate a formal resection. Also, HCC tumors are commonly multifocal and are underestimated by current CT or MRI imaging.\textsuperscript{127} Furthermore, recurrence rates are high at 5 years after resection (>50%). Hence, OLT is an appealing treatment because it removes both the cancer and the cirrhotic liver that leads to cancer. More than 7000 liver transplantations are performed each year in the United States, with 1-year survival rates approaching 90%. In May 2017, approximately 14,463 patients were on the waiting list for liver transplantation.\textsuperscript{128}

Initial series of OLT for HCC reported in the 1990s included advanced cases of HCC, and the 5-year survival rates were only 20% to 50%.\textsuperscript{88} This compared poorly with overall 5-year survival rates of 70% to 75% for OLT in the Organ Procurement and Transplantation Network/United Network for Organ Sharing (OPTN/UNOS) database. Mazzaferrro and colleagues at Milan subsequently showed that survival rates were markedly improved when OLT was limited to patients with early-stage HCC (stage I or stage II) with one tumor ≤5 cm, or up to three tumors no larger than 3 cm, along with the absence of gross vascular invasion or extrahepatic spread.\textsuperscript{129} Multiple studies have validated these findings, and many groups have proposed an expansion of the Milan criteria.\textsuperscript{81}

As noted previously, the 6- to 40-point MELD score was adopted by OPTN/UNOS in 2002 for allocation of deceased donor liver organs in the United States. In an attempt to prioritize patients with preserved liver function and progressive HCC, patients with stage II HCC are allocated exception points (currently 28 MELD points which activates 6 months after liver transplant listing, increasing every 3 months and capping at 34 points as long as they continue to meet transplant criteria). This allocation has had a positive effect for HCC liver transplant candidates, leading to decreased waiting list dropout and increased transplant rates with excellent long-term outcomes.\textsuperscript{130} The goal is to better equate death rates on the liver transplant waiting list for patients with stage I or stage II HCC with rates for patients with chronic liver disease without HCC. Although indications for liver transplantation have increased, the supply of donor livers has failed to keep pace with the numbers of potential recipients. A partial solution has been the use of living donor grafts. This is especially true in Asia where the incidence of HCC is high and the rate of cadaveric donation is low. Living donor grafts include right and left lobes, as well as dual grafts from separate donors to provide adequate hepatic mass to the recipient. The use of living donor grafts also allows for transplant programs to push the boundaries by accepting patients beyond the Milan criteria with good results.\textsuperscript{112}

**Radiofrequency Ablation**

In 1891, d’Arsonval discovered that radiofrequency (RF) waves delivered as an alternating electric current (>10 kHz) could pass through living tissue without causing pain or neuromuscular excitation. The resistance of the tissue to the rapidly alternating current produced heat. This discovery contributed to the development of the surgical application of electrocautery. In 1908, Beer used RF coagulation to destroy urinary bladder tumors. Cushing and Bovie later applied RF ablation to intracranial tumors. In 1961, Lounsberry studied the histologic changes of the liver after RF ablation (RFA) in animal models. He found that RF caused local tissue destruction with uniform necrosis. In the early 1990s, two groups proposed that RFA can be an effective method for destroying unresectable malignant liver tumors.\textsuperscript{132,133} Both groups found that RFA produced lesions with well-demarcated areas of necrosis without viable tumor cells present. Clinical reports after short-term follow-up suggested that RFA was safe and effective in the treatment of liver tumors.\textsuperscript{134-136} However, Abdalla and colleagues examined data for 358 consecutive patients with colorectal liver metastases treated with curative intent over a 10-year period (1992 to 2002).\textsuperscript{137} Liver-only recurrence after RFA was four times the rate after resection (44% vs. 11% of patients), and RFA alone or in combination with resection did not provide survival rates comparable to those with resection alone. Nonetheless, RFA remains a common procedure that can be performed by a percutaneous, minimally invasive laparoscopic, or open approach.\textsuperscript{138,139} It also has been used successfully to ablate small HCCs as a bridge to liver transplantation.\textsuperscript{140} Results were reported for the first randomized clinical trial involving RFA treatment for HCC in 291 Chinese patients with three or fewer HCC tumors ranging in size from 3 to 7.5 cm.\textsuperscript{141} Patients were randomly assigned to treatment arms of RFA alone (n = 100), transarterial chemoembolization (TACE) alone (n = 95), or combined TACE plus RFA (n = 96). At a median follow-up of 28.5 months, median survival was 22 months in the RFA group, 24 months in the TACE group, and 37 months in the TACE plus RFA group. Patients treated with TACE plus RFA had significantly better overall survival than those treated with TACE alone (P < .001) or RFA alone (P < .001). Sucandy and colleagues examined long-term 5- and 10-year overall survival in 320 patients that had hepatic RFA for HCC or colorectal liver metastases (CLM).\textsuperscript{142} The majority of patients (71%) had a single tumor ablation. Minimum 5-year follow-up was available in 89% patients, with a median follow-up of 115.3 months. In the HCC group, the 5- and 10-year overall survivals were 38.5% and 23.4%, respectively, while in the CLM group, the 5- and 10-year overall survivals were 27.6% and 15%, respectively.

**Ethanol Ablation, Cryosurgery, and Microwave Ablation**

Percutaneous ethanol injection has been shown to be a safe and effective treatment for small HCCs.\textsuperscript{138} The ethanol usually is delivered by percutaneous injection under ultrasound or CT guidance. Percutaneous ethanol injection also is used to treat small HCC tumors as a bridge to liver transplantation in some centers to avoid patient dropout.\textsuperscript{10} Although cryosurgery was used in the late 1980s and 1990s for ablation of liver tumors, many have abandoned this approach in favor of RFA because of the latter’s fewer side effects and ease of use. Microwave ablation is a thermal ablative technique used in the management of unresectable liver tumors to produce a coagulation necrosis.
In a multicenter phase 2 U.S. trial using a 915-MHz microwave generator, 87 patients underwent 94 ablation procedures for 224 hepatic tumors.\textsuperscript{143} Forty-five percent of the procedures were performed using an open approach, 7% laparoscopically, and 48% percutaneously. The average tumor size was 3.6 cm (range, 0.5 to 9.0 cm). At a mean follow-up of 19 months, 47% of the patients were alive with no evidence of disease. Local recurrence at the ablation site occurred in 2.7% of tumors, and regional recurrence occurred in 43% of patients. There were no procedure-related deaths. Further studies are required to define the role of this technology in relation to the other ablation options available.

**Chemoembolization and Hepatic Artery Pump Chemoperfusion**

Chemoembolization is the process of injecting chemotherapeutic drugs combined with embolization particles into the hepatic artery that supplies the liver tumor using a percutaneous, transfemoral approach. It is most commonly used for treatment of unresectable HCC. Three randomized trials and a meta-analysis have shown a survival benefit with chemoembolization.\textsuperscript{144-147} In a study by Lo and colleagues, 80 Asian patients were randomly assigned to receive either chemoembolization with cisplatin in lipiodol or symptomatic treatment only.\textsuperscript{144} Chemoembolization resulted in a marked tumor response, and the actuarial survival was significantly better in the chemoembolization group (1- and 3-year survival of 57% and 26%, respectively) than in the control group (1- and 3-year survival of 32% and 3%, respectively). In another randomized trial, a Barcelona group compared chemoembolization with doxorubicin versus supportive care and showed that chemoembolization significantly improved survival.\textsuperscript{145} Finally, in a large prospective cohort study of 8510 patients with unresectable HCC in Japan who received transcatheter arterial lipiodol chemoembolization, the 5-year survival rate was 26% and median survival time was 34 months.\textsuperscript{146} The TACE-related mortality rate after the initial therapy was 0.5%. Complications of TACE include liver dysfunction or liver failure, hepatic abscess, and hepatic artery thrombosis. Multiple studies also have shown promising results for chemoembolization with drug-eluting beads in treatment of HCC.\textsuperscript{148}

In the 1990s, hepatic artery pump chemoperfusion with floxuridine for colorectal cancer metastases to the liver was used both for treatment of inoperable disease and in the adjuvant setting.\textsuperscript{149} However, in the modern era of improved chemotherapeutic options, this treatment modality is seldom used outside of a clinical trial.

**Yttrium-90 Microspheres**

Selective internal radioembolization or transarterial radioembolization (TARE) is a promising new treatment modality for patients with inoperable primary or metastatic liver tumors. The treatment is a minimally invasive transcatheter therapy in which radioactive microspheres are infused into the hepatic arteries via a transfemoral percutaneous approach. The yttrium-90 microspheres are directly injected into the hepatic artery branches that supply the tumor. Once infused, the microspheres deliver doses of high-energy, low-penetration radiation selectively to the tumor. The main indications are inoperable HCC\textsuperscript{150} and colorectal cancer hepatic metastases for which systemic chemotherapy has failed.\textsuperscript{151,152} In a study involving 137 patients with unresectable chemorefractory liver metastases treated with radioembolization, there was a response rate of 42.8% (2.1% complete response, 40.7% partial response) according to World Health Organization criteria.\textsuperscript{152} One-year survival rate was 47.8%, and 2-year survival rate was 30.9%. Median survival was 457 days for patients with colorectal tumor metastases, 776 days for those with neuroendocrine tumor metastases, and 207 days for those with noncolorectal, nonneuroendocrine tumor metastases. The two products available in the United States are SIR-Spheres (Sirius, Sydney, Australia) and TheraSphere (Nordian, Ottawa, Canada).

**Stereotactic Radiosurgery and Intensity-Modulated Radiation Therapy**

Although stereotactic radiosurgery (with CyberKnife and other systems) is in widespread use for brain and spinal tumors, body application to HCC or metastatic liver tumors has only recently occurred. In a phase 1 study, 31 patients with unresectable HCCs and 10 with unresectable cholangiocarcinomas completed a six-fraction course of stereotactic body radiotherapy.\textsuperscript{153} The treatment was well tolerated, and median survival was 11.7 and 15.0 months for the two groups, respectively. A similar safety profile was observed in a study in the Netherlands.\textsuperscript{154} Further clinical trials are required to define the future role of stereotactic radiosurgery in treatment of HCC and metastatic tumors. Intensity-modulated radiation therapy (IMRT) is another technological advancement that facilitates the targeted delivery of external-beam radiation. Early clinical data suggested favorable outcomes with IMRT for the treatment of patients with unresectable HCC, and ongoing trials are further examining the role of IMRT for these locally advanced tumors.

**Downstaging**

In more advanced-stage patients not eligible for MELD exception points, hepatic-directed therapy including TACE and tumor ablation with radiofrequency, microwave, and ethanol ablation have been found to be effective in shrinking tumors to meet Milan criteria (downstaging). Multiple centers have used downstaging to allow for OLT in patients whose tumors responded and shrank to meet eligibility criteria.\textsuperscript{155,156}

**Systemic Chemotherapy**

Chemotherapy has not demonstrated great efficacy in patients with HCC, especially in patients with significant cirrhosis. For treatment of HCC, the multikinase inhibitor sorafenib has shown some efficacy in a phase 3 randomized international multicenter trial. The SHARP trial (Sorafenib HCC Assessment Randomized Protocol) enrolled 602 patients with Child’s class A cirrhosis and inoperable HCC. At interim analysis, the trial was discontinued because a survival benefit was found in the treatment group. The median overall survival for patients receiving sorafenib was 10.7 months versus 7.9 months for patients in the control arm. Based on these findings, sorafenib received accelerated Food and Drug Administration approval for the treatment of advanced unresectable HCC.\textsuperscript{157} Future studies will likely examine the role of other molecularly targeted agents and combinations of sorafenib with other treatment modalities.

**HEPATIC RESECTION SURGICAL TECHNIQUES**

**Nomenclature**

Due to the confusion in language with regard to anatomic descriptions of hepatic resections, a common nomenclature was introduced at the International Hepato-Pancreato-Biliary Association
meeting in Brisbane, Australia, in 2000 (Table 31-8). The goal was to provide universal terminology for liver anatomy and hepatic resections because there was much overlap among the designations for hepatic lobes, sections, sectors, and segments used by surgeons worldwide (Fig. 31-21). The most common or prevailing anatomic pattern was used as the basis for naming liver anatomy, and the surgical procedure nomenclature adopted for hepatic resections was based on the assigned anatomic terminology. Adoption of a common language should enable hepatic surgeons to better understand and interpret liver surgery publications from different continents and disseminate their knowledge to the next generation of hepatobiliary surgeons. Nonetheless, even today, the literature is full of both old and new liver resection terminology, so the surgeon in training must be familiar with all the various classifications.

Table 31-8

Brisbane 2000 liver terminology

<table>
<thead>
<tr>
<th>OLDER HEPATIC RESECTION TERMINOLOGY</th>
<th>BRISBANE 2000 HEPATIC RESECTION TERMINOLOGY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right hepatic lobectomy</td>
<td>Right hepatectomy or right hemihepatectomy (V, VI, VII, VIII)</td>
</tr>
<tr>
<td>Left hepatic lobectomy</td>
<td>Left hepatectomy or left hemihepatectomy (II, III, IV)</td>
</tr>
<tr>
<td>Right hepatic trisegmentectomy</td>
<td>Right trisectionectomy or extended right hepatectomy (or hemihepatectomy, IV, V, VI, VII, VIII)</td>
</tr>
<tr>
<td>Left hepatic trisegmentectomy</td>
<td>Left trisectionectomy or extended left hepatectomy (or hemihepatectomy, II, III, IV, V, VIII)</td>
</tr>
<tr>
<td>Left lateral segmentectomy</td>
<td>Left lateral sectionectomy or bisegmentectomy (II, III)</td>
</tr>
<tr>
<td>Right posterior lobectomy</td>
<td>Right posterior sectionectomy (VI, VII)</td>
</tr>
<tr>
<td>Caudate lobectomy</td>
<td>Caudate lobectomy or segmentectomy (I)</td>
</tr>
</tbody>
</table>

ALTERNATIVE “SECTOR” TERMINOLOGY

- Right anterior sectorectomy
- Right posterior sectorectomy or right lateral sectorectomy
- Left medial sectorectomy or left paramedian sectorectomy (bisegmentectomy, III, IV)
- Left lateral sectorectomy (segmentectomy, II)

Techniques and Devices for Dividing the Hepatic Parenchyma

Hepatic resection surgery has evolved over the past 50 years. A better understanding of liver anatomy and physiology, coupled with improved anesthesia techniques and widespread use of intraoperative ultrasound, has led to virtually “bloodless” liver surgery in the modern era (the year 2000 to the present). Innovations in technology have expanded the list of liver parenchymal transection devices and hemostatic agents (Table 31-9). Use of each device or agent has a learning curve, and undoubtedly every experienced hepatic surgeon has his or her personal preferences.

One major trend has been the application of vascular stapling devices for division of the hepatic and portal veins. Based on early reports of successful stapling of extrahepatic

Table 31-9

Techniques and devices for dividing liver parenchyma and achieving hemostasis

- Blunt fracture and clips
- Monopolar cautery (Bovie)
- Bipolar cautery
- Argon beam coagulator
- CUSA ultrasonic dissector
- Hydro-Jet water-jet dissector
- Harmonic Scalpel, AutoSonix ultrasonic transsector-coagulator
- LigaSure tissue fusion system
- SurgRx EnSeal tissue sealing and transection system
- Gyrus PK cutting forceps
- Endovascular staplers
- TissueLink sealing devices
- Habib 4X Laparoscopic sealer
- InLine bipolar linear coagulator
- Topical agents (fibrin glues, Surgicel, Gelfoam, Avitene, Tisseel, Floseal, Crosseal)
vessels, stapling devices are now being used in the parenchymal transection phase, which remains a source of potential blood loss due to back bleeding from the middle hepatic vein.\textsuperscript{167,168} One advantage of the stapling technique is the speed with which the transection can be performed, which minimizes surface bleeding and period of ischemia for the remnant liver. However, a major disadvantage of the stapling technique is the cost of multiple stapler cartridges. This is balanced by the decreased expenses reported with avoidance of ICU admission and blood transfusion, as well as shortened operating room time. Another consideration in the use of staplers for parenchymal transection is the potential for bile leaks. However, in a large series of 101 consecutive right hemihepatectomies performed using the stapling technique, there was only one reported bile leak (1\%), which sealed after ERCP.\textsuperscript{166}

Steps in Commonly Performed Hepatic Resections
A fundamental understanding of hepatic anatomy is vital for any surgeon with the desire to perform hepatobiliary surgery. Each hepatic resection surgery can be broken down into a series of orderly steps. The key to being a proficient hepatic surgeon is not to operate swiftly but rather to accomplish the operation by completing the steps in an orchestrated fashion. Mastery of the operative steps coupled with knowledge of liver anatomy and the common anatomic variants provides the foundation for safe hepatic surgery. There are many different techniques and sequences for accomplishing each of the anatomic (and nonanatomic) hepatic operations. The authors present their preferred approach in a stepwise fashion for right hepatic lobectomy (right hemihepatectomy), left hepatic lobectomy (left hemihepatectomy), and left lateral sectectomy (left lateral sectionectomy). Provision of a detailed approach for every type of liver resection is beyond the scope of this chapter, and readers are referred to several excellent descriptions.\textsuperscript{169}

Steps Common to All Open Major Hepatic Resections
1. Make the skin incision—right subcostal with or without a partial or complete left subcostal extension across the midline, depending on the patient’s habitus and liver/tumor anatomy.
2. Open and explore the abdomen, and place a fixed table retractor (e.g., Thompson or Bookwalter).
3. Examine the liver with bimanual palpation. Perform liver ultrasound, and confirm the operation to be performed.
4. Take down the round and falciform ligaments, and expose the anterior surface of the hepatic veins.
5. For a left hepatectomy, divide the left triangular ligament; for a right hepatectomy, mobilize the right lobe from the right coronary and triangular ligaments.
6. Open the gastrohepatic ligament, palpate the porta hepatitis, and assess for accessory or replaced hepatic arteries.
7. Perform a cholecystectomy; leave the gallbladder with the cystic duct intact if the gallbladder is involved by the tumor.

Right Hepatic Lobectomy (Right Hepatectomy or Hemihepatectomy)
8. Mobilize the liver from the anterior aspect of the IVC in “piggyback” fashion; ligate the short hepatic veins up to the right hepatic vein (RHV).
9. Perform a right hilar dissection—gently lower the hilar plate, then doubly ligate and divide the right hepatic artery (RHA), superior to the right side of the common bile duct.
10. Doubly ligate and divide a replaced or accessory RHA if present.

Left Hepatic Lobectomy (Left Hepatectomy or Hemihepatectomy)
8. Widely open the gastrohepatic ligament flush with the undersurface of the left lateral section and the caudate lobe.
9. Doubly ligate and divide a replaced or accessory left hepatic artery (LHA) if present.
10. Clamp the round ligament (ligament teres) and pull it anteriorly as a handle to expose the left hilum.
11. Divide any existing parenchymal bridge between segments III and IVB.
12. Dissect the left hilum at the base of the umbilical fissure and lower the hilar plate anterior to the left portal pedicle.
13. Incise the peritoneum overlying the hilum from the left side, and doubly ligate the LHA (after test clamping and confirming a palpable pulse in the RHA).
14. Dissect the portal vein at the base of the umbilical fissure (it will take a nearly 90° bend from the transverse to the umbilical portion).
15. Expose the portal vein, identifying the right and left branches. Control the small portal vein branch off the LPV to the caudate lobe to allow the exposure of additional length. Divide the LPV either with a vascular stapler or between vascular clamps.
16. Ligate and divide the ligamentum venosum caudally.
17. Identify the long extrahepatic course of the left hepatic duct (LHD) behind the portal vein. Ligate and divide the LHD at the umbilical fissure.
18. Fold the left lateral segment up and back to the right, exposing the window at the base of the left hepatic vein (LHV) as it enters the IVC. This is facilitated by dividing any loose areolar tissue overlying the ligamentum venosum, which is divided proximally.
19. Pass a large, blunt right-angle clamp in the window between the RHV and the MHV, and hug the back of the MHV, aiming for the deep edge of the LHV. Do not force it or perforate the IVC or MHV.
20. Pass the silastic tube of a Jackson-Pratt drain through this window.
21. Notch or divide the caudate process crossing to the left hepatic lobe and bring the drain up and through this notch.
22. Hang the liver over the drain by pulling up as you divide through the liver parenchyma.
23. Repeat ultrasound and confirm the transection plane on the anterior surface, staying close to the demarcated line. Do not bisect the MHV as it passes tangentially from the left to the right lobe.
24. Cauterize down approximately 1 cm in the liver parenchyma, then switch to a hydro-jet dissection device in combination with Bovie electrocautery and suture ligation.
25. Continue parenchymal division until the left/middle hepatic veins are encountered.
26. Divide the LHV and MHV between vascular clamps and suture the ligate the LHV/MHV.
27. Check the transected edge of the liver for surgical bleeding; ensure hemostasis of the transected edge with an argon beam coagulator and suture ligation.
28. Inspect the transection surface for bile leaks. These should be clipped or suture ligated. Apply dilute solution of hydrogen peroxide to facilitate the visualization of bile leaks.
29. Perform completion ultrasound to confirm RPV inflow and RHV outflow.
30. Apply tissue sealant to the transected surface of the liver. Place a Jackson-Pratt drain in the left subphrenic space and close the abdomen (Fig. 31-23).

Comments Because the right posterior duct arises from the left hepatic duct (LHD) in approximately 20% of cases (see Fig. 31-9) and the right anterior duct comes off the LHD in approximately 5% of cases, it is vital to divide the LHD at the base of the umbilical fissure and not more centrally in the hilum as it bifurcates. If the LHD were divided as it appears to bifurcate from the right hepatic duct, then approximately 20% to
25% of the time, either the right posterior or right anterior duct would be transected. After the LHD is divided as described earlier (Step 17), the liver parenchyma is scored and divided horizontally approximately 1 cm above the left hilum; the surgeon thus assumes that an aberrant right anterior or posterior duct is coming off the LHD in the hilum and preserves it. Then as the parenchymal transection reaches the left side of the gallbladder fossa, the transection plane turns vertical to run parallel to Cantlie’s line (or the left edge of the gallbladder bed). The left lobe of the liver will be well demarcated at this point (after the vascular inflow has been divided), which guides the transection plane on the anterior surface. In general, the transection plane should be close to the demarcation line to minimize the amount of devascularized liver remaining. When dividing the LHV and MHV, the surgeon should keep in mind that they have a common trunk approximately 90% of the time. If it is not easy to open the window deep to the MHV and LHV, then division of the MHV and LHV can be accomplished after the parenchymal transection.

**Left Lateral Segmentectomy (Left Lateral Sectionectomy)**

8. Widely open the gastrohepatic ligament flush with the undersurface of the left lateral section and the caudate lobe.
9. Doubly ligate and divide a replaced or accessory LHA if present.
10. Clamp the round ligament and pull it anteriorly as a handle to expose the left hilum.
11. Divide any existing parenchymal bridge between segments III and IVB.
12. Carry the dissection down from the end of the round ligament, and the segment III pedicle will be encountered.
13. Incise the peritoneal reflection on the left side of the round ligament as it inserts into the umbilical fissure. This will facilitate encircling the segment III and II pedicles, which can be divided separately. When encircling the segment II pedicle, take care to avoid injury to the caudate inflow vessels coming off the LPV.
14. Divide the liver parenchyma, staying flush on the left side of the falciform ligament using Bovie electrocautery.
15. Divide the LHV as the parenchymal transection is complete.
16. A Pringle maneuver usually is not required for a left lateral sectionectomy because complete devascularization occurs before transection and little back bleeding is encountered.

**Comments**

If the segment III and II LHA branches are large, they can be individually ligated in the left hilum before the pedicles (with portal vein and hepatic duct branches) are taken. If the tumor is more peripheral in the left lateral segment, then the segment III and II pedicles can be divided with a vascular stapler inside the liver during the parenchymal transection.

**Pringle and Ischemic Preconditioning**

Pringle described clamping of the portal triad a century ago in the landmark paper “Notes on the Arrest of Hepatic Hemorrhage Due to Trauma.” Although the Pringle maneuver was initially described for controlling bleeding due to traumatic liver injury, it is commonly used during elective hepatic resections. The goal is to minimize blood loss and hypotension, which add significant morbidity to the operation. Furthermore, intraoperative blood transfusion has been shown to be an independent risk factor for increased postoperative infection as well as worse patient survival in some studies. Therefore, all efforts should be made to minimize blood loss during hepatic resection.

Although the liver has been shown to tolerate up to 1 hour of warm ischemia, some technical variations of the Pringle maneuver include intermittent vascular occlusion with cycles of approximately 15 minutes on and 5 minutes off. Experimental and clinical studies have demonstrated the efficacy of intermittent vascular occlusion in decreasing ischemia/reperfusion injury compared with continuous vascular occlusion, with less elevation of postoperative liver enzyme levels. Another variation is selective hemihepatic vascular occlusion, which can reduce the severity of visceral congestion and total liver ischemia. In one prospective trial of total versus selective portal triad clamping, both techniques of inflow clamping were found to be equally effective for patients with normal livers, but greater liver damage was observed with total inflow occlusion in patients with cirrhotic livers.

In an attempt to decrease the ischemic damage associated with inflow occlusion, some hepatic surgeons have advocated the use of ischemic preconditioning. Ischemic preconditioning refers to the brief interruption of blood flow to an organ, followed by a short reperfusion period, and then a more prolonged period of ischemia. In a randomized clinical trial involving 100 patients undergoing major hepatic resection, Clavien and colleagues reported significantly less liver injury in the group who received ischemic preconditioning with a 10-minute clamp, a 10-minute reperfusion, and then a 30-minute clamp than in those who received a 30-minute clamp alone. Patients with steatosis also were especially protected by ischemic preconditioning, and the mechanism was shown to be related in part to preservation of the adenosine triphosphate content of liver tissue.

**Preoperative Portal Vein Embolization**

The observation that tumor thrombosis of a major portal vein branch induced ipsilateral lobar atrophy and contralateral lobe hypertrophy led to the concept of intentional preoperative portal vein embolization (PVE) to induce compensatory hypertrophy of the remnant liver. This procedure was first described in the 1980s and is accomplished via a percutaneous, transhepatic route. Numerous studies have subsequently confirmed that PVE is effective in inducing hypertrophy of nonembolized hepatic segments. PVE usually is performed in the setting of a planned right trisectionectomy or extended left hepatectomy (also referred to as a left trisectionectomy; resection of Couinaud’s segments II, III, IV, V, and VIII of the liver) or extended hepatic lobectomy when it is thought that the patient’s remnant liver will be too small to support liver function. The future liver remnant volume (e.g., the volume of segments II, III, and I) in a patient undergoing a planned right trisectionectomy can be directly measured by helical CT and then divided by the total estimated liver volume to calculate the percentage of the future liver remnant. If the future liver remnant is thought to be too small, then PVE should be considered to increase the size of the future liver remnant. In general, surgery is planned approximately 4 weeks after PVE to allow adequate time for hypertrophy.

There is no universal agreement on what constitutes a future liver remnant adequate to avoid postoperative liver failure. It is thought that 25% to 30% of the total liver volume is adequate in patients with a normal liver. Vauthey and associates reported that major postoperative complications were increased when the estimated future liver remnant was <25%. Farges and colleagues conducted a prospective study to assess the benefits of PVE before right hepatectomy. They demonstrated
that PVE had no beneficial effect on the postoperative course in patients with normal livers but significantly reduced postoperative complications in patients with chronic liver diseases.\textsuperscript{185} A larger remnant may be necessary even in patients with normal livers when a complex hepatectomy is planned or when the background liver is steatotic.\textsuperscript{186} This is especially relevant with the increased incidence of fatty liver disease. A larger remnant may also be needed when patients have received preoperative chemotherapy. Some have suggested that 40\% of the total hepatic volume should remain to minimize postoperative complications in patients who have undergoing liver disease or who have received preoperative chemotherapy for colorectal cancer metastases.\textsuperscript{187,188} In a recent study encompassing 112 patients who underwent PVE, major complications, hepatic insufficiency, length of hospital stay, and 90-day mortality rate were significantly greater in patients with a standardized future liver remnant of \(\leq 20\%\) or a degree of hypertrophy of \(< 5\%\) than in patients with higher values.\textsuperscript{189} In another study, the authors performed PVE during neoadjuvant chemotherapy for colorectal cancer metastases. After a median wait of 30 days after PVE, patients receiving neoadjuvant chemotherapy showed median liver growth of 22\% in the contralateral (nonembolized) lobe compared with 26\% for those not receiving chemotherapy (not a statistically significant difference), which indicated that liver growth occurs after PVE even when cytotoxic chemotherapy is administered.\textsuperscript{190} PVE-related complications occur at a relatively low rate and include bleeding, hemobilia, liver abscess, incompletely embolization, and small bowel obstruction. To augment the ability to increase liver function reserve in patients who have undergone PVE, some groups have added ipsilateral hepatic artery embolization and ipsilateral hepatic vein embolization.\textsuperscript{191}

**Staged Hepatectomy, ALPPS, and Repeat Hepatic Resection for Recurrent Liver Cancer**

A two-stage hepatectomy is a sequential resection strategy to remove all metastatic liver tumors when it is impossible to resect all disease in a single operative procedure. The first-stage hepatectomy usually consists of clearance of the left hemiliver by nonanatomic resection, followed by right portal vein ligation or embolization to induce left lobe hypertrophy.\textsuperscript{192,193} This is followed by a second-stage major right hepatectomy or extended right hepatectomy to resect the right liver metastases. This approach is most commonly used in cases of initially unresectable colorectal hepatic metastases and has yielded very good results.\textsuperscript{193}

Another technique to increase future liver remnant to avoid post-hepatectomy liver failure is known as “Associating Liver Partition and Portal Vein Ligation for Staged Hepatectomy (ALPPS)” and was described in 2012. This technique consists of operative portal vein ligation with in situ liver transection along the future line of resection leaving the arterial and hepatic vein branches intact.\textsuperscript{194} There have been several technical modifications to the original description, and this remains a hot topic that is actively debated (ALPPS vs. PVE) at the current liver surgery meetings.

The majority of patients undergoing hepatic resection for colorectal cancer metastases experience a recurrence. For those with limited disease recurrence confined to the liver, repeat hepatectomy is a reasonable option and can be performed with low morbidity and mortality in experienced hands.\textsuperscript{195} In one study, 126 patients who underwent a second liver resection for colorectal cancer metastases had 1-, 3-, and 5-year survival rates of 86\%, 51\%, and 34\%, respectively. By multivariate analysis, the presence of more than one lesion and a tumor size of \(> 5\) cm were independent prognostic indicators of reduced survival.\textsuperscript{196} In another study, 40 patients underwent a second hepatectomy for liver metastases from colorectal cancer and experienced a survival benefit similar to that from the first hepatectomy; however, the results suggested that this approach should be limited to those patients who do not have extrahepatic disease and for whom \(> 1\) year has elapsed since the first operation.\textsuperscript{197} A meta-analysis of 21 studies examining clinical outcomes after first and second liver resections for colorectal cancer metastases showed that repeat hepatectomy was safe and provided a survival benefit equal to that from the first liver resection.\textsuperscript{198}

Repeat hepatectomy also has been performed in patients with HCC. Nakajima and colleagues reported on follow-up of 94 patients who underwent curative liver resection for HCC from 1991 to 1996.\textsuperscript{199} Of these, 57 patients had isolated recurrent disease in the liver. Twelve of these 57 patients underwent repeat hepatic resection, whereas the other 45 patients received ablation therapy. The overall survival rate in those undergoing a second hepatectomy was 90\% at 2 years; however, the disease-free survival rate was only 31\% at 2 years, significantly lower than the 62\% rate after initial hepatectomy. Likewise, in another group of 84 patients who underwent second hepatectomy for recurrent HCC, the overall 5-year survival rate was 50\%, but the recurrence-free survival rate was only 10\%.\textsuperscript{200} In a report of 67 patients undergoing a second resection for HCC, overall 1-, 3-, and 5-year survival rates were 93\%, 70\%, and 56\%, respectively.\textsuperscript{201} Multivariate analysis showed that absence of portal invasion at the second resection, single HCC at primary hepatectomy, and disease-free interval of \(\geq 1\) year after primary hepatectomy were independent prognostic factors after the second resection.

**LAPAROSCOPIC LIVER RESECTION**

Cherqui and colleagues first reported in 2000 that laparoscopic hepatic surgery was feasible.\textsuperscript{202} Since this initial report, laparoscopic liver surgery has expanded from the simple unroofing of hepatic cysts to resection of peripheral benign lesions to formal anatomic lobectomies for malignancy and laparoscopic hepatectomy for living donor liver transplantation. While minimally invasive approaches have been widely adopted in other areas of abdominal surgery, there was initial apprehension regarding laparoscopic liver resection (LLR), hampering its widespread adoption.\textsuperscript{203} Great strides have been made in the past decade with techniques of laparoscopic liver resection,\textsuperscript{204,205} and two International Laparoscopic Liver Resection Consensus Conferences have been convened in Louisville (2008) and Morioka (2014).\textsuperscript{206,207} Indications for liver resection should not be altered by the availability of minimally invasive liver resection techniques. Currently, over 9500 cases of laparoscopic liver resection have been reported worldwide, with over 50\% of the cases being done for malignancy.\textsuperscript{208}

Pure laparoscopic and hand-assisted laparoscopic liver resection are the two most commonly used techniques for minimally invasive liver resection surgery.\textsuperscript{209,210} While robotic liver resection is being used by several groups.\textsuperscript{211} Advantages of the hand-port include tactile feedback, facilitation of liver mobilization, and ease of ability to control bleeding. Also, when doing a large parenchymal resection, the hand-port can be comparable in size to the extraction port utilized in the purely laparoscopic
approach. The hand-assisted approach can be ideal for surgeons beginning the transition to laparoscopic liver resection and for more experienced laparoscopic HPB surgeons doing laparoscopic major hepatectomies or as an alternative to conversion to open surgery.

Benefits of laparoscopic liver resection include less blood loss, decreased morbidity, decreased postoperative pain and narcotic requirements, faster return of bowel function, and shorter length of hospital stay compared to open hepatic resection. Long-term oncologic outcomes for HCC and CLM have been shown to be comparable for laparoscopic vs. open liver resection using propensity score matching and meta-analysis studies. The learning curve for laparoscopic liver resection has been reported to be around 60 cases, although this may be greater for laparoscopic major hepatectomy. Cost analysis has shown that laparoscopic liver resection is cost effective compared to open liver resection, where the added cost of the operating room disposables are more than offset by the savings associated with ~50% reduction in the hospital length of stay.

Initial experience with laparoscopic living-donor hepatectomy for transplantation was with left lateral segmentectomy during liver allograft procurement for pediatric transplants. With continued advances in laparoscopic liver surgery, LLR has been applied to adult-to-adult donor right hepatectomy, although this approach remains controversial.

In summary, laparoscopic liver resection can now be performed safely by experienced surgeons in selected patients. Compared to open hepatic resection, the laparoscopic approach has benefits of less blood loss, reduced postoperative pain, and a shorter length of hospital stay, with similar oncologic outcomes for resection of HCC and limited colorectal liver metastases.

REFERENCES

Entries highlighted in bright blue are key references.


40. Renner EL. How to decide when to list a patient with acute liver failure for liver transplantation? Clichy or King’s College criteria, or something else? J Hepatol. 2007;46(4):554-557.


156. Yao FY, Kerlan RK, Jr, Hirose R, et al. Excellent outcome following down-staging of hepatocellular carcinoma prior to


ANATOMY

Gallbladder

The gallbladder is a pear-shaped sac that measures around 7 to 10 cm long, with an average capacity of 30 to 50 mL. When obstructed, the gallbladder can distend markedly and contain up to 300 mL of fluid. The gallbladder is located in an anatomic fossa on the inferior surface of the liver. Cantle’s line, a vertical plane running from the gallbladder fossa anteriorly to the inferior vena cava (IVC) posteriorly divides the liver into right and left lobes. The gallbladder itself is divided into four anatomic areas: the fundus, the body, the infundibulum, and the neck. The fundus is the rounded, blind end that normally extends 1 to 2 cm beyond the liver’s margin and contains most of the smooth muscle of the organ. The body functions as the main storage area and contains most of the elastic tissue allowing for distention. As the body tapers towards the neck of the gallbladder, a mucosal outpouching is present at the junction of the neck and the cystic duct, known as the infundibulum or Hartmann’s pouch. Beyond this, the neck of the gallbladder lies in the deepest part of the gallbladder fossa and can extend slightly into the free portion of the hepatoduodenal ligament, where it connects with the cystic duct (Fig. 32-1).1

The same peritoneal lining that covers the liver extends to cover the fundus and the inferior surface of the gallbladder. Occasionally, part or all of the gallbladder is embedded deep inside the liver parenchyma (an intrahepatic gallbladder). Rarely, the gallbladder has a complete peritoneal covering on all sides and is suspended in a mesentery off the inferior surface of the liver.

The mucosal lining of the gallbladder is formed by a single, highly redundant, simple columnar epithelium that contains cholesterol and fat globules. The mucus secreted into the gallbladder originates in tubuloalveolar glands that are found in the mucosal lining of the infundibulum and neck of the gallbladder, but are absent from the body and fundus. The epithelial lining of the gallbladder is supported by a lamina propria. The gallbladder differs histologically from the rest of the gastrointestinal (GI) tract in that it lacks a muscularis mucosa and submucosa. The muscular layer has circular, longitudinal, and oblique fibers, but without well-defined layers. The adventitia contains connective tissue, nerves, vessels, lymphatics, and adipocytes. The gallbladder is covered by serosa except where the gallbladder is embedded in the liver.

The cystic artery that supplies the gallbladder is usually a branch of the right hepatic artery (>90% of the time). The course of the cystic artery may vary, but it nearly always is found within the hepatocystic triangle (triangle of Calot), the area bound by the cystic duct, common hepatic duct, and the inferior edge of the liver. When the cystic artery reaches the neck of the gallbladder, it divides into anterior and posterior divisions. Venous return is carried either through small veins that enter directly into the liver or, rarely, to a large cystic vein that carries blood back to the portal vein. Gallbladder lymphatics drain into nodes
Key Points

1. The physiology of the gallbladder, biliary tree, and sphincter of Oddi are regulated by a complex interplay of hormones and neuronal inputs designed to coordinate bile release with food consumption. Dysfunctions related to this activity are linked to the development of gallbladder pathologies as described in this chapter.

2. In Western countries, the most common type of gallstones are cholesterol stones. The pathogenesis of these stones relates to supersaturation of bile with cholesterol and subsequent precipitation.

3. Laparoscopic cholecystectomy has been demonstrated to be safe and effective, and it has become the treatment of choice for symptomatic gallstones. Knowledge of the various anatomic anomalies of the cystic duct and artery is critical in guiding the dissection of these structures and avoiding injury to the common bile duct during cholecystectomy.

4. Common bile duct injuries, although uncommon, can be devastating to patients. Proper exposure of the hepatocystic (Calot’s) triangle to obtain the critical view of safety and careful identification of the anatomic structures are keys to avoiding these injuries. Once a bile duct injury is diagnosed, the best outcomes are seen at large referral centers with experienced biliary surgeons, and patient transfer may be required.

5. The main risk factor for gallbladder disease in Western countries is cholelithiasis. The main complications include cholecystitis, cholechocholithiasis, cholangitis, and biliary pancreatitis. Cholelithiasis is also the major risk factor for the development of gallbladder cancer.

6. Carcinomas of the gallbladder or bile ducts generally have a poor prognosis because patients usually present late in the disease process and have poor response to chemotherapy and radiation. Surgery offers the best chance for survival and has good long-term outcomes in patients with early-stage disease.

at the neck of the gallbladder. Frequently, a visible lymph node (Lund’s or Mascagni’s node, often referred to as Calot’s node) overlies the insertion of the cystic artery into the gallbladder wall. The gallbladder receives parasympathetic, sympathetic and sensory innervation through nerve fibers running largely through the gastro hepatic ligament. Parasympathetic (cholinergic) fibers arise from the hepatic branches of the vagus nerve to stimulate activity in the gallbladder, bile ducts, and liver. These vagal branches also have peptide-containing nerves containing agents such as substance P, somatostatin, enkephalins, and vasoactive intestinal polypeptide (VIP).

The sympathetic and sensory braches of the gallbladder, liver, and bile ducts pass through the celiac plexus and control gallbladder relaxation and mediate the pain of biliary colic.

Bile Ducts

The extrahepatic biliary tree consists of the right and left hepatic ducts, the common hepatic duct, the cystic duct, and the common bile duct. Exiting the liver, the left hepatic duct is longer than the right and has a greater propensity for dilatation as a consequence of distal obstruction. The two ducts join close to their emergence from the liver to form the common hepatic duct. The common hepatic duct typically extends 1 to 4 cm, has a diameter of approximately 4 mm, and lies anterior to the portal vein and to the right of the hepatic artery.

The cystic duct exits the gallbladder and joins the common hepatic duct at an acute angle to form the common bile duct. The segment of the cystic duct immediately adjacent to the gallbladder neck bears a variable number of mucosal folds called the spiral valves of Heister. While they do not have any valvular function, they can make cannulation of the cystic duct difficult. The length and course of the cystic duct can be quite variable. It may be short or absent and have a high union with the hepatic duct, or it may be long and running parallel to, behind, or spiraling around to the common hepatic duct before joining it, sometimes as far distally as at the duodenum. Variations of the cystic duct and its point of union with the common hepatic duct are surgically important and misidentification can lead to bile duct injuries (Fig. 32-2).

The union of the cystic duct and the common hepatic duct marks the start of the common bile duct. This segment is typically about 7 to 11 cm in length and 5 to 10 mm in diameter,
though its diameter can increase slightly with age and following cholecystectomy. The upper third (supraduodenal portion) passes downward in the free edge of the hepatoduodenal ligament, to the right of the hepatic artery and anterior to the portal vein. The middle third (retroduodenal portion) of the common bile duct curves behind the first portion of the duodenum and diverges laterally from the portal vein and the hepatic arteries. The lower third (pancreatic portion) can curve behind the head of the pancreas in a groove, or traverse through it to enter the wall of the second portion of the duodenum. The duct then runs obliquely downward within the wall of the duodenum for 1 to 2 cm before opening on a papilla of mucous membrane (ampulla of Vater), about 10 cm distal to the pylorus.

The union of the common bile duct and the main pancreatic duct follows one of three configurations. In about 70% of people, these ducts unite outside the duodenal wall and traverse the duodenal wall as a single duct. In about 20%, they join within the duodenal wall and have a short or no common duct, but open through the same opening into the duodenum. In about 10%, they exit via separate openings into the duodenum, termed pancreas divisum. The sphincter of Oddi, a thick coat of circular smooth muscle, surrounds the common bile duct at the ampulla of Vater (Fig. 32-3). It controls the flow of bile, and in some cases pancreatic juice, into the duodenum.

The extrahepatic bile ducts are lined by a columnar mucosa with numerous mucous glands that are concentrated in the common bile duct. A fibro areolar tissue containing scant smooth muscle cells surrounds the mucosa. A distinct muscle layer is not present in the human common bile duct. The arterial supply to the bile ducts is derived from the gastroduodenal and the right hepatic arteries, with major trunks running along the medial and lateral walls of the common duct (sometimes referred to as 3 o’clock and 9 o’clock). The nerve supply to the common bile duct is the same as for the gallbladder, with the density of nerve fibers and ganglia increasing near the sphincter of Oddi.1,2

Figure 32-2. Variations of the cystic duct anatomy. A. Low junction between the cystic duct and common hepatic duct. B. Cystic duct adherent to the common hepatic duct. C. High junction between the cystic and the common hepatic duct. D. Cystic duct drains into right hepatic duct. E. Long cystic duct that joins common hepatic duct behind the duodenum. F. Absence of cystic duct. G. Cystic duct crosses posterior to common hepatic duct and joins it anteriorly. H. Cystic duct courses anterior to common hepatic duct and joins it posteriorly.

Figure 32-3. The sphincter of Oddi.
Anatomic Variants

The classic description of the extrahepatic biliary tree and its arteries applies only in about one-third of patients. The gallbladder may have abnormal positions, be intrahepatic, be rudimentary (a small, nonfunctional hypoplastic remnant), or have anomalous forms or duplications. A partially or completely intrahepatic gallbladder is associated with an increased incidence of cholelithiasis, and may be encountered at the time of cholecystectomy. Isolated congenital absence of the gallbladder is very rare, with a reported incidence of 0.03%. Before the diagnosis is made, the presence of an intrahepatic gallbladder or anomalous position must first be ruled out. Duplication of the gallbladder with two separate cavities and two separate cystic ducts has an incidence of about one in every 4000 persons. This occurs in two major varieties: the more common form in which each gallbladder has its own cystic duct that empties independently into the same or different parts of the extrahepatic biliary tree, and the less common variant in which the two cystic ducts merge before they enter the common bile duct. Duplication is only clinically important when some pathologic process affects one or both organs. Even rarer variants include a left-sided gallbladder (often with a cystic duct that empties into the left hepatic or common bile duct), retrodisplacement of the gallbladder at the posterior-inferior surface of the liver, transverse positioning of the gallbladder, or a floating gallbladder in which the gallbladder is hanging by a mesentery (Fig. 32-4).

Additional small bile ducts (of Luschka) may drain directly from the liver fossa into the body of the gallbladder. If present, but not recognized at the time of a cholecystectomy, a bile leak and subsequent accumulation of bile (biloma) may occur in the abdomen. An accessory right hepatic duct occurs in about 5% of cases. Variations in how the common bile duct enters the duodenum are described earlier, in the “Bile Ducts” section.

Anomalies of the hepatic artery and the cystic artery are quite common, occurring in as many as 50% of cases. While the right hepatic artery usually originates from the proper hepatic branch of the celiac trunk, up to 20% of patients will have a replaced right hepatic artery coming off the superior mesenteric artery. In about 5% of cases, there are two right hepatic arteries, one from the proper hepatic artery and the other from the superior mesenteric artery (accessory right hepatic artery). While the right hepatic artery typically runs posterior to the bile ducts, variations may allow it to course anterior to the common duct, making it vulnerable during surgical procedures, particularly if it runs parallel to the cystic duct or in the mesentery of the gallbladder. The cystic artery arises from the right hepatic artery in about 90% of cases, but it may arise from the left hepatic, common hepatic, gastroduodenal, or superior mesenteric arteries (Fig. 32-5).

Figure 32-5. Variations in the arterial supply to the gallbladder. A. Cystic artery from right hepatic artery, about 80% to 90%. B. Cystic artery off the right hepatic artery arising from the superior mesenteric artery (accessory or replaced), about 10%. C. Two cystic arteries, one from the right hepatic, the other from the common hepatic artery, rare. D. Two cystic arteries, one from the right hepatic, the other from the left hepatic artery, rare. E. The cystic artery branching from the right hepatic artery and running anterior to the common hepatic duct, rare. F. Two cystic arteries arising from the right hepatic artery, rare.

PHYSIOLOGY

Bile Formation and Composition

The liver produces bile continuously and excretes it into the bile canaliculi. Bile leaves the liver through the right and left hepatic ducts, into the common hepatic duct and then the common bile duct. With an intact sphincter of Oddi, tonic contraction diverts bile flow into the gallbladder for storage, while mealtime stimulation allows for its passage into the duodenum. The normal adult consuming an average diet produces 500 to 1000 mL of bile a day. The secretion of bile is responsive to neurogenic, hormonal, and chemical stimuli. Parasympathetic stimulation from the hepatic branches of the vagus nerve increases secretion of bile, whereas sympathetic nerve stimulation via the celiac plexus results in decreased bile flow. Hydrochloric acid, partly digested proteins, and fatty acids entering the duodenum from the stomach after a meal stimulate the release of secretin from the S-cells of the duodenum, and increases bile production and flow.

Bile is mainly composed of water, mixed with bile salts and acids, cholesterol, phospholipids (lecithin), proteins, and bilirubin. It also contains several minor components such as
electrolytes and vitamins. Sodium, potassium, calcium, and chlorine have the same concentration in bile as in plasma or extracellular fluid. The pH of hepatic bile is usually neutral or slightly alkaline, though a high protein diet will shift the bile to a more acidic pH. The primary bile salts, cholate and chenoxycholate, are synthesized in the liver from cholesterol metabolism. They are conjugated with taurine and glycine and act within the bile as anions (bile acids) that are balanced by sodium. These bile acids are then excreted into the bile by hepatocytes and aid in the digestion and absorption of fats in the intestines. About 80% of the secreted conjugated bile acids are reabsorbed in the terminal ileum. The remainder is dehydroxylated (deconjugated) by gut bacteria, forming the secondary bile acids deoxycholate and lithocholate. These are absorbed in the colon and can then be transported back to the liver. Eventually, about 95% of the bile acid pool is reabsorbed, the so-called enterohepatic circulation. Only a small amount (5%) is excreted in the stool, allowing the relatively small quantity of bile acids produced to have maximal effect.

The color of the bile is due to the presence of the pigment bilirubin (orange or yellow) and its oxidized form, biliverdin (green), which are the metabolic products of the breakdown of hemoglobin, and are present in bile in concentrations 100 times greater than in plasma. Bilirubin conjugated in the liver can be excreted through the urine as urobilinogen (yellow). Remaining excess bile pigment passes into the intestines where bacteria convert it into stercobilinogen (brown), which is excreted through the stool.

**Gallbladder Function**

The gallbladder, bile ducts, and the sphincter of Oddi act together to store and regulate the flow of bile. The main function of the gallbladder is to concentrate and store hepatic bile in order to deliver it in a coordinated fashion to the duodenum in response to a meal.

**Absorption and Secretion.** In the fasting state, approximately 80% of the bile secreted by the liver is stored in the gallbladder. This storage is made possible by the fact that the gallbladder mucosa has the greatest absorptive power per unit area of any structure in the body. It rapidly absorbs sodium, chloride, and water against significant gradients, concentrating the bile as much as 10-fold and leading to a marked change in bile composition. This rapid absorptive capacity is one of the protective mechanisms that prevent a potentially dangerous rise in pressure within the biliary system as bile is produced and stored. In addition, gradual relaxation of the gallbladder as well as routine emptying of the gallbladder’s excess bile stores during the fasting period also play a role in maintaining a low resting intraluminal pressure in the biliary tree.

The mucosal cells of the gallbladder itself secrete at least two important products into the gallbladder lumen: glycoproteins and hydrogen ions. The mucosal glands in the infundibulum and the neck of the gallbladder secrete mucus glycoproteins that are believed to protect the mucosa from the corrosive action of bile and to facilitate the passage of bile through the cystic duct. This same mucus creates the colorless “white bile” seen in hydrops of the gallbladder as a result of cystic duct obstruction blocking the entry of bile pigments into the gallbladder. The transport of hydrogen ions by the gallbladder epithelium also plays an important role in decreasing the pH of stored bile. This acidification helps prevent the precipitation of calcium salts, which can act as a nidus for stone formation.

**Motor Activity.** Normal gallbladder filling is facilitated by tonic contraction of the sphincter of Oddi, which creates a small but effective pressure gradient between the bile ducts and the gallbladder. In association with phase II of the interdigestive migrating myenteric motor complex (MMC) in the gut, the gallbladder repeatedly empties small volumes of bile into the duodenum. This process is mediated at least in part by the hormone motilin. In response to a meal, the gallbladder delivers larger volumes to the intestine by a combination of gallbladder contraction and synchronized sphincter of Oddi relaxation. One of the main stimuli to this coordinated effort of gallbladder emptying is the hormone cholecystokinin (CCK). CCK is released endogenously from the enteroeendocrine cells in the duodenum in response to a meal. When stimulated by eating, the gallbladder empties 50% to 70% of its contents within 30 to 40 minutes. Over the following 60 to 90 minutes, the gallbladder gradually refills as CCK levels drop. Other minor hormonal and neural pathways also are involved in the coordinated action of the gallbladder and the sphincter of Oddi. Defects in the motor activity of the gallbladder that inhibit correct emptying are thought to play a role in cholesterol nucleation and gallstone formation.

**Neurohormonal Regulation.** Neurally mediated reflexes are very important in maintaining the functions of the gallbladder, sphincter of Oddi, stomach, and duodenum to coordinate the flow of bile into the intestines at the correct times. The vagus nerve stimulates contraction of the gallbladder by parasympathetic innervation while splanchnic sympathetic nerves from the celiac plexus are inhibitory to its motor activity. For this reason, parasympathomimetic or cholinergic drugs, including nicotine and caffeine, contract the gallbladder. Conversely, anticholinergic drugs such as atropine lead to gallbladder relaxation. Antral distention of the stomach causes both gallbladder contraction and relaxation of the sphincter of Oddi.

In addition to neural inputs, hormonal receptors are located on the smooth muscles, vessels, nerves, and epithelium of the gallbladder and biliary tree. CCK is a peptide that comes from the enteroeendocrine cells of the duodenum and proximal jejunum. CCK is released into the bloodstream in response to the presence of hydrochloric acid, fat, and amino acids in the duodenum. CCK has a plasma half-life of 2 to 3 minutes and is metabolized by both the liver and the kidneys. CCK acts directly on smooth muscle receptors of the gallbladder and stimulates gallbladder contraction. It also relaxes the terminal bile duct, the sphincter of Oddi, and the duodenum to allow forward bile flow. CCK stimulation of the gallbladder and biliary tree is also mediated by cholinergic vagal neurons. For this reason, patients who have had a vagotomy may have a diminished response to CCK stimulation, resulting in an increase in the size and volume of the gallbladder.

Hormones such as vasoactive intestinal polypeptide (VIP) and somatostatin are potent inhibitors of gallbladder contraction. Patients treated with somatostatin analogues and those with somatostatinomas have a high incidence of gallstones, presumably due to the inhibition of gallbladder contraction and emptying. Other hormones such as substance P and enkephalin affect gallbladder motility, but their exact physiologic role is less clear.

**Sphincter of Oddi**

The sphincter of Oddi regulates the flow of bile and pancreatic juice into the duodenum, prevents the regurgitation of duodenal contents into the biliary tree, and diverts bile into
the gallbladder. It is a complex structure that is functionally independent from the duodenal musculature and creates a high-pressure zone between the bile duct and the duodenum. The sphincter of Oddi spans approximately 4 to 6 mm in length and has a basal resting pressure of about 13 mmHg above the duodenal pressure. On manometry, the sphincter shows phasic contractions with a frequency of about four per minute and amplitude of 12 to 140 mmHg. The spontaneous motility of the sphincter of Oddi is regulated by the interstitial cells of Cajal through intrinsic and extrinsic inputs from hormones and neurons acting on the smooth muscle cells. Relaxation occurs in response to raising levels of the gastrointestinal hormones CCK, glucagon, and secretin. This leads to diminished amplitude of phasic contractions and reduced basal pressure of the sphincter, allowing increased flow of bile into the duodenum. During fasting, the sphincter of Oddi activity is coordinated with the periodic partial gallbladder emptying that occurs during phase II of the migrating myoelectric motor complexes. Pharmacologic administration of certain gastrointestinal hormones, such as glucagon, can temporarily decrease sphincter of Oddi baseline pressure and facilitate diagnostic studies.

**DIAGNOSTIC STUDIES**

A variety of diagnostic modalities are available for the patient with suspected disease of the gallbladder or bile ducts. In 1924, the diagnosis of gallstones was revolutionized by the introduction of oral cholecystography by Graham and Cole. For decades, it was the mainstay of investigation for gallstones. It involved oral administration of a radiopaque compound that is absorbed, excreted by the liver, and passed into the gallbladder. Stones are noted on a film as filling defects in a visualized, pacified gallbladder. In the later half of the 20th century, biliary imaging improved dramatically with the development of hepatobiliary scintigraphy (radionucleotide scanning), as well as transhepatic and endoscopic retrograde cholangiography (ERCP), which allowed for more detailed imaging of the biliary tree. Later, ultrasonography, computed tomography (CT), and magnetic resonance imaging (MRI) would further improve the ability to image the biliary tract.

**Blood Tests**

When patients with suspected diseases of the gallbladder or the extrahepatic biliary tree are evaluated, a complete blood count and liver function tests are routinely requested. An elevated white blood cell (WBC) count may indicate or raise suspicion of acute cholecystitis (infection within the gallbladder). If associated with an elevation of bilirubin, alkaline phosphatase, and transaminases, cholangitis (infection within the biliary tree) should be suspected. Cholestasis (an obstruction to bile flow) is generally characterized by an elevation of conjugated bilirubin and a rise in alkaline phosphatase, but it may have no transaminitis. Such a pattern may suggest choledocholithiasis (stones in the common bile duct) or an obstructing lesion such as a stricture or cholangiocarcinoma. In patients with simple symptomatic choledolithiasis, biliary colic, or chronic cholecystitis (a chronic inflammatory state of the gallbladder without infection), blood tests will often be normal.

**Transabdominal Ultrasonography**

Transabdominal ultrasound is the initial investigation of any patient suspected to have disease of the biliary tree. It is noninvasive, painless, does not submit the patient to radiation, and can be performed on critically ill patients. Adjacent organs can also frequently be examined at the same time. However, its reliability and interpretation are dependent upon the skills and experience of the operator. In addition, obese patients, patients with ascites, and patients with distended bowel may be difficult to examine with ultrasound as the quality of the images obtained in these situations can be poor.

Ultrasound will show stones in the gallbladder with sensitivity and specificity of >90%, and can also reliably detect other pathologies of the biliary tree. Stones are acoustically dense and reflect the ultrasound waves back to the ultrasonic transducer. Because stones block the passage of sound waves to the region behind them, they also produce an acoustic shadow (Fig. 32-6). Stones move with changes in position. Polyps, on the other hand, may be calcified and reflect shadows, but they do not move with change in posture. Some stones form a layer in the gallbladder; others a form sediment or sludge. A thickened gallbladder wall, pericholecystic fluid, and local tenderness with direct pressure by the ultrasound probe over the fundus of the gallbladder (sonographic Murphy’s sign) may indicate acute cholecystitis. When a stone obstructs the neck of the gallbladder, the gallbladder may become very large, but thin walled. A contracted, thick-walled gallbladder can be indicative of chronic cholecystitis.

The extrahepatic bile ducts are also well visualized by transabdominal ultrasound, with the exception of the retroduodenal portion. Dilation of the biliary tree in a patient with jaundice suggests an extrahepatic obstruction as the cause for the jaundice. Frequently, the site and, sometimes, the cause of the obstruction can be determined by ultrasound. Small stones in the common bile duct frequently get lodged at the distal end of it, behind the duodenum, and are, therefore, difficult to detect. A dilated common bile duct on ultrasound, small stones in the gallbladder, and a classic clinical presentation allows one to assume that a stone or stones are causing the obstruction. Periampullary tumors can be difficult to diagnose on ultrasound, but above the retroduodenal portion, the level of obstruction and the cause may be visualized quite well. Ultrasound can also be helpful in evaluating tumor invasion and flow in the portal vein, an important guideline for resectability of periampullary tumors.

**Computed Tomography**

Abdominal CT scans are frequently used in the workup of undifferentiated abdominal pain and thus often diagnose gallbladder...
disease. CT scanning is inferior to ultrasonography in diagnosing gallstones but is similar in sensitivity for acute cholecystitis. The major application of CT scan, however, is to define the course and status of the extrahepatic biliary tree and adjacent structures, and to evaluate for alternate causes of a patient's clinical presentation. CT is also the initial test of choice in evaluating patients with suspected malignancy of the gallbladder, the extrahepatic biliary system, or nearby organs such as the head of the pancreas. Use of CT scan is an integral part of the differential diagnosis of obstructive jaundice of unknown origin (Fig. 32-7).

Hepatobiliary Scintigraphy
Hepatobiliary scintigraphy, or hepatobiliary iminodiacetic acid (HIDA) scanning, is another option for noninvasive evaluation of the liver, gallbladder, bile ducts, and duodenum that provides both anatomic and functional information. Technetium-labeled derivatives of iminodiacetic acid are injected intravenously, taken up by the Kupffer cells in the liver, and excreted in the bile. Uptake by the liver is usually detected within 10 minutes, and the gallbladder, bile ducts, and duodenum are typically visualized within 60 minutes in fasting subjects. The primary use of biliary scintigraphy is in the diagnosis of acute cholecystitis, which appears as a nonvisualized gallbladder, with prompt filling of the common bile duct and duodenum. The lack of gallbladder filling is due to inflammatory closure of the cystic duct preventing bile backflow into the gallbladder (Fig. 32-8). Evidence of cystic duct obstruction on biliary scintigraphy is highly diagnostic for acute cholecystitis. The sensitivity and specificity for the diagnosis are about 95% each. False-positive results can occur in patients in the nonfasting state, those receiving parenteral nutrition, or in the setting of gallbladder stasis, recent narcotic use, or alcoholism. Filling of the gallbladder and common bile duct with delayed or absent filling of the duodenum indicates an obstruction at the ampulla. Biliary leaks as a complication of surgery of the gallbladder or the biliary tree can be confirmed and frequently localized by biliary scintigraphy.

Magnetic Resonance Imaging
Available since the mid-1990s, MRI provides anatomic details of the liver, gallbladder, and pancreas similar to those obtained from CT. Many MRI techniques (i.e., heavily T2-weighted sequences, pulse sequences with or without contrast materials) can generate high-resolution anatomic images of the biliary tree and the pancreatic duct. MRI with magnetic resonance cholangiopancreatography (MRCP) offers a focused, noninvasive test for the diagnosis of biliary tract and pancreatic disease (Fig. 32-9). It has a sensitivity and specificity of 95% and 89%, respectively, for detecting choledocholithiasis. In many centers, MRCP is the preferred imaging modality for precise evaluation of biliary and pancreatic duct pathology, reserving endoscopic retrograde cholangiopancreatography (ERCP) for therapeutic purposes only.

Figure 32-7. Computed tomography scan of the upper abdomen from a patient with cancer of the distal common bile duct. The cancer obstructs the common bile duct as well as the pancreatic duct. 1 = the portal vein; 2 = a dilated intrahepatic bile duct; 3 = dilated cystic duct and the neck of the gallbladder; 4 = dilated common hepatic duct; 5 = the bifurcation of the common hepatic artery into the gastroduodenal artery and the proper hepatic artery; 6 = dilated pancreatic duct; 7 = the splenic vein.

Figure 32-8. HIDA scanning. A. Normal HIDA scan showing filling of the extrahepatic biliary tree and gallbladder (white arrow). B. HIDA scan in a patient with acute cholecystitis showing no filling of the gallbladder.
Endoscopic Retrograde Cholangiopancreatography

While the use of endoscopic retrograde cholangiopancreatography (ERCP) in biliary disease is particularly valuable for its therapeutic capabilities, its diagnostic role should not be overlooked. Using a side-viewing endoscope, the common bile duct can be cannulated through the ampulla of Vater and a cholangiogram performed using fluoroscopy (Fig. 32-10). The procedure requires at least intravenous (IV) sedation and in some cases general anesthesia. The advantages of ERCP include direct visualization of the ampullary region and direct access to the distal common bile duct for cholangiography or choledochoscopy. The test is rarely needed for uncomplicated gallstone disease. However, for cases of choledocholithiasis, obstructive jaundice, biliary strictures, or cholangitis, ERCP has the advantage of being both diagnostic and therapeutic. If ductal stones are identified on the endoscopic cholangiogram, biliary sphincterotomy and stone extraction can be performed, clearing the common bile duct of stones. If another etiology such as a biliary stricture is found, diagnostic brushings can be obtained at the time of the procedure. In the hands of experts, the success rate of common bile duct cannulation and cholangiography is >90%. Notable complications of diagnostic ERCP include pancreatitis, which occurs in approximately 3.5% of patients, as well as rare occurrences of bleeding, perforation, or infection (cholangitis).16,17

Endoscopic Choledochoscopy

The development of small fiber-optic cameras that can be threaded through endoscopes used for endoscopic retrograde cholangiopancreatography (ERCP) has facilitated the development of intraductal endoscopy. By providing direct visualization of the biliary and pancreatic ducts, this technology has been shown to increase the effectiveness of ERCP in the diagnosis of certain biliary diseases.18 Intraductal endoscopy has been shown to have therapeutic applications that include biliary stone lithotripsy and directed stone extraction in high-risk surgical patients.19 It can also allow for direct visualization and sampling of concerning lesions in order to evaluate for malignancy (Fig. 32-11). Studies have thus far shown intraductal endoscopy to be safe and effective, though complications such as bile duct perforation, minor bleeding, and cholangitis have
been described. Further refinement of this technology will likely enhance ERCP as a diagnostic and therapeutic tool.

**Endoscopic Ultrasound**
Endoscopic ultrasound (EUS) has improved significantly in recent years and offers additional diagnostic utility to the workup of biliary disease. It requires a specialized 30° endoscope with either a radial or linear ultrasound transducer at its tip. The results are operator dependent and require a skilled endoscopist but offer noninvasive imaging of the bile ducts and adjacent structures. Endoscopic ultrasound can also be used to identify choledocholithiasis. It is useful for evaluation of the retroduodenal portion of the bile duct, which is difficult to visualize with transabdominal ultrasonography. Although EUS is less sensitive than ERCP for biliary stones, the technique is less invasive as it does not require cannulation of the sphincter of Oddi. EUS is also of particular value in the evaluation of tumors near or behind the duodenum and their resectability. Using a linear EUS scope that has a biopsy channel, fine-needle aspiration (FNA) of tumors or lymph nodes, therapeutic injections, or drainage procedures under direct ultrasonic guidance can be performed.

**Percutaneous Transhepatic Cholangiography**
In settings in which the biliary tree cannot be accessed endoscopically, antegrade cholangiography can be performed by accessing the intrahepatic bile ducts percutaneously with a small needle under fluoroscopic guidance. Once the position in a bile duct has been confirmed, a guidewire is inserted and a catheter is passed over the wire (Fig. 32-12). Through the catheter, an antegrade cholangiogram can be obtained and therapeutic interventions such as tissue sampling, biliary drain insertions, or stent placements performed. Percutaneous transhepatic cholangiography (PTC) can also be performed through previously placed percutaneous biliary drainage tubes, if present. PTC has little role in the management of patients with uncomplicated gallstone disease but can be useful in patients with bile duct strictures or tumors, as it can define the anatomy of the biliary tree proximal to the affected segment. As with any invasive procedure, there are potential risks. For PTC, these are mainly bleeding, cholangitis, bile leak, and other catheter-related problems.

**GALLSTONE DISEASE**

**Prevalence and Incidence**
Gallstone disease (cholelithiasis) is one of the most common afflictions of the digestive tract. Autopsy reports show that gallstones are present in between 10% and 15% of adults. The prevalence of gallstones is related to many factors, including diet, age, gender, BMI, and ethnic background with increased prevalence in patients of Native American and Latin American descent. Certain conditions also predispose to the development of gallstones including pregnancy, non-HDL hyperlipidemia, Crohn’s disease, and certain blood disorders such as hereditary spherocytosis, sickle cell disease, and thalassemia. Surgeries that alter the normal neural or hormonal regulation of the biliary tree including terminal ileal resection and gastric or duodenal surgery increase the risk of cholelithiasis. Rapid weight loss following bariatric surgery or lifestyle changes can also precipitate gallstone formation by creating an imbalance in bile composition. Medications such as somatostatin analogues and estrogen-containing oral contraceptives are also associated with an increased risk of developing gallstones. Women are three times more likely to develop gallstones than men, and first-degree relatives of patients with gallstones have a twofold greater prevalence, possibly indicating a genetic predisposition.

**Natural History**
Despite the high prevalence of cholelithiasis, most patients will remain asymptomatic from their gallstones throughout life. For unknown reasons, some patients progress to a symptomatic stage, with typical symptoms of postprandial right upper quadrant pain (biliary colic) caused by a stone obstructing the cystic duct. In addition to pain, gallstones may progress to cause complications such as acute cholecystitis, choledocholithiasis, cholangitis, gallstone pancreatitis, gallstone ileus, and gallbladder cancer. Rarely, one of these complications of gallstones may be the initial presenting picture.

Gallstones in patients without biliary symptoms are commonly diagnosed incidentally during unrelated abdominal imaging or at the time of surgery for an unrelated diagnosis. Several studies have examined the likelihood of developing biliary colic or developing significant complications of gallstone disease after incidental diagnosis in the asymptomatic patient. About 80% of these patients will remain symptom free. However, 2% to 3% will become symptomatic per year (i.e., develop biliary colic). Once symptomatic, patients tend to have recurring bouts of biliary colic. Complicated gallstone disease (cholecystitis, choledocholithiasis, gallstone pancreatitis, etc.) develops in 3% to 5% of symptomatic patients per year. Because few patients develop complications without previous biliary symptoms, prophylactic cholecystectomy in asymptomatic persons with gallstones is rarely indicated. Exceptions exist for individuals who will be isolated from medical care for extended periods of time, or in populations with increased risk of gallbladder cancer, in which case a prophylactic cholecystectomy may be advisable. The presence of porcelain gallbladder, marked by significant calcifications thought to be related to
gallstones, is a rare premalignant condition and is an absolute indication for cholecystectomy, even when asymptomatic.

**Gallstone Formation**
Gallstones form as a result of solids settling out of solution. The major organic solutes in bile are bilirubin, bile salts, phospholipids, and cholesterol. Gallstones are classified by their cholesterol content as either cholesterol stones or pigment stones. Pigment stones can be further classified as either black or brown. In Western countries, about 80% of gallstones are cholesterol stones and about 15% to 20% are black pigment stones. Brown pigment stones account for only a small percentage. Both types of pigment stones are more common in Asia.

**Cholesterol Stones.** Pure cholesterol stones are uncommon and account for <10% of all stones. They usually occur as a single large stone with a smooth surface. The majority of cholesterol stones are mixed but are at least 70% cholesterol by weight in addition to variable amounts of bile pigments and calcium. These stones are usually multiple, of variable size, and may be

---

**Figure 32-12.** Schematic diagram of percutaneous transhepatic cholangiogram and drainage for obstructing proximal cholangiocarcinoma. 

A. Dilated intrahepatic bile duct is entered percutaneously with a fine needle. 
B. Small guidewire is passed through the needle into the duct. 
C. A plastic catheter has been passed over the wire, and the wire is subsequently removed. A cholangiogram can be performed through the catheter. 
D. An external drainage catheter in place. 
E. Long wire placed via the catheter and advanced past the tumor and into the duodenum. 
F. Internal stent has been placed through the tumor.
hard and faceted or irregular, multilobed, and soft (Fig. 32-13). Colors range from whitish yellow to green or black. Most cholesterol stones (>90%) are radiolucent, though some have a high calcium carbonate component and become radioopaque.

The primary event in the formation of cholesterol stones is supersaturation of bile with cholesterol. Cholesterol is highly nonpolar and its solubility in water and bile depends on the relative concentration of cholesterol, bile salts, and lecithin (the main phospholipid in bile). Cholesterol is secreted into bile and is surrounded by bile salts and phospholipids to form a soluble vesicle complex. When cholesterol hyperssecretion is present, either through increased intake or dysfunctional processing, supersaturation occurs. When cholesterol concentrations exceed the ability of the bile salts and phospholipids to maintain solubility, the cholesterol precipitates out of solution into a solid, forming a cholesterol stone (Fig. 32-14).26 Cholesterol hyperssecretion is almost always the cause of supersaturation rather than reduced secretion of phospholipid or bile salts.2

**Pigmented Stones.** Pigmented stones contain <20% cholesterol and are dark because of the presence of calcium bilirubinate. Black and brown pigment stones have little in common and should be considered as separate entities.

Black pigment stones are usually small, brittle, dark, and sometimes spiculated. They are formed by supersaturation of unconjugated bilirubin within the bile. Deconjugation of bilirubin occurs normally in bile at a slow rate. Thus, excessive levels of conjugated bilirubin excretion, as occurs in hemolytic disorders like hereditary spherocytosis and sickle cell disease will lead to an increased rate of production of unconjugated bilirubin. Cirrhosis and hepatic dysfunction may also lead to increased secretion of unconjugated bilirubin directly from the liver. The insoluble unconjugated bilirubin will then precipitate with calcium as insoluble calcium bilirubinate, forming a pigment stone. Due to their high calcium content, pigment stones are often radiopaque. Like cholesterol stones, they almost always form in the gallbladder. In Asian countries such as Japan, black stones account for a much higher percentage of gallstones than in the Western hemisphere.

Brown stones are usually <1 cm in diameter, brownish-yellow, soft, and often mushy. They may form either in the gallbladder or in the bile ducts secondary to bacterial infection and bile stasis. Bacteria such as *Escherichia coli* secrete β-glucuronidase that enzymatically cleaves conjugated bilirubin to produce the insoluble unconjugated bilirubin. This unconjugated bilirubin then precipitates with calcium, and along with dead bacterial cell bodies, forms soft brown stones in the biliary tree. Brown stones are typically found in Asian populations and are associated with stasis secondary to parasite infection with *Ascaris lumbricoides*.

---

**Figure 32-13.** Gallbladder with cholesterol stones. **A.** Stones of multiple shapes and sizes. **B.** Solitary large stone. **C.** Multiple stones of varying composition. (Reproduced with permission from Slesinger MH, Fordtran JS: Gastrointestinal Diseases. Philadelphia, PA: Elsevier/Saunders; 1989.)

**Figure 32-14.** The three major components of bile plotted on triangular coordinates, cholesterol, bile salts and phospholipids (lecithin). A given point represents the relative molar ratios of each. The area labeled “micellar liquid” shows the range of concentrations in which cholesterol is fully solubilized. The shaded area directly above this region corresponds to a metastable zone, supersaturated with cholesterol. Above the shaded area, bile has exceeded the solubilization capacity of cholesterol and precipitation of cholesterol crystals and stones occurs.
Symptomatic Gallstones

Symptomatic Cholelithiasis. Patients with symptomatic gallstone disease typically present with recurrent attacks of pain. The pain develops when a stone obstructs the cystic duct, resulting in a progressive increase of tension in the gallbladder wall as it contracts in response to a meal. This postprandial right upper quadrant or epigastric pain is often referred to as biliary colic. If untreated, about two-thirds of these patients will develop chronic noninfectious inflammation of the gallbladder wall, termed chronic cholecystitis. The pathologic changes, which often do not correlate well with symptoms, vary from an apparently normal gallbladder with minor chronic inflammation in the mucosa, to a shrunken, nonfunctioning gallbladder with transmural fibrosis and adhesions to nearby structures. The mucosa is initially normal or hypertrophied but later becomes atrophied, with the epithelium protruding into the muscle coat, leading to the formation of the so-called Aschoff-Rokitansky sinuses.

Clinical Manifestations The chief symptom associated with symptomatic cholelithiasis is pain (biliary colic). The pain is constant and increases in severity over the first half hour or so after a meal and can last 1 to 5 hours. It is located in the epigastrium or right upper quadrant and frequently radiates to the right upper back or between the scapulae (Fig. 32-15). The pain is severe and comes on abruptly, typically during the night or after a fatty meal. It often is associated with nausea and sometimes vomiting. Patients generally suffer discrete, recurrent attacks of pain, between which they feel well. Physical examination may reveal mild right upper quadrant tenderness during an episode of pain. If the patient is pain free, the physical examination is usually unremarkable. Laboratory values, such as WBC count and liver function tests, are usually normal in patients with uncomplicated gallstones.

Atypical presentations of gallstone disease are common and a high index of suspicion for biliary disease must be maintained when evaluating patients with abdominal complaints. Association with meals is present in only about 50% of patients. Some patients report milder attacks of pain but relate it to meals. The pain may be located primarily in the back or the left upper or right lower quadrant. Bloating and belching may be present and associated with the attacks of pain. In patients with atypical presentations, other conditions that may be causing upper abdominal pain should be ruled out, even in the presence of gallstones. These include but are not limited to peptic ulcer disease, gastroesophageal reflux disease, herpes zoster, abdominal wall hernias, inflammatory bowel disease, diverticular disease, pancreatitis, liver disease, renal calculi, pleuritic pain, and cardiac pain.

When the pain lasts >24 hours without resolving, an impacted stone in the cystic duct or acute cholecystitis (see later “Acute Cholecystitis” section) should be suspected. An impacted stone without cholecystitis will result in what is called hydrops of the gallbladder. Bile will be unable to enter the gallbladder due to the obstructed cystic duct, but the gallbladder epithelium will continue to secrete mucus, and the gallbladder will become distended with clear-white mucinous material. The gallbladder may be palpable but usually is not tender. Hydrops of the gallbladder may result in edema of the gallbladder wall, inflammation, infection, and perforation. Although hydrops may persist with few consequences, early cholecystectomy is generally indicated to avoid complications.

Diagnosis The diagnosis of symptomatic cholelithiasis or chronic cholecystitis depends on the presence of typical symptoms and the demonstration of stones on diagnostic imaging. An abdominal ultrasound is the standard diagnostic test for gallstones as it is noninvasive and highly sensitive (see earlier “Ultrasonography” section). Gallstones are occasionally identified on abdominal CT scans that were obtained as part of a broader workup of abdominal pain. In these cases, if the patient has typical symptoms, it is reasonable to proceed with intervention. Stones diagnosed incidentally on CT or plain radiographs in patients without symptoms should be left in place. Occasionally, patients with typical attacks of biliary pain have no evidence of stones on ultrasound but have evidence of sludge in the gallbladder. If a patient has attacks of typical biliary pain and sludge is detected, cholecystectomy is warranted.

In addition to sludge and stones, cholesterolosis and adenomyomatosis of the gallbladder may cause typical biliary symptoms and may be detected on ultrasound or CT. Cholesterolosis is caused by the accumulation of cholesterol in macrophages in the
gallbladder lamina propria, either locally or as polyps. It produces the classic studded macroscopic appearance of a “strawberry gallbladder.” Adenomyomatosis (cholecystitis glandularis proliferans) is characterized on microscopy by hypertrophic smooth muscle bundles and by the ingrowths of mucosal glands into the muscle layer (epithelial sinus formation). Granulomatous polyps develop in the lumen at the fundus, and the gallbladder wall is thickened. Septae or strictures may be seen within the gallbladder. In symptomatic patients, cholecystectomy is the treatment of choice for patients with these conditions.29

Treatment. Nonsurgical management of gallstone disease using medications or lithotripsy has had disappointing long-term results. These modalities are not considered to be part of the primary treatment algorithm for gallstone disease.30 Surgical cholecystectomy offers the best long-term results for patients with symptomatic gallstones. About 90% of patients with typical biliary symptoms and stones are rendered symptom free after cholecystectomy. For patients with atypical symptoms such as dyspepsia, flatulence, belching, bloating, and dietary fat intolerance, the results are not as favorable. The laparoscopic approach has been proven to be safe and effective and has become the standard of care for symptomatic gallstone disease, replacing open cholecystectomy in routine cases.30,31

Due to the possibility of developing complications related to gallstone disease, patients with symptomatic cholelithiasis should be offered elective cholecystectomy. While waiting for surgery, and if surgery has to be postponed, the patient should be advised to avoid dietary fats and large meals. Diabetic patients with symptomatic gallstones should be encouraged to have a cholecystectomy promptly, as they are more prone to developing severe acute cholecystitis. Pregnant women with symptomatic gallstones who cannot be managed expectantly with diet modifications can safely undergo laparoscopic cholecystectomy. The operation should be performed during the second trimester if possible.

Acute Cholecystitis. Acute cholecystitis, or infection of the gallbladder, is associated with gallstones in 90% to 95% of cases. Rarely, acalculous cholecystitis can occur, usually in patients with other acute systemic diseases (see later “Acalculous Cholecystitis” section). Obstruction of the cystic duct by a gallstone is the initiating event that leads to gallbladder distention, inflammation, and edema of the gallbladder wall. In <1% of acute cholecystitis, the cause is a tumor obstructing the cystic duct. Why inflammation develops only occasionally with cystic duct obstruction is unknown, but it is probably related to the duration of obstruction. Initially, acute cholecystitis is an inflammatory process, probably mediated by the mucosal toxin lysolecithin, a product of lecithin, as well as bile salts and platelet-activating factor. An increase in prostaglandin synthesis amplifies the inflammatory response. In acute cholecystitis, the gallbladder wall becomes grossly thickened and reddish with subserosal hemorrhages. Pericholecystic fluid often is present. The mucosa may show hyperemia and patchy necrosis. In severe cases, about 5% to 10%, the inflammatory process progresses and leads to ischemia and necrosis of the gallbladder wall. More frequently, the gallstone is dislodged and the inflammation resolves.

Not all episodes of uncomplicated acute cholecystitis involve infection. Secondary bacterial contamination is thought to occur in only 15% to 30% of patients. With some severe infections, gangrenous cholecystitis can develop, and an abscess or perforation may occur. When they happen, perforations are usually contained in the subhepatic space by the omentum and adjacent organs. However, free perforation with peritonitis, intrahepatic perforation with intrahepatic abscesses, and perforation into adjacent organs (duodenum or colon) with cholecystoenteric fistula have been described. When gas-forming organisms are part of the secondary bacterial infection, gas may be seen in the gallbladder lumen and in the wall of the gallbladder on abdominal radiographs and CT scans, an entity called emphysematous cholecystitis.

Clinical Manifestations. About 80% of patients with acute cholecystitis give a history compatible with chronic cholecystitis. Acute cholecystitis often begins as an attack of biliary colic with relapsing and remitting pain in the right upper quadrant or epigastrium that may radiate to the right back or interscapular area. In contrast to biliary colic, the pain of acute cholecystitis does not subside. It is unremitting, may persist for several days, and is usually more severe than the pain associated with uncomplicated gallstone disease. The patient is often febrile, complains of anorexia, nausea, and vomiting, and may be reluctant to move as the inflammatory process creates focal peritonitis. On physical examination, tenderness and guarding are usually present in the right upper quadrant. A mass, the gallbladder and adherent omentum, is occasionally palpable; however, guarding may prevent identification of this. Murphy’s sign, an inspiratory arrest with deep palpation in the right subcostal area, is characteristic of acute cholecystitis.

Laboratory evaluation commonly reveals a mild to moderate leukocytosis (12,000–15,000 cells/mm³). However, a normal WBC does not rule out the diagnosis. An unusually high WBC count (>20,000 cells/mm³) suggests a complicated form of cholecystitis such as gangrenous cholecystitis, perforation, or associated cholangitis. In uncomplicated acute cholecystitis, serum liver chemistries are usually normal, but a mild elevation of serum bilirubin (<4 mg/dL) may be present along with mild elevation of alkaline phosphatase, transaminases, and amylase.28 Severe jaundice is suggestive of obstruction of the bile ducts. This can be a result of common bile duct stones or severe pericholecystic inflammation secondary to impaction of a stone in the infundibulum of the gallbladder that mechanically obstructs the bile duct, known as Mirizzi’s syndrome (Fig. 32-16). In elderly patients and in those with diabetes mellitus, acute cholecystitis may have a subtle

Figure 32-16. Mirizzi’s syndrome. Impaction of a large stone in the neck of the gallbladder causing obstruction at the level of the confluence of the cystic duct and common hepatic duct.
presentation resulting in a delay in diagnosis. These patients may also have higher rates of treatment related morbidity compared to younger and healthier patients.

The differential diagnosis for acute cholecystitis includes but is not limited to peptic ulcer disease, pancreatitis, appendicitis, hepatitis, perihepatitis (Fitz-Hugh–Curtis syndrome), myocardial ischemia, pneumonia, pleuritis, and herpes zoster involving the intercostal nerve.

**Diagnosis** Ultrasonography is considered the most useful initial radiologic test for diagnosing acute cholecystitis, with a sensitivity and specificity of 70% to 90%. Ultrasound is effective at documenting the presence or absence of stones, and it can show gallbladder wall thickening and pericholecystic fluid, both of which are highly suggestive of acute cholecystitis (Fig. 32-17). Focal tenderness over the gallbladder when compressed by the sonographic probe (sonographic Murphy’s sign) also supports the diagnosis of acute cholecystitis. Biliary scintigraphy (HIDA scanning) may be of help in atypical cases if the diagnosis remains in question after initial workup. Lack of filling of the gallbladder after 4 hours indicates an obstructed cystic duct and, in the clinical setting of suspected acute cholecystitis, confirms the diagnosis with a reported sensitivity above 90%.32 Conversely, a normal HIDA scan with clear filling of the gallbladder rules out the diagnosis of acute cholecystitis. CT scans are frequently performed on patients with acute abdominal pain of unknown etiology, as they can evaluate for a number of potential pathologic processes at once. In patients with acute cholecystitis, a CT scan can demonstrate thickening of the gallbladder wall, pericholecystic fluid, and the presence of gallstones, but it is somewhat less sensitive than ultrasonography.

**Treatment** Patients who present with acute cholecystitis should receive IV fluids, broad-spectrum antibiotics, and analgesia. The antibiotics should cover gram-negative enteric organisms as well as anaerobes. Although the inflammation in acute cholecystitis may be sterile in some patients, it is difficult to know who is secondarily infected. Therefore, antibiotics have become a standard part of the initial management of acute cholecystitis in most centers.

Cholecystectomy is the definitive treatment for acute cholecystitis. In the past, the timing of cholecystectomy has been a matter of debate. Early cholecystectomy performed within 72 hours of the onset of the illness is preferred over delayed cholecystectomy that is performed 6 to 10 weeks after initial medical treatment and recuperation. Several studies have shown that unless the patient is unfit for surgery, early cholecystectomy should be recommended as soon as possible, as it offers the patient a definitive solution in one hospital admission, quicker recovery times, similar complication rates, and an earlier return to work.33,34

Laparoscopic cholecystectomy is the procedure of choice for acute cholecystitis. The conversion rate to an open cholecystectomy has fallen in recent years to less than 5% as laparoscopic equipment and experience has improved.35 While laparoscopic cholecystectomy for acute cholecystitis may be more tedious and take longer than an elective cholecystectomy for symptomatic choledolithiasis, the laparoscopic approach remains safe and effective, even in the setting of acute and sometimes severe inflammation. Open cholecystectomy must remain an option in particularly difficult cases, or in patients suspected of having prohibitive intraabdominal adhesions, but it is rarely the primary treatment choice.

When patients are medically unfit for surgery due to the severity of their illness or medical comorbidities, they can be treated with antibiotics and biliary decompression with cholecystostomy tube placement, which is usually effective in stabilizing the patient.36 For those who do recover after cholecystostomy, the tube can be removed once the track is mature (approximately 4 weeks) and cholangiography through it shows a patent cystic duct. Elective laparoscopic cholecystectomy can be scheduled within approximately 6 to 8 weeks, assuming their medical fitness recovers.37 Failure to improve after cholecystostomy may be due to gangrene of the gallbladder or perforation, in which case, damage control surgery may be unavoidable.

**Choledocholithiasis.** Common bile duct (CBD) stones may be small or large, single or multiple, and are found in 6% to 12% of patients with stones in the gallbladder. The incidence increases with age. About 20% to 25% of patients above the age of 60 with symptomatic gallstones have stones in the common bile duct as well as in the gallbladder.38 The vast majority of ductal stones in...
Western countries are formed within the gallbladder and migrate down the cystic duct into the common bile duct. These are classified as secondary CBD stones, in contrast to the primary CBD stones that form in the bile duct itself. Secondary stones are usually cholesterol stones, whereas primary stones are usually of the brown pigment type. The primary stones are associated with biliary stasis and infection, and they are more commonly seen in Asian populations. Biliary stasis leading to the development of primary CBD stones can be caused by biliary strictures, papillary stenosis, tumors, or other (secondary) stones.

**Clinical Manifestations** Choleodochal stones may be silent and often are discovered incidentally. They may cause complete or incomplete obstruction, or they may manifest with cholangitis or gallstone pancreatitis. The typical pain caused by a stone in the bile duct is very similar to that of biliary colic caused by impaction of a stone in the cystic duct. Nausea and vomiting are common. Physical examination may be normal, but mild epigastric or right upper quadrant tenderness as well as mild icterus are common. The symptoms may also be intermittent, such as pain and transient jaundice caused by a stone that temporarily impacts the ampulla but subsequently moves away, acting as a ball valve. A small stone may pass through the ampulla spontaneously with resolution of symptoms. Finally, the stones may become completely impacted, causing severe progressive jaundice. Elevation of serum bilirubin, alkaline phosphatase, and transaminases are commonly seen in patients with bile duct stones. However, in about one-third of patients with common bile duct stones, the liver chemistries are normal, particularly if the obstruction is incomplete or intermittent.

**Diagnosis** Ultrasonography is useful for documenting stones in the gallbladder (if still present), as well as determining the size of the common bile duct. As stones in the bile ducts tend to move down to the distal part of the common duct behind the duodenum, bowel gas can preclude their detection on ultrasonography. A dilated common bile duct (>8 mm in diameter) on ultrasonography in a patient with gallstones, jaundice, and biliary pain is highly suggestive of common bile duct stones. If the presence of bile duct stones is in question, magnetic resonance cholangiopancreatography (MRCP) provides excellent anatomic detail and has a sensitivity and specificity of 95% and 89%, respectively, for detecting choledocholithiasis. Endoscopic retrograde cholangiopancreatography (ERCP) is highly effective at diagnosing choledocholithiasis and in experienced hands, cannulation of the ampulla of Vater and diagnostic cholangiography are achieved in >90% of cases. However, due to the risks associated with the procedure, it is rarely used as a purely diagnostic modality, rather being reserved for cases in which a therapeutic intervention such as stone extraction or sphincterotomy is planned. Endoscopic ultrasound has been demonstrated to be as good as ERCP for detecting common bile duct stones (sensitivity of 95% and specificity of 97%). However, EUS has fewer therapeutic capabilities and requires endoscopic expertise, making it less desirable except in specific clinical scenarios. Percutaneous transhepatic cholangiography (PTC) is rarely needed in patients with common bile duct stones but can be performed for both diagnostic and therapeutic reasons in patients with contraindications to endoscopic or surgical approaches.

**Treatment** For patients with symptomatic gallstones and suspected common bile duct stones, bile duct clearance and cholecystectomy are indicated. This may be safely achieved either with preoperative ERCP followed by surgery or by going directly to surgery with intraoperative cholangiogram and common bile duct exploration to address retained stones. Both approaches are considered safe and effective, and no formal recommendation exists to definitively support one over the other.

If upfront laparoscopic cholecystectomy is pursued, the surgery should include an intraoperative cholangiogram to document the presence or absence of bile duct stones. If stones are identified, laparoscopic common bile duct exploration via the cystic duct or with formal choledochotomy allows the stones to be retrieved in the same setting (see “Cholecystectomy” section). If the expertise and/or instrumentation for laparoscopic common bile duct exploration are not available, the patient can be awoken and scheduled for ERCP with sphincterotomy the following day. An open common bile duct exploration is an option if the endoscopic and laparoscopic methods are not feasible. If a choledochotomy is performed, primary repair can be considered in large ducts, while smaller ducts should be repaired over a T-tube. To do this, a standard T-tube should be modified by cutting the ends short enough to allow placement within the duct and dividing the T longitudinally to facilitate easy removal from the duct later on (Fig. 32-18). If a common

---

**Figure 32-18.** T-tube placement. A. A standard T-tube that has been cut and modified for use in the biliary tract. B. The T-tube is placed through a ductotomy in the common bile duct with the defect closed over the tube. The opposite end is brought out through the abdominal wall for decompression of the bile ducts.
bile duct exploration was performed and a T tube left in place, a T-tube cholangiogram should be obtained before its removal, at least several weeks after its placement.

In very severe cases, stones impacted in the ampulla may be unable to be cleared by endoscopic approaches or common bile duct exploration (open or laparoscopic). In these cases, transduodenal sphincterotomy can be considered. If one is entirely unable to disimpact the duct, choledochooduodenostomy or Roux-en-Y choledochojejunostomy may be the only option to restore biliary continuity.42

If the stones were left in place at the time of surgery or diagnosed shortly after the cholecystectomy, they are classified as retained. Those diagnosed months or years later are termed recurrent (Fig. 32-19). Retained or recurrent stones following cholecystectomy are best treated endoscopically. A generous sphincterotomy will allow for stone retrieval as well as spontaneous passage of stones. Alternately, retained stones can be cleared via a mature T-tube tract (4 weeks) if one was placed at the time of surgery. To do this, the T-tube is removed and a catheter passed through the tract into the common bile duct. Under fluoroscopic guidance, the stones can be retrieved with baskets or balloons. A similar approach will allow for stone clearance by percutaneous transhepatic cholecystostomy (PTC) if there is no other way to reach the duct. Repeat surgery should be a last resort if other interventions have failed.

Cholangitis. Cholangitis is one of the main complications of choledochal stones. Acute cholangitis is an ascending bacterial infection associated with partial or complete obstruction of the bile ducts.43 Hepatic bile is sterile, and bile in the bile ducts is kept sterile by continuous antegrade bile flow and by the presence of antibacterial substances in bile, such as immunoglobulin. Mechanical hindrance to bile flow facilitates ascending bacterial contamination from the bowel. Positive bile cultures are common in the presence of bile duct stones as well as with other causes of obstruction. Biliary bacterial contamination alone does not lead to clinical cholangitis; the combination of both significant bacterial contamination and biliary obstruction is required for its development. Gallstones are the most common cause of obstruction in cholangitis. Other causes include primary sclerosing cholangitis, benign and malignant strictures, parasites, instrumentation of the ducts, and indwelling stents, as well as partially obstructed biliary-enteric anastomoses. The most common organisms cultured from bile in patients with cholangitis include E coli, Klebsiella pneumoniae, Streptococcus faecalis, Enterobacter, and Bacteroides fragilis.43

Clinical Manifestations Cholangitis may present as anything from a mild, self-limited episode to a fulminating, potentially life-threatening septicemia. Patients with gallstone-induced cholangitis are most commonly older and female. The most common presentation is fever, epigastric or right upper quadrant pain, and jaundice. These classic symptoms, known as Charcot’s triad, are present in about two-thirds of patients. The illness can progress rapidly with the development of shock and altered mental status, known as Reynolds’ pentad (e.g., fever, jaundice, right upper quadrant pain, septic shock, and mental status changes). However, the presentation may be atypical, with little if any fever, jaundice, or pain. This occurs most commonly in the elderly, who may have unremarkable symptoms until the process is already quite advanced. Patients with indwelling stents are at particularly high risk for cholangitis, though rarely become jaundiced as a patent stent will prevent the obstruction of bile flow. On abdominal examination, the findings are indistinguishable from those of acute cholecystitis.44

Diagnosis Leukocytosis, hyperbilirubinemia, and elevation of alkaline phosphatase and transaminases are common and, when present, support the clinical diagnosis of cholangitis. Ultrasonography is helpful, as it will document the presence of gallbladder stones, demonstrate dilated ducts, and possibly pinpoint a site of obstruction. CT scanning and MRI can show pancreatic and periampullary masses, if present, in addition to the ductal dilatation. However, abdominal imaging will rarely elucidate the exact cause of cholangitis, and the initial diagnosis is generally made clinically.

Treatment The initial treatment of patients with cholangitis includes broad-spectrum IV antibiotics to cover enteric organisms and anaerobes, fluid resuscitation, and rapid biliary...
decompression. This is most often accomplished through ERCP and sphincterotomy. ERCP will show the level and the reason for the obstruction, allow for culture of the bile, permit the removal of stones if present, and accomplish drainage of the bile ducts. Placement of drainage catheters or stents can also be performed if needed. In cases in which ERCP is not available, PTC, EUS, or surgical drainage can be utilized. The selection of the appropriate approach will depend on the type and location of the suspected obstruction as well as the availability of local resources and expertise. Cholecystostomy tubes are not indicated in the acute management of cholangitis as the primary source of the infection is extrinsic to the gallbladder.

Patients with cholangitis can deteriorate rapidly and may require intensive care unit monitoring and vasopressor support. However, most patients will respond to biliary decompression and supportive measures. In the current era, acute cholangitis is associated with an overall mortality rate of approximately 5%. When associated with renal failure, cardiac impairment, hepatic abscesses, and malignacies, the morbidity and mortality rates are much higher. Patients who have suffered an episode of acute cholangitis related to gallstone disease should be recommended to undergo elective cholecystectomy approximately 6 weeks after the resolution of their cholangitis. Those whose cholangitis was related to another cause of biliary obstruction should be followed and treated for the specific etiology of their obstruction but do not necessarily require cholecystectomy if gallstones were not the causative etiology of their cholangitis. Patients with indwelling stents and cholangitis usually require repeated imaging and stent exchange to mitigate the risk of recurrent infections.

Gallstone Pancreatitis. Gallstones in the common bile duct can provoke attacks of acute pancreatitis through transient or persistent obstruction of the pancreatic duct by a stone passing through or impacted in the ampulla. The exact mechanism by which obstruction of the pancreatic duct leads to pancreatitis is unclear, but it may be related to increased ductal pressures causing leakage of pancreatic enzymes into the glandular tissue. The initial management of gallstone pancreatitis is supportive, including admission for bowel rest, IV hydration, and pain control. Antibiotics are not indicated in the absence of signs of infected pancreatic necrosis. Imaging of the biliary tree with ultrasound, CT, or MRCP is essential to confirm the diagnosis. When gallstones are present and the pancreatitis is mild and self-limited, the stone has probably passed. For these patients, a cholecystectomy with intraoperative cholangiogram is indicated as soon as the pancreatitis has clinically resolved. It is strongly recommended that cholecystectomy be performed during the same admission whenever possible due to the high rate of recurrence and increased morbidity of subsequent attacks of pancreatitis. If gallstones are present obstructing the duct and the pancreatitis is severe, an ERCP with sphincterotomy and stone extraction may be necessary. This must be balanced with the risk of ERCP-induced pancreatitis and thus is usually only employed if supportive measures are failing.

Gallstone Ileus. Gallstone ileus can occur when a large gallstone erodes through the wall of the gallbladder directly into the intestine via a choledochoenteric fistula (Fig. 32-20A). These stones can then pass through the intestinal tract until they reach an area of fixed obstruction. Proximal stones can become impacted in the pylorus or proximal duodenum causing gastric outlet obstruction (Bouveret syndrome). Those that travel distally may become lodged at surgical anastomoses or the ileocecal valve, where they can become impacted and cause small bowel obstruction. Gallstone ileus is responsible for less than 1% of all intestinal obstructions.

These patients present with symptoms of obstipation, nausea, and abdominal pain. Plain films may show an obstructive bowel gas pattern but may fail to identify a radiolucent stone. Ultrasound evaluation may be limited by extensive bowel gas. CT is highly sensitive and specific for gallstone ileus and will help to determine the location of the obstruction. Management of gallstone ileus focuses on relieving the intestinal obstruction and removing the stone. In cases of very proximal obstructions in the stomach or duodenum, endoscopic retrieval can be effective. For more distal stones, surgical enterolitotomy can be accomplished either laparoscopically or open. This procedure entails the removal of the stone through...
an enterotomy that is then either repaired or resected depending on its size (Fig. 32-20B). Stones that have successfully traversed the ileocecal valve are likely to pass without further intervention. The role of pursuing cholecystectomy and/or choledochoenteric fistula closure at the time of enterolithotomy or addressing it at a later time remains a topic of debate, but it should be considered to reduce the risk of recurrence. 47

**Cholangiohepatitis**

Cholangiohepatitis, also known as recurrent pyogenic cholangitis, is endemic to the Orient. It also has been encountered in Asian population in the United States, Europe, and Australia. It affects both sexes equally and occurs most frequently in the third and fourth decades of life. Cholangiohepatitis is caused by bacterial contamination (commonly *E. coli*, *Klebsiella* species, *Bacteroides* species, or *Enterococcus faecalis*) of the biliary tree, and often it is associated with biliary parasites such as *Clonorchis sinensis*, *Opisthorchis viverrini*, and *A. lumbricoides*. Bacterial enzymes cause deconjugation of bilirubin, which precipitates as bile sludge. The sludge and dead bacterial cell bodies form brown pigment stones, the nucleus of which may contain an adult *Clonorchis* worm, an ovum, or an ascariid. These stones can form throughout the biliary tree and cause partial obstructions that contribute to repeated bouts of cholangitis, biliary strictures, further stone formation, infection, hepatic abscesses, or liver failure (secondary biliary cirrhosis). 58

Patients with cholangiohepatitis usually present with pain in the right upper quadrant or epigastrum, fever, and jaundice. Relapsing symptoms are one of the most characteristic features of the disease. The episodes may vary in severity but, without intervention, will gradually lead to malnutrition and hepatic insufficiency. An ultrasound may detect stones in the biliary tree, pneumobilia from infection by gas-forming organisms, liver abscesses, and, occasionally, strictures. The gallbladder may be thickened and inflamed in about 20% of patients but rarely contains gallstones. ERCP or MRCP can be utilized for biliary imaging for cholangiohepatitis. They can detect obstructions and define strictures and stones. ERCP (or PTC if necessary) has the additional benefit of allowing for emergent decompression of the biliary tree in the septic patient. Hepatic abscesses may be drained percutaneously. The long-term goal of therapy is to extract stones and debris and relieve strictures. It may take several procedures, and in severe, refractory cases in which stones and strictures cannot be relieved, it may require a hepatochojunostomy to reestablish biliary–enteric continuity. Occasionally, resection of involved areas of the liver may offer the best form of treatment. Recurrences are common, and the prognosis is poor once hepatic insufficiency has developed. 49

**PROCEDURAL INTERVENTIONS FOR GALLSTONE DISEASE**

**Percutaneous Transhepatic Cholecystostomy Tubes**

In cases in which a patient with cholecystitis is deemed to be too ill to safely undergo cholecystectomy, a cholecystostomy tube may be placed into the gallbladder to decompress and drain a distended, inflamed, hydropic, or purulent gallbladder. 50 Surgical cholecystostomy with a large catheter placed under local anesthesia is rarely required today. Rather, percutaneous transhepatic cholecystostomy (PTC) tubes are most often pigtail catheters inserted percutaneously under ultrasound guidance. 50

![Figure 32-21. Percutaneous cholecystostomy. A pigtail catheter has been placed through the abdominal wall, the right lobe of the liver, and into the gallbladder.](image)

The catheter is inserted over a guidewire that has been passed through the abdominal wall, the liver, and into the gallbladder (Fig. 32-21). By passing the catheter through the liver, the risk of uncontrolled bile leak around the catheter and into the peritoneal cavity is minimized. The catheter can be removed when the inflammation has resolved and the patient’s condition has improved. A patent cystic duct should be confirmed by a tube cholangiogram prior to its removal. Interval cholecystectomy should be considered if the patient’s fitness has improved, particularly in individuals whose etiology of cholecystitis was gallstones.

**Endoscopic Interventions**

Endoscopic advances in the last few decades have made endoscopy and ERCP a valuable therapeutic tool in the management of gallstone disease, particularly in the setting of common bile duct stones or abnormalities. Using a 90-degree side-viewing endoscope, the duodenum can be entered and the ampulla of Vater on the medial wall of the second portion of the duodenum visualized. This can then be cannulated to allow wire and catheter access to the biliary tree, facilitating retrograde cholangiogram, diagnostic brushings, stenting, dilations, or fluoroscopically guided basket or balloon retrieval of common bile duct stones. When CBD stones are present, endoscopic sphincterotomy should be performed, which will allow for passage of larger stones both at the time of bile duct clearance and in the case of any ongoing choledocholithiasis (Fig. 32-22). In the hands of experts, ERCP has high rates of successful cannulation and bile duct clearance, and it is a safe and tolerable procedure. Debate remains when comparing ERCP to surgical common bile duct exploration in terms of timing and outcomes for cholecystectomy, but both are considered acceptable treatments. 41 In special cases, such as the presence of Roux-en-Y anatomy or a previous hepaticojejunostomy, ERCP can be difficult. However, such anatomy does not preclude the option for endoscopic intervention. Laparoscopic-assisted ERCP (in which the remnant stomach is accessed surgically and the endoscope passed into the duodenum) or double-balloon ERCP can be utilized to reach the biliary tree.

**Cholecystectomy**

Cholecystectomy is one of the most common abdominal surgeries performed in Western countries, with over 750,000
being performed each year in the United States alone. Carl Langenbuch performed the first successful open cholecystectomy in 1882, and for >100 years, it was the standard treatment for symptomatic gallbladder stones. In 1987, laparoscopic cholecystectomy was introduced by Philippe Mouret in France and quickly revolutionized the treatment of gallstone disease. It not only supplanted open cholecystectomy, but it also more or less ended attempts for noninvasive management of gallstones (such as extracorporeal shock wave or cholangioscopic lithotripsy) or medical therapies (such as bile salts). Laparoscopic cholecystectomy offers a cure for gallstones with a minimally invasive procedure, minor pain and scarring, and early return to full activity. Today, laparoscopic cholecystectomy is the treatment of choice for symptomatic gallstones and the complications of gallstone disease.

Few absolute contraindications exist to laparoscopic cholecystectomy, but they include hemodynamic instability, uncontrolled coagulopathy, or frank peritonitis. In addition, patients with severe obstructive pulmonary disease (COPD) or congestive heart failure (e.g., cardiac ejection fraction <20%) might not tolerate the increased intraabdominal pressures of pneumoperitoneum with carbon dioxide and may require open cholecystectomy. Conditions formerly believed to be relative contraindications such as acute cholecystitis, gangrene and empyema of the gallbladder, biliary-enteric fistulae, obesity, pregnancy, ventriculoperitoneal shunts, cirrhosis, and previous upper abdominal procedures are now considered risk factors for a potentially difficult cholecystectomy, but they do not preclude an attempt at laparoscopy. While laparoscopic outcomes have steadily improved and laparoscopic cholecystectomy has been shown multiple times to be safe and feasible, conversion to an open operation should always remain an option, and it is not a failure. Conversion to open may be necessary if the patient is unable to tolerate pneumoperitoneum, a complication occurs that cannot be fixed laparoscopically, important anatomic structures cannot be clearly identified, or when no progress is made over a set period of time. In the elective setting, conversion to an open procedure is needed in about 5% of patients. Emergent procedures or patients with complicated gallstone disease can be more challenging, and the incidence of conversion has been reported to be between 10% and 30%. The possibility of conversion to open should always be discussed with the patient preoperatively.

Serious complications of cholecystectomy are rare. The mortality rate for laparoscopic cholecystectomy is about 0.1%. Wound infection and cardiopulmonary complication rates are considerably lower following laparoscopic cholecystectomy than are those for an open procedure. While laparoscopic cholecystectomy has historically been associated with a higher rate of injury to the bile ducts than the open approach, modern data appears to show this trend disappearing as familiarity with laparoscopic techniques and technologies have improved.

Patients undergoing cholecystectomy should have a complete blood count and liver function tests preoperatively. Prophylaxis against deep venous thrombosis with either low molecular weight heparin or compression stockings is indicated.
The patient should be instructed to empty their bladder before coming to the operating room to avoid the need for urinary catheterization. An orogastric tube can be placed if the stomach is distended with gas, but it is generally removed at the end of the operation.

**Laparoscopic Cholecystectomy.** The patient is typically positioned supine with the operating surgeon standing at the patient’s left side. Split-leg positioning with the surgeon standing between the patient’s legs can also provide ergonomic access to the right upper quadrant. Tucking one arm can be helpful if a cholangiogram is planned to allow easier maneuvering of the fluoroscopy machine around the patient. Pneumoperitoneum is established with carbon dioxide gas, either with an open technique (Hasson), optical viewing trocar, or closed-needle technique (Veress). Typical access is at the supraumbilical region, though in the case of previous surgery or scars, alternate access sites should be considered. Once an adequate pneumoperitoneum is established, a 5- or 10-mm trocar is inserted through the supraumbilical incision, through which a 5- or 10-mm 30° laparoscope is introduced. Traditionally, three additional ports are then placed with a 10- or 12-mm port in the epigastrium, a 5-mm port in the right midclavicular line, and a 5-mm port in the right flank (Fig. 32-23). Additional ports may be placed as needed to aid with retraction in difficult cases.

Through the lateral-most port, the assistant uses a locking instrument to grasp the gallbladder fundus and retract it over the liver edge and upward towards the patient’s right shoulder. This will help visualize the body of the gallbladder and the hilar area. Exposure may be facilitated by placing the patient

---

**Figure 32-23.** Laparoscopic cholecystectomy. **A.** The trocar placement. **B.** The fundus has been grasped and retracted cephalad to expose the proximal gallbladder and the hepatoduodenal ligament. Another grasper retracts the gallbladder infundibulum posterolaterally to better expose the triangle of Calot (hepatocystic triangle bound by the common hepatic duct, cystic duct, and liver margin). **C.** Intraoperative photo of the critical view of safety. The hepatocystic triangle has been cleared of fat and fibrous tissue, the lower one-third of the gallbladder is separated from the liver to expose the cystic plate, and two and only two structures are seen entering the gallbladder. **D.** A clip is being placed on the cystic duct–gallbladder junction. **E.** A small opening has been made in the cystic duct, and a cholangiogram catheter is being inserted. **F.** Additional clips have been placed, the cystic duct has been divided, and the cystic artery is being divided.
in reverse Trendelenburg position with slight tilting of the table to bring the right side up. Through the midclavicular port, the surgeon uses a grasper in the left hand to retract the gallbladder infundibulum laterally and expose the neck of the gallbladder and hepatoduodenal ligament. It may be necessary to take down any adhesions between the omentum, duodenum, or colon to the gallbladder in order to reach the infundibulum. The majority of the dissection can then be performed with the right hand through the epigastric port, utilizing a combination of electrocautery and sharp and blunt dissection.

Dissection starts at the infundibulum of the gallbladder, just above the takeoff of the cystic duct. The peritoneum, fat, and loose areolar tissue around the gallbladder and the cystic duct–gallbladder junction is dissected off and reflected inferiorly toward the bile duct. This is continued until the gallbladder neck and the proximal cystic duct are clearly identified. The next step is the identification of the cystic artery, which usually runs parallel to and somewhat behind the cystic duct, and often lies behind a prominent lymph node (Lund’s node, often called Calot’s node). At this point, a critical view of safety should be obtained. This requires that the hepatocystic triangle is cleared of fat and fibrous tissue, the lower third of the gallbladder is separated from the liver to expose the cystic plate, and two and only two structures (cystic duct and cystic artery) are going into the gallbladder (see Fig. 32-19). At this point, an intraoperative cholangiogram can be performed if indicated (see “Intraoperative Cholangiogram” section).

With a critical view of safety obtained, the cystic duct and artery are clipped with two clips at the base and one clip on the gallbladder side. They can then be safely divided. Sometimes, a very dilated cystic duct may be too large for clips. Such ducts can be successfully managed by ligation with an endoloop, laparoscopic stapler, or suture closure. Finally, the gallbladder is dissected off the liver bed using electrocautery while watching for potential abnormal posterior branches of ducts or arteries. Before the gallbladder is completely removed from the liver edge, it can be used as a retractor for a final evaluation of the operative field. The surgeon should be sure to evaluate for bleeding points or bile staining, and confirm placement of the clips on the cystic duct and artery. The gallbladder is then divided from its final attachments and removed either through the epigastric or umbilical incision, often with the aid of a retrieval bag. The fascial defect and skin incision may need to be enlarged in order to remove the specimen, particularly if the stones are large or the gallbladder is very inflamed. Any bile or blood that has accumulated during the procedure should be cleaned away, and if stones were spilled, they should be retrieved and removed. If the gallbladder was severely inflamed or gangrenous, or if any bile or blood is expected to accumulate, a closed-suction drain can be placed through one of the 5-mm ports and left underneath the right liver lobe close to the gallbladder fossa, though this is not routinely required.

Open Cholecystectomy. The same surgical principles apply for laparoscopic and open cholecystectomies. Open cholecystectomy has become an uncommon procedure, usually performed either as a conversion from laparoscopic cholecystectomy or as a second procedure in patients who require laparotomy for another reason. The approach can either be through a midline laparotomy, or more commonly through a right subcostal incision. The gallbladder is dissected free from the liver bed, usually starting at the fundus and working proximally toward the hepatocystic triangle. Once the cystic artery and cystic duct have been dissected and clearly identified, they are ligated and divided, and the gallbladder is removed. In particularly difficult cases, in which the gallbladder is partially obliterated or ductal or arterial anatomy cannot be identified, a partial cholecystectomy may be performed. This includes removal of as much gallbladder mucosa as possible and attempted closure of the cystic duct stump with wide drainage of the area.

Intraoperative Cholangiogram. Intraoperative cholangiogram is an optional but valuable tool for evaluating the extrahepatic bile ducts, identifying common bile duct stones, or clarifying aberrant ductal anatomy. The use of routine versus selective cholangiography remains a topic of debate with a lack of definitive evidence on either side. However, routine intraoperative cholangiography will detect stones in approximately 7% of patients, and it assists with outlining anatomy and detecting injury. Selective intraoperative cholangiogram should be performed when the patient has a history of abnormal liver function tests, pancreatitis, jaundice, a large duct and small stones, a dilated duct on preoperative ultrasonography, or if preoperative endoscopic cholangiography for the aforementioned reasons was unsuccessful. Although there is no consensus recommendation on the use of routine versus selective cholangiography, all surgeons performing cholecystectomy should be familiar with the procedure. If a cholangiogram is to be performed, a clip is placed on the proximal cystic duct, and a small incision is made on its anterior surface, just inferior to the clip. A cholangiogram catheter is passed into the cystic duct and secured with a clamp or clip. The fluoroscopy machine is then positioned over the patient and a cholangiogram performed by injection of contrast through the cholangiocatheter during live fluoroscopic dynamic imaging. An ideal cholangiogram includes filling of the right and left hepatic ducts, emptying into the duodenum, and no visualized filling defects (Fig. 32-24). Care must be taken not to introduce air bubbles into the system during contrast injection as these will appear as filling defects on the cholangiogram images. If no contrast is visualized in the duodenum, a dose of glucagon can be utilized to relax the sphincter of Oddi and facilitate contrast flow. Once the cholangiogram is completed, the catheter is removed. Laparoscopic ultrasonography is as accurate as intraoperative cholangiography in detecting common bile duct stones, and it is less invasive. However, it requires more skill to perform and interpret and is not always readily available.

Common Bile Duct Exploration

Common bile duct stones that are detected pre- or intraoperatively may be managed with common bile duct exploration (CBDE) at the time of the cholecystectomy. While preoperative ERCP is also an appropriate option for known bile duct stones, laparoscopic CBDE can be used as a primary approach to choledocholithiasis safely and with good outcomes, even in higher risk populations such as the elderly. If stones in the duct are small, they may sometimes be simply flushed into the duodenum with saline irrigation via the cholangiography catheter. This can be facilitated by the administration of IV glucagon to relax the sphincter of Oddi. If irrigation is unsuccessful, several options exist to clear the duct, including fluoroscopic or endoscopic approaches.

With access to the cystic duct by a small ductotomy, a balloon catheter is used to dilate the cystic duct, and a wire basket can be passed down the common bile duct under fluoroscopic
guidance to catch and remove the stones (Fig. 32-25). Alternatively, endoscopic evaluation with a flexible choledoscope will allow for direct visualization and retrieval of the stones within the common duct. To do this, reliable catheter access must be obtained with an introducer sheath placed either through one of the laparoscopic ports or a new stab incision in the anterior abdominal wall. The cystic duct should first be dilated with a small balloon catheter to allow for passage of the introducer and scope and for effective retrieval of larger stones. Once the scope is within the common bile duct, irrigation is used to distend the lumen. Stones may then be caught in a wire basket under direct visualization or simply pushed into the duodenum. Once the common bile duct has been cleared of stones, the cystic duct is ligated below the level of the ductotomy and divided, and the cholecystectomy is completed.

While the cystic duct is the preferred route of access for common bile duct exploration, occasionally an incision into the common bile duct itself (choledochotomy) is necessary. The flexible choledochoscope is then passed into the duct for visualization and clearance of stones. The choledochotomy can be closed primarily if the duct is very large, or over a T-tube. If available, common bile duct exploration can be highly advantageous as it provides the opportunity to treat the entirety of the disease in a single event, rather than subjecting patients to multiple procedures. However, the procedure can be technically challenging to perform and requires the availability of the proper equipment and surgical expertise.

Common Bile Duct Drainage Procedures
In very rare cases in which stones or obstructions cannot be cleared by either ERCP with sphincterotomy or CBDE, and the patient is suffering clinical effects from their common duct stones, an additional choledochal drainage procedure may become necessary. In the case of an open operation, transduodenal sphincterotomy can be attempted by incising the duodenum transversely and cutting the sphincter of Oddi at the 11 o’clock position, taking care to avoid injury to the pancreatic duct. The impacted stones can then be manually removed or simply allowed to pass through the sphincterotomy.

Bypass procedures can also be used to restore continuity of bile flow in the setting of irretrievable impacted stones. For short distance bypasses, a Cholecodochoduodenostomy is performed by mobilizing the second part of the duodenum (a Kocher maneuver) and anastomosing it side to side with the common bile duct (Fig. 32-26A-C). If the distance is too great to safely complete a cholecodochoduodenostomy without tension, a choledochojejunostomy can be done by bringing up a roughly 45-cm limb of jejunum and anastomosing it end to side to the common bile duct (Fig. 32-26D-E). If the entirety of the extrahepatic biliary tree must be bypassed, hepaticojejunostomy allows for drainage of the hepatic ducts directly a loop of jejunum (Fig. 32-26F-G). These choledochal drainage procedures can also be used to manage common bile duct strictures or as a palliative procedure for malignant obstruction in the peripancreatic region.

OTHER BENIGN DISEASES AND LESIONS
Biliary Dyskinesia and Sphincter of Oddi Dysfunction
Biliary dyskinesia is an umbrella term that refers to disorders affecting the normal motility and function of the gallbladder and sphincter of Oddi. These disorders are becoming increasingly recognized as improvements in imaging allow for more detailed evaluations of biliary tract function. Patients with biliary dyskinesia may present with typical biliary type symptoms, but without evidence of stones or sludge on abdominal imaging.
A decreased gallbladder ejection fraction on HIDA scanning (EF <35%) is considered diagnostic of biliary dyskinesia. In these patients, studies suggest that symptoms will be improved or resolved by cholecystectomy in up to 90% of cases.\(^\text{62}\)

Sphincter of Oddi dysfunction can occur as a primary presentation of episodic biliary type pain with abnormal liver function tests or as recurrent biliary type pain after cholecystectomy. More severe cases may present with recurrent jaundice or pancreatitis. If other causes are ruled out, such as retained stones, strictures or periampullary tumors, a stenotic or dyskinetic sphincter of Oddi should be suspected. A benign stenosis of the outlet of the common bile duct is usually associated with inflammation, fibrosis, or muscular hypertrophy. The pathogenesis is unclear, but trauma from the passage of stones, sphincter motility disorders, and congenital anomalies have been suggested. A dilated common bile duct that is difficult to cannulate during ERCP or delayed emptying of contrast from the biliary tree after ERCP are useful diagnostic features. Ampullary manometry and specific provocation tests are available in specialized units to aid in the diagnosis. Once identified, sphincterotomy will typically yield good results.\(^\text{63}\)

**Acalculous Cholecystitis**

Acalculous cholecystitis is an acute inflammation of the gallbladder that occurs in the absence of gallstones. It is a rare entity that typically develops in critically ill patients in the intensive care unit.\(^\text{64}\) Patients on parenteral nutrition, with extensive burns, sepsis, major operations, multiple trauma, or prolonged illness with multiple organ system failure are at risk for developing acalculous cholecystitis. The cause is unknown, but gallbladder distention, bile stasis, and ischemia have been implicated as causative factors. After resection, pathologic examination of the gallbladder wall after an episode of acalculous cholecystitis reveals edema of the serosa and muscular layers, with patchy thrombosis of arterioles and venules.\(^\text{65}\)

The ability to recognize the symptoms and signs of acalculous cholecystitis can depend on the condition and mental status of the patient, but acalculous cholecystitis can be similar to acute calculous cholecystitis, with right upper quadrant pain and tenderness, fever, and leukocytosis. In the sedated or unconscious patient, the clinical features are often masked, but fever and elevated WBC count, as well as elevation of alkaline phosphatase and bilirubin, are indications for
Figure 32-26. Biliary enteric anastomoses. There are three types. I. Choledochoduodenostomy. A. The distal common bile duct is opened longitudinally, as is the duodenum. B. Interrupted sutures are placed between the common bile duct and the duodenum. C. Completed choledochoduodenostomy. II. Choledochojejunostomy. D. The common bile duct and small bowel are divided. E. A limb of jejunum is brought up in a Roux-en-Y configuration and anastomosed to the bile duct. III. Hepaticojejunostomy. F. The entire extrahepatic biliary tree has been resected and the reconstruction completed with a Roux-en-Y limb of jejunum. G. Percutaneous transhepatic stents are placed across hepaticojejunostomy (optional).
further investigation. Ultrasonography is usually the diagnostic test of choice, as it can be done bedside in the intensive care unit. It can demonstrate the distended gallbladder with thickened wall, biliary sludge, pericholecystic fluid, and the presence or absence of abscess formation. CT scanning can aid in the diagnosis of acalculous cholecystitis and additionally allows a more general evaluation of the abdomen and chest to rule out other sources of infection. A HIDA scan can also be useful if it shows nonvisualization of the gallbladder, but it is less sensitive and can have higher false-positive rates in patients who are in a prolonged fasting state, on total parenteral nutrition, or have liver disease. Once the diagnosis is confirmed, acalculous cholecystitis requires urgent intervention as rapid deterioration can occur. This should include early broad-spectrum antibiotics and fluid resuscitation. If the patient is stable to undergo an abdominal operation, laparoscopic cholecystectomy is the most definitive treatment, and it can be safely performed even in the setting of severe acute inflammation. However, if patients are critically ill and unfit for surgery, percutaneous cholecystostomy is the best treatment choice (see Fig. 32-18). About 90% of patients will improve with a percutaneous cholecystostomy tube. Interval cholecystectomy can be discussed with the patient after they have recovered from their acute illness, but it is not strictly required in the absence of gallstone or other identified gallbladder pathology.

**Choledochal (Biliary) Cysts**

Choledochal cysts are congenital cystic dilatations of the extrahepatic and/or intrahepatic biliary tree. They are rare, with an incidence of between 1:100,000 and 1:150,000 in populations of Western countries, but are more common in populations of Eastern countries occurring in as many as 1:1000 individuals. Choledochal cysts affect females three to eight times more often than males. Although frequently found in infancy or childhood, nearly one-half are diagnosed in adults. The cause is unknown, but it is believed that weakness of the bile duct wall and increased pressure secondary to partial biliary obstruction can contribute to biliary cyst formation. More than 90% of patients have an anomalous pancreaticobiliary duct junction, with the pancreatic duct joining the common bile duct outside the duodenal wall, creating a long common channel (>1.5 cm). This may allow free reflux of pancreatic secretions into the biliary tract, leading to inflammatory changes, increased biliary pressure, and cyst formation. The cysts are lined with cuboidal epithelium and can vary in size from small dilations to giant cystic masses. The typical clinical triad of biliary cysts includes abdominal pain, jaundice, and a palpable mass, though this constellation is seen in less than one-half of patients. Adults may present with cholangitis. Blood tests will often be normal though elevations of transaminases can be seen in cases of infection or obstruction. Ultrasonography or CT scanning will confirm the diagnosis, but ERCP or MRCP are essential to formally assess the biliary anatomy and to plan the appropriate surgical treatment. The risk of cholangiocarcinoma in patients with choledochal cysts is 20- to 30-fold higher than in the general population and varies with the patient’s age and type of cyst. For this reason, excision is recommended whenever possible when high-risk choledochal cysts are diagnosed.

Choledochal cysts are classified into five types depending on the location and structure of the cysts. The subcategories of choledochal cysts are defined in Fig. 32-27. Type I cysts (fusiform CBD dilations) are the most common form, accounting for approximately 50% of cases, and have the highest risk of malignancy (>60%). For types I and II (saccular diverticula of the common bile duct), excision of the cystic dilations in the extrahepatic biliary tree, including cholecystectomy, with either simple cyst excision or duct resection with Roux-en-Y hepaticojejunostomy is ideal. Type III cysts (intraduodenal) create a treatment challenge as full resection would require pancreaticoduodenectomy. Given that type III cyst are associated with the lowest malignancy risk of any choledochal cyst (~2%), sphincterotomy and surveillance is generally recommended over formal excision. In Type IV (multiple cysts), excision of all cystic tissue and reconstruction is again recommended. For type IVa, which is characterized by multiple cysts with intrahepatic involvement, additional segmental resection of the liver may be required if intrahepatic stones, strictures, or abscesses are present. Type V choledochal cysts (Caroli disease) are very rare and account for less than 1% of patients with choledochal cysts. These cysts are multiple and can affect the entire liver. In advanced stages, this may result in cirrhosis and liver failure necessitating liver transplantation.

**Primary Sclerosing Cholangitis**

Primary sclerosing cholangitis (PSC) is an uncommon disease characterized by inflammatory strictures involving the intrahepatic and extrahepatic biliary tree. It is a progressive disease that eventually results in secondary biliary cirrhosis. Sometimes, biliary strictures are clearly secondary to bile duct stones, acute cholangitis, previous biliary surgery, or toxic agents, and are termed secondary sclerosing cholangitis. However, primary sclerosing cholangitis is a disease entity of its own, with no clear attributing cause. Autoimmune reaction, chronic low-grade bacterial or viral infection, toxic reaction, and genetic factors have all been suggested to play a role in its pathogenesis. PSC is commonly associated with other autoimmune diseases including ulcerative colitis in about two-thirds of patients, Riedel’s thyroiditis, and retroperitoneal fibrosis. The human leukocyte antigen haplotypes HLA-B8, DR3, DQ2, and DRw52A, commonly found in patients with autoimmune diseases, also are more frequently seen in patients with primary sclerosing cholangitis than in controls. The mean age of presentation for PSC is 30 to 45 years, and men are affected twice as often as women. Most patients are symptomatic when diagnosed, and may complain of intermittent jaundice, fatigue, weight loss, pruritus, or abdominal pain. Initial presentation with acute cholangitis is rare without preceding biliary tract intervention or surgery. A minority of patients are diagnosed incidentally by elevated liver function tests, particular when found in a patient with ulcerative colitis. While the clinical presentation and laboratory results may suggest the PSC, ERCP revealing multiple dilatations and strictures (beading) of the intra- and extrahepatic biliary tree confirms the diagnosis. The hepatic duct bifurcation is often the most severely affected segment. A liver biopsy may not be diagnostic, but it is important to determine the degree of hepatic fibrosis and the presence of cirrhosis.

The clinical course in sclerosing cholangitis is highly variable, but cyclic remissions and exacerbations are typical. Some patients will remain asymptomatic for years, while others progress rapidly with the obliterative inflammatory changes leading to secondary biliary cirrhosis and liver failure. In patients with associated ulcerative colitis, the course of each disease seems independent of the other and colectomy has no effect on the
course of primary sclerosing cholangitis. Of the patients with sclerosing cholangitis, 10% to 15% will develop cholangiocarcinoma, which can present at any time during the disease process and does not necessarily correlate with the extent of the sclerosing cholangitis or the development of liver failure. Cholangiocarcinoma in the setting of PSC frequently follows an aggressive course. Patients need to be followed by serial ERCP and liver biopsies to evaluate for the development of complications such as strictures, cancers, or cirrhosis.

There is no known curative treatment for primary sclerosing cholangitis and medical management is largely supportive. Corticosteroids, immunosuppressants, ursodeoxycholic acid, and antibiotics have been attempted with disappointing results. If biliary strictures occur, they can be dilated and stented either endoscopically or percutaneously. These measures have given short-term improvements in symptoms and serum bilirubin levels but provide long-term results in less than half of patients. Surgical management with resection of the extrahepatic biliary tree and hepaticojejunostomy has produced reasonable results in patients with extrahepatic and bifurcation strictures, but without cirrhosis or significant hepatic fibrosis. In patients with primary sclerosing cholangitis and advanced liver disease, liver transplantation is the only option. It offers excellent results, with overall 5-year survival as high as 85%. Unfortunately, recurrence of PSC can occur in 10% to 20% of patients and may require retransplantation.

**Bile Duct Strictures**

Benign bile duct strictures can have numerous causes. However, the vast majority are related to operative injury, most commonly during cholecystectomy. Other causes include fibrosis due to chronic pancreatitis, common bile duct stones, acute cholangitis, biliary obstruction due to cholecystolithiasis (Mirizzi’s syndrome), sclerosing cholangitis, cholangiohepatitis, and strictures of a biliary-enteric anastomosis. Bile duct strictures that go unrecognized or are improperly managed can lead to severe complications such as recurrent cholangitis, secondary biliary cirrhosis, and portal hypertension.

Bile duct strictures most commonly result in recurrent episodes of cholangitis but may present with isolated jaundice without infection. Liver function tests usually show evidence of cholestasis with elevations of bilirubin and alkaline phosphatase. Imaging with ultrasound or CT can show dilated bile ducts proximal to the stricture, as well as provide information about the level of the stenosis. MRCP gives more detailed anatomic information about the location and the degree of dilatation. If the diagnosis remains in question, cholangiography (endoscopic or more rarely percutaneous) will outline the biliary tree, define the stricture and its location, and allow for therapeutic interventions (Fig. 32-28). The treatment of biliary strictures depends on the location and the cause of the stricture. Percutaneous or endoscopic dilatation and/or stent placement will provide good results in more than one half of
patients. For persistent or complex strictures, surgical resection and reconstruction with Roux-en-Y choledochojejunostomy or hepaticojejunostomy may be necessary and will result in good or excellent outcomes in 80% to 90% of patients.70 Choledochoduodenostomy may be a choice for strictures in the distal-most part of the common bile duct if a tension-free repair can be achieved.

**INJURY TO THE BILIARY TRACT**

**Gallbladder**

Injuries to the gallbladder itself are uncommon but can occur in the setting of penetrating trauma (gunshot or stab wounds) or medical procedures (liver biopsy or surgery). Nonpenetrating trauma to the gallbladder is extremely rare but can cause contusion, avulsion, laceration, rupture, or traumatic cholecystitis. Regardless of the etiology of gallbladder injury, the treatment of choice is cholecystectomy. The prognosis is typically good but depends on the extent of related injury, as damage to nearby organs is not uncommon.

**Extrahepatic Bile Ducts**

Rarely, penetrating trauma to the extrahepatic bile ducts does occur, and it is usually associated with trauma to other viscera. The vast majority of injuries to the extrahepatic biliary system, however, are iatrogenic, usually occurring during cholecystectomy. These injuries are among the most feared and litigated complications in surgery, and can result in significant morbidity.71,72 Biliary tract injury can also occur during common bile duct exploration, division or mobilization of the duodenum during gastrectomy, or dissection of the hepatic hilum during liver resections.

The incidence of bile duct injury during cholecystectomy is estimated to be relatively low (about 0.2%).73 While initial experience with laparoscopic cholecystectomy appeared to show a higher rate of injury to the bile ducts compared to the open approach, these trends appear to be disappearing as laparoscopic technology and familiarity with the techniques of the procedure have improved.53 A number of different factors are thought to be associated with bile duct injury during laparoscopic cholecystectomy. These include acute or chronic inflammation, obesity, anatomic variations, and surgical technique. Inadequate exposure or failure to correctly identify structures before ligating or dividing them are the most common causes of significant biliary injury (see “Anatomic Variants” section). Excessive cephalad retraction of the gallbladder may align the cystic duct with the common bile duct, and the latter may then be mistakenly clipped and divided. Careless use of electrocautery can lead to thermal injury. Dissection deep into the liver parenchyma may cause injury to intrahepatic ducts, and poor clip placement close to the hilar area or to structures not well visualized can result in a clip across a bile duct.74,75

Techniques to avoid injury to the bile ducts during cholecystectomy are important to understand. The use of an angled, 30° or 45° laparoscope instead of an end-viewing camera will help visualize the anatomic structures, in particular those around the triangle of Calot. An angled scope also will aid in the proper placement of clips. The routine use of intraoperative cholangiography during every cholecystectomy as a method to prevent bile duct injury remains controversial.55 Nonetheless, the frequency of bile duct injuries is cut by 50% when an intraoperative cholangiogram is performed. Critical to the successful use of cholangiography is accurate interpretation of the imaging. It is important to check that the whole biliary system fills with contrast, including both major ducts on the right and the left hepatic duct, and that there is no extravasation of contrast. While routine use may reduce or limit the extent of injury, or help identify it early, it does not seem to prevent it entirely.76 No consensus recommendation exists on the use of selective versus routine cholangiography.

Perhaps the most universally agreed upon method for mitigating the risk of bile duct injury during laparoscopic cholecystectomy is obtaining the critical view of safety. This requires that the hepatocystic triangle is dissected free of fat and fibrous tissue, the lower third of the gallbladder is separated from the cystic plate, and there are two and only two structures running into the gallbladder, the cystic duct, and the cystic artery (see Fig. 32-23).54 Newer technologies such as fluorescence cholangiography to help identify biliary anatomy intraoperatively have shown promising early results, though large-scale applications remain to be seen.77

**Diagnosis.** Only about 25% of major bile duct injuries (common bile duct or hepatic duct) are recognized at the time of surgery. In these cases, intraoperative bile leakage, recognition of the correct anatomy, or an abnormal cholangiogram led to the diagnosis of a bile duct injury. In those that go unrecognized at the time of surgery, more than half will re-present within the first month postoperatively, though some can present months or years later with strictures, cholangitis, or cirrhosis from a remote bile duct injury.
Bile duct injuries typically result in either leaks or obstructions related to strictures. Bile leak, most commonly from the cystic duct stump, a transected aberrant right hepatic duct, or a lateral injury to the main bile duct, usually presents with abdominal pain, fever, and a mild elevation of liver function tests. If a drain was placed at the time of surgery, bilious fluid may be seen. A CT scan or ultrasound can show either a fluid collection in the gallbladder fossa (biloma), or free fluid ( bile) in the peritoneum (Fig. 32-29A). ERCP (Fig. 32-29B) or HIDA scan can be utilized to better localize the site of the bile leak.

Obstruction or stricture should be suspected in patients with progressive elevations of liver function tests or jaundice after cholecystectomy. CT scan or ultrasound can demonstrate the dilated part of the biliary tree, and may identify the level of the bile duct obstruction. MRI cholangiography, if available, provides an excellent, noninvasive delineation of the biliary anatomy both proximal and distal to the injury. Endoscopic or percutaneous cholangiography may also be helpful to confirm the diagnosis, depending on the location and type of injury.

Management. The management of bile duct injuries depends on the type, extent, and level of the injury, as well as the timing of its diagnosis. Initial proper treatment of bile duct injury can avoid the development of further complications or bile duct strictures. If an injury is discovered that exceeds the capacity of the available surgical expertise, the patient should be transferred to a tertiary care center. In these situations, drains should be placed in the surgical bed and antibiotics initiated. If a complete obstructive transection has occurred, it may also be necessary to place a percutaneous transhepatic drainage catheter to decompress the biliary tree prior to transfer.

If identified at the time of surgery, bile leaks from small bile ducts (<3 mm) or those draining a single hepatic segment can safely be ligated. If the injured duct is ≥4 mm, however, it is likely to drain multiple segments or an entire lobe and thus needs to be repaired or reimplanted. Minor injuries to the common bile duct or the common hepatic duct are traditionally managed with placement of a T-tube that has been modified by cutting the ends to allow for its placement in and removal from the bile duct (see Fig. 32-18). If the injury is small, the T-tube may be placed through it as if it were a formal choledochotomy.

In more extensive injuries, the T-tube should be placed through a separate choledochotomy and the injury closed over the T-tube end to minimize the risk of subsequent stricture formation.

Major bile duct injuries identified intraoperatively such as complete transection of the common hepatic or common bile duct are best managed at the time of injury. In many of these major injuries, the bile duct has not only been transected, but a variable length of the duct may have been removed with the surgical specimen. This injury usually requires reconstruction with a biliary-enteric anastomosis, and is best performed as soon as possible following the injury. If there is no or minimal loss of ductal length, a duct-to-duct repair may be done over a T-tube that is placed through a separate incision. In any repair that is chosen, it is critical to perform a tension-free anastomosis to minimize the high risk of postoperative stricture formation.

Bile leaks identified postoperatively can usually be managed with percutaneous drainage of intra-abdominal fluid collections followed by endoscopic biliary stenting. With most leaks, regardless of the location, stenting of the common bile duct will provide a low resistance route for bile flow into the duodenum, decreasing flow through the leak and allowing it to heal. This is particularly effective for leaks from the cystic duct stump. Rarely, surgery is required to repair a large leak if endoscopic interventions have failed or peritonitis is present.

Major bile duct injuries diagnosed in the later postoperative period may not be amenable to immediate reconstruction due to acute inflammation. They may need to be managed with transhepatic biliary catheter placement for biliary decompression as well as percutaneous drainage of intra-abdominal bile collections, if any. When the acute inflammation has resolved 6 to 8 weeks later, operative repair is performed.

Patients with bile duct stricture from an injury or as a sequela of previous repair usually present with either progressive elevation of liver function tests or cholangitis. The initial management usually includes endoscopic attempts at dilation or stenting. Balloon dilatation of a stricture usually requires multiple procedures and rarely provides long-term relief. Self-expanding metal or plastic stents can provide temporary or, in the high-risk patient, permanent drainage of the biliary tree. If the stricture is unable to be addressed endoscopically,
percunaneous transhepatic biliary drainage catheter placement may be necessary for decompression, and to define the anatomy, location and extent of the damage. Definitive treatment of refractory biliary strictures entails resection of the affected segment and reconstruction with a biliary-enteric anastomosis.

Outcome. Good results can be expected in the majority of patients with bile duct injuries, with the best results coming when the injury is recognized immediately and repaired by an experienced biliary tract surgeon. The perioperative mortality rate is reported to be less than 10%, but common morbidities associated with bile duct repairs include cholangitis, external biliary fistula, bile leak, subhepatic and subphrenic abscesses, and hemobilia. Restenosis of a biliary enteric anastomosis occurs in about 10% of patients, and typically presents within 2 years, but can manifest up to 20 years after the initial procedure. In general, treatment of proximal strictures is associated with a lower success rate than distal ones. The worst results are seen in patients with many operative revisions and in those who have evidence of liver failure or portal hypertension. However, previous repair does not preclude a successful outcome, particularly in patients with good liver function. In the most severe cases, patients with refractory strictures and deteriorating liver function may become candidates for liver transplant.

TUMORS

Carcinoma of the Gallbladder

Cancer of the gallbladder is a rare malignancy that occurs predominantly in the elderly. It is an aggressive tumor, with a poor prognosis that is usually not diagnosed until it has become advanced and is causing symptoms. The median survival for gallbladder cancer is around 6 months with a reported 5-year survival rate of 5%. In a minority of cases, early cancers are identified incidentally following cholecystectomy for cholelithiasis, in which case, 5-year survival is over 80%.

Incidence. Gallbladder cancer is the sixth most common GI malignancy in Western countries. It accounts for 2% to 4% of all malignant GI tumors, with about 4000 new cases diagnosed annually in the United States. It is two to six times more common in females than males, and the peak incidence is in the seventh decade of life. Its occurrence in random autopsy series is about 0.4%, but 0.3% to 3% of patients undergoing cholecystectomy for gallstone disease are found incidentally to have gallbladder cancer.

There are also significant ethnic variations in the incidence of gallbladder cancer, with rates being particularly high in native populations of the United States, Mexico, and Chile. The overall incidence of gallbladder cancer in the United States is approximately 1.5 cases per 100,000 residents. For Native American females with gallstones, the incidence is around 7.1 per 100,000. For women in the native populations of Chile, gallbladder cancer occurs in 27.3 per 100,000 individuals. Asian populations, particularly those of Korean descent, are also at increased risk of developing gallbladder cancer.

Etiology. The pathogenesis of gallbladder cancer has not been fully defined but is likely related to a combination of chronic inflammation, infection, genetics, and environmental exposures such as heavy metals and tobacco. Cholelithiasis is the most important risk factor for gallbladder carcinoma, and up to 85% of patients with carcinoma of the gallbladder have gallstones. However, <3% of patients with gallstones have gallbladder cancer, and the 20-year risk of developing cancer remains low; <0.5% for the overall population and 1.5% for high-risk groups. Larger stones (>3 cm) are associated with a 10-fold increased risk of cancer. The risk of developing cancer of the gallbladder is higher in patients with symptomatic than asymptomatic gallstones, and it is more commonly seen in the setting of cholesterol stones.

Polyloid lesions of the gallbladder, which are present in as many as 5% of adults, are also associated with increased risk of cancer. This is particularly true for polyps measuring >10 mm, which carry a 25% risk of malignancy. Solitary or sessile polyps, or those showing rapid growth on serial imaging, particularly if in the presence of gallstones or age >50 are also concerning for malignancy. When such findings are identified, the patient should have their gallbladder removed, even if they are asymptomatic. Polyps that are not removed should be monitored on serial imaging. The finding of a “Porcelain” gallbladder, or dense circumferential calcifications of the gallbladder wall, is associated with an approximately 10% risk of gallbladder carcinoma. While this condition was previously considered to be an absolute indication for cholecystectomy, more recent studies suggest that given the low rate of malignancy, observation is safe and acceptable. Nevertheless, resection remains a reasonable option, particularly if the patient is symptomatic, and the decision should ultimately be made only after discussing risks and benefits of each approach with the patient.

Patients with certain types of choledochal cysts also have an increased risk of developing cancer anywhere in the biliary tree, but the incidence is highest in the gallbladder and cholecystectomy should be performed with any surgical intervention on the choledochal cyst. Primary sclerosing cholangitis, anomalous pancreaticobiliary duct junction, and exposure to carcinogens (azotoluene, nitrosamines) also are associated with cancer of the gallbladder, and screening with abdominal ultrasound should be considered in these patients.

Pathology. Between 80% and 90% of gallbladder cancers are adenocarcinomas. Squamous cell, adenosquamous, oat cell, and other anaplastic lesions rarely occur. The histologic subtypes of gallbladder adenocarcinomas include papillary, nodular, and tubular. Less than 10% are of the papillary type, but these are associated with an overall better outcome, as they are most commonly diagnosed while localized to the gallbladder. Cancer of the gallbladder can spread through lymphatics, venous drainage, or by direct invasion into the liver parenchyma. Lymphatic flow from the gallbladder drains first to the cystic duct node (Lund’s node or Calot’s node), then pericholedochal and hilar nodes, and finally to the peripancreatic, duodenal, perportal, celiac, and superior mesenteric artery nodes. The gallbladder veins drain directly into the adjacent liver, usually segments Vb and V, where tumor invasion is common (Fig. 32-30). The gallbladder wall differs histologically from the intestines in that it lacks a muscularis mucosa and submucosa. Lymphatics are present in the subserosal layer only. Therefore, cancers that have not grown through the muscular layer have minimal risk of nodal disease. Unfortunately, only a small portion of gallbladder cancers (10–25%) are identified while they are still localized to the gallbladder. The majority will already have nodal involvement, extension into adjacent liver, or distant metastasis at the time of diagnosis.

Clinical Manifestations and Diagnosis. Signs and symptoms of carcinoma of the gallbladder are generally indistinguishable from those associated with cholecystitis and cholelithiasis, and this can lead to delays in treatment or misdiagnosis. These include abdominal discomfort, right upper quadrant pain, nausea, and vomiting. Jaundice, weight loss, anorexia, ascites,
and abdominal masses are less common presenting symptoms. Common misdiagnoses include chronic cholecystitis, acute cholecystitis, choledocholithiasis, hydrops of the gallbladder, and pancreatic cancer. Laboratory findings, if abnormal, are most often consistent with biliary obstruction. Ultrasonography often reveals a thickened, irregular gallbladder wall (>3mm) with hypervascularity or a mass replacing the gallbladder. It may also visualize tumor invasion of the liver, lymphadenopathy, or a dilated biliary tree. The sensitivity of ultrasonography in detecting gallbladder cancer ranges from 70% to 100%. A CT scan may be helpful in identifying a gallbladder mass and evaluating for nodal spread or local invasion into adjacent organs or vasculature. If questions about local invasion remain, MRCP allows for complete assessment of biliary, vascular, nodal, hepatic, and adjacent organ involvement.\(^6\) Endoscopic ultrasound (EUS) can be a useful tool in staging and evaluating for local invasion, as well as obtaining tissue diagnosis through fine needle aspiration (FNA). Tissue diagnosis can also be obtained by CT or ultrasound-guided biopsy of the tumor, though this is not required prior to cholecystectomy if the tumor appears resectable on imaging. In jaundiced patients, a percutaneous transhepatic cholangiogram may be helpful to delineate the extent of biliary tree involvement. The role of PET scanning in gallbladder cancer is yet to be fully defined but can be utilized in both staging and surveillance.

**Treatment.** Surgical resection remains the only curative option for gallbladder cancer. While most patients are unresectable at the time of diagnosis, if preoperative staging suggests a potentially resectable tumor, exploration for tissue diagnosis, formal pathologic staging, and possible curative resection are warranted.

Tumors limited to the lamina propria or muscular layer of the gallbladder (T1) are usually identified incidentally, after laparoscopic cholecystectomy for gallstone disease. There is near universal agreement that simple laparoscopic cholecystectomy is an adequate treatment for T1 lesions and results in a near 100% overall 5-year survival rate. When the tumor invades the perimuscular connective tissue without extension beyond the serosa or into the liver (T2 tumors), an extended cholecystectomy should be performed.\(^6\) This includes additional resection of liver segments IVb and V, as well as lymphadenectomy of the celiac and paraaortic lymph nodes. Given the extent of this operation, an open approach is standard. One-half of patients with T2 tumors are found to have nodal disease on pathologic examination, highlighting the importance of regional lymphadenectomy as part of surgery for T2 cancers.\(^6\) For tumors that grow beyond the serosa, or invade the liver or other adjacent organs (T3), there is a higher likelihood of intraperitoneal or distant spread. However, if no peritoneal or nodal involvement is found, complete tumor excision with an extended right hepatectomy and caudate lobectomy with lymphadenectomy must be performed for adequate tumor clearance. In addition, if a T2 or T3 tumor is identified incidentally after laparoscopic cholecystectomy, and the patient is returning to the OR for liver resection and lymphadenectomy, the previous laparoscopic port sites must also be excised due to the high risk of recurrence in these locations. T4 tumors are those that have grown into major blood vessels or two or more structures outside the liver, and they are typically considered unresectable.

Due to the high frequency of late diagnosis, palliative procedures for unresectable cancer, jaundice, or duodenal obstructions remain the most frequently performed surgery for gallbladder cancers. Today, patients with obstructive jaundice can frequently be managed with either endoscopic or percutaneously placed biliary stents. Various regimens of neoadjuvant, adjuvant, and definitive chemoradiotherapy have been trialed in gallbladder cancer. Overall, benefits have been marginal, but treatment may improve survival time by several months. These therapies can be offered to patients in conjunction with resection for curative intent or as definitive therapy, but no standard recommendation exists for their use.\(^85-87\)

**Prognosis.** Most patients with gallbladder cancer have unresectable disease at the time of diagnosis. The overall 5-year survival rate of all patients with gallbladder cancer is <5%, with a median survival of 6 months.\(^87\) However, patients with T1 disease treated with cholecystectomy have an excellent prognosis (85–100% 5-year survival rate). The 5-year survival rate for T2 lesions treated with an extended cholecystectomy (liver segment IVb/V resection) and lymphadenectomy is >70% compared to 25% to 40% for T2 patients treated with simple cholecystectomy. Patients with advanced (T3 or T4) but resectable gallbladder cancer are reported to have 5-year survival rates of 20% to 50%, supporting aggressive resection in those patients who can tolerate surgery. The median survival for patients with distant metastasis at the time of presentation is only 1 to 3 months.

Recurrence after resection of gallbladder cancer occurs most commonly in the liver or in the celiac or retropancreatic nodes. The prognosis for recurrent disease is very poor, and the main goal of follow-up is to provide palliative care. The most common problems are pruritus and cholangitis associated with obstructive jaundice, bowel obstruction secondary to carcinoma, and pain. Death occurs most commonly secondary to biliary sepsis or liver failure.

---

**Figure 32-30.** Computed tomography scan of a patient with gallbladder cancer. The image shown is at the level of the liver hilum. The portal vein is bifurcating into the left and right portal branch. The tumor has invaded segment IV of the liver (arrowheads) and obstructed the common hepatic duct, resulting in intrahepatic ductal dilatation (arrows).
Cholangiocarcinoma
Cholangiocarcinoma is a rare tumor arising from the biliary epithelium and may occur anywhere along the biliary tree. About half are located at the hepatic duct bifurcation (Klatskin tumors), with 40% occurring more distally and 10% being intrahepatic. Surgical resection offers the only chance for cure, but unfortunately many patients have advanced disease at the time of diagnosis. Therefore, palliative procedures aimed to provide biliary drainage and prevent liver failure and cholangitis are often the only therapeutic possibilities available.

Incidence. The autopsy incidence of bile duct carcinoma is about 0.3%. The overall incidence of cholangiocarcinoma in the United States is about 1 per 100,000 people per year, with approximately 2500 new cases diagnosed annually. The disease has a slight male predominance and an average age of presentation between 50 and 70 years.

Etiology. Most cases of extrahepatic cholangiocarcinoma develop de novo with no identifiable risk factors. However, there is an increased risk of cholangiocarcinoma in patients with choledochal cysts, ulcerative colitis, hepatolithiasis, biliary-enteric anastomoses, hepatitis B and C, cirrhosis, biliary tract infections with Clonorchis (liver flukes), and chronic typhoid carriers. Exposure to dietary nitrosamines, Thorotrast, or dioxin also puts patients at increased risk for cholangiocarcinoma. Patients with primary sclerosing cholangitis have a 5% to 10% lifetime risk of developing cholangiocarcinoma with typical disease onset in their 40s. For this reason, these patients require regular screening. Features common to most risk factors include biliary stasis, bile duct stones, and infection.

Pathology. Over 95% of bile duct cancers are ductal adenocarcinomas with the vast majority occurring in the extrahepatic biliary tree. Morphologically, they are divided into nodular (the most common type), scirrhous, diffusely infiltrating, or papillary. Anatomically, they are divided into distal, perihilar or intrahepatic tumors. Intrahepatic cholangiocarcinomas make up approximately 10% of cases and are typically treated like hepatocellular carcinoma, with hepatectomy when possible and transplant when unresectable. About half of all cholangiocarcinomas are located in the perihilar region with the remaining 40% occurring more distally in the common bile duct.

Perihilar cholangiocarcinomas, also referred to as Klatskin tumors, are further classified based on anatomic location by the Bismuth-Corlette classification (Fig. 32-31). Type I tumors are confined to the common hepatic duct, but type II tumors involve the bifurcation without involvement of the secondary intrahepatic ducts. Type IIIa and IIIb tumors extend into the right and left secondary intrahepatic ducts, respectively. Type IV tumors involve both the right and left secondary intrahepatic ducts.

Clinical Manifestations and Diagnosis. Painless jaundice is the most common initial presentation in patients with cholangiocarcinoma. Pruritus, mild right upper quadrant pain, anorexia, fatigue, and weight loss may also be present. Cholangitis is the presenting symptom in about 10% of patients. Except for jaundice, physical examination is usually normal in patients with cholangiocarcinoma. Occasionally, asymptomatic patients are found to have cholangiocarcinoma while being evaluated for elevated liver function tests. Tumor markers, such as CA 125 and carcinoembryonic antigen (CEA), can be elevated in cholangiocarcinoma but tend to be nonspecific because they also increase in other GI and gynecologic malignancies. The tumor marker most commonly used to aid the diagnosis of cholangiocarcinoma is CA 19-9, which has a sensitivity of 79% and specificity of 98% if the serum value is >129 U/mL. However, mild elevations in CA 19-9 can also be seen in cholangitis, biliary obstruction, other GI and gynecologic neoplasms, and patients who lack the Lewis blood type antigen.

The initial workup for suspected cholangiocarcinoma includes abdominal imaging with ultrasound or CT scanning. Perihilar tumors will cause dilatation of the intrahepatic biliary tree, but a normal or collapsed gallbladder and extrahepatic bile ducts distal to the tumor. Distal bile duct cancer will lead to dilatation of the extra- and intrahepatic bile ducts as well as the gallbladder. Initial imaging is important to determine the level of obstruction and to rule out the presence of bile duct stones as the cause of the obstructive jaundice (Fig. 32-32). It is usually difficult to visualize the tumor itself on ultrasound, CT, or even MRCP, but any of these modalities can provide an outline of biliary anatomy, an estimate of the level of obstruction, evaluation of portal vein patency, and screening for nearby lymphadenopathy. Detailed evaluation of the biliary anatomy and tumor itself is best completed through cholangiography. ERCP is generally adequate, but in cases where the proximal extent of the tumor remains in question, PTC may be required to determine resectability.

Tissue diagnosis may be difficult to obtain. Current diagnostic techniques including fine-needle aspiration (percutaneous or endoscopic), and biliary brushings have been shown to have a low sensitivity in detecting malignancy, anywhere between 15% and 60%. Cholangioscopy with direct visualization and sampling of intraluminal masses may be able to improve diagnosis rates but is only available in specialized centers (see Fig. 32-10). Patients with potentially resectable disease should, therefore, be offered surgical exploration based on radiographic findings and clinical suspicion.
**Treatment.** Surgical excision is the only potentially curative treatment for cholangiocarcinoma. In the past one to two decades, improvements in surgical techniques have resulted in lower mortality and better outcomes for patients undergoing aggressive surgical excision for cholangiocarcinoma.\(^9\)

Despite improvements in ultrasonography, CT scanning, and MRI, more than one-half of patients who are explored are found to have peritoneal implants, nodal or hepatic metastasis, or locally advanced disease that precludes resection. Patients suspected of having resectable disease should first undergo diagnostic laparoscopy. Those who are found to have previously unidentified metastatic disease should undergo cholecystectomy and surgical bypass for biliary decompression.\(^3\)

For curative resection, the location and local extent of the tumor dictates the extent of the surgery required. Distal bile duct tumors are often resectable but may require pancreaticoduodenectomy (Whipple procedure). For patients with distal bile duct cancer found to be unresectable on surgical exploration, Roux-en-Y hepaticojunostomy, cholecystectomy, and gastrojejunostomy to prevent gastric outlet obstruction should be performed. Perihilar tumors involving the bifurcation or proximal common hepatic duct (Bismuth-Corlette type I or II) with no signs of vascular involvement are candidates for local tumor excision with portal lymphadenectomy. If the tumor involves the right or left hepatic duct (Bismuth-Corlette type IIIa or IIIb), right or left hepatic lobectomy, respectively, should also be performed. Frequently, resection of the adjacent caudate lobe is required because of direct extension into caudate biliary radicals or parenchyma.\(^9\) Type IV Klatskin tumors, those with more extensive involvement of both hepatic ducts and intrahepatic spread, are often considered unresectable or only treatable with liver transplantation.

The best outcomes in perihilar cholangiocarcinoma are seen in patients who undergo neoadjuvant chemoradiation followed by liver transplantation. However, there are very strict inclusion criteria for transplantation, and few patients qualify.\(^3\) Patients with primary sclerosing cholangitis who develop cholangiocarcinoma should be treated with liver transplant whenever possible.

Nonoperative biliary decompression can be performed for patients with unresectable disease on initial presentation. Endoscopic placement of expandable metal stents is often the preferred approach. For very proximal or intrahepatic tumors, percutaneous drainage catheters may be necessary to fully decompress the biliary tree (see Fig. 32-12). There is a significantly higher risk of cholangitis in patients with drainage catheters or stents compared to those with surgical bypasses. In addition, stent occlusion is not uncommon. Nevertheless, operative intervention is not warranted in patients with metastatic disease.\(^9\)

There is no proven role for adjuvant chemotherapy in the treatment of cholangiocarcinoma. Adjuvant radiation therapy has also not been shown to increase either quality of life or survival in resected patients. Patients with unresectable disease can be offered palliative chemotherapy, typically with gemcitabine and cisplatin, but the response rates are low (10–20%), and the survival benefit is marginal. The combination of radiation and chemotherapy may be more effective than either treatment alone for unresectable disease, but no data from randomized trials are available. Giving chemoradiation to these patients can be difficult because of the high incidence of cholangitis. External-beam radiation has not been shown to be an effective treatment for unresected disease. The use of interstitial (intraoperative) radiation, brachytherapy with iridium-192 via percutaneous or endoscopic stents, and combined interstitial and external-beam radiation for unresectable cholangiocarcinoma has been reported with some encouraging results. However, no randomized, prospective trials have been reported.\(^9\) Photodynamic therapy has been proposed as a palliative measure for patients with unresectable disease and

**Figure 32-32.** A. An endoscopic retrograde cholangiogram in a patient with cancer of the common hepatic duct (arrowheads). The common bile duct is of normal size, as is the cystic duct (arrow), but the proximal biliary tree is dilated. The gallbladder is not visualized because of tumor obstructing its neck. B. An ultrasound from the same patient showing dilated ducts and tumor obstructing the common hepatic duct (arrow). The walls of the bile ducts adjacent to the obstruction are thickened by tumor infiltration (arrowheads).
has been found to prolong survival and improve quality of life in patients with biliary stents.\textsuperscript{95,96}

**Prognosis.** Most patients with perihilar cholangiocarcinoma present with advanced, unresectable disease. Median survival in this population is between 5 and 8 months. The most common causes of death are hepatic failure and cholangitis. The overall 5-year survival rate for patients with resectable perihilar cholangiocarcinoma is between 10% and 30%, but for patients with negative margins, it may be as high as 40%. The operative mortality for perihilar cholangiocarcinoma is 6% to 8%. Patients with distal cholangiocarcinoma are more likely to have resectable disease and improved prognosis compared to perihilar cholangiocarcinoma. The overall 5-year survival rate for resectable distal disease is 30% to 50%, and the median survival is 32 to 38 months. Patients who receive liver transplantation for cholangiocarcinoma can experience 5-year disease free survival rates as high as 68%.

The greatest risk factors for recurrence after resection are the presence of positive margins and lymph node–positive tumors. Therapy for recurrent disease concentrates on palliation of symptoms and additional surgery is not recommended for patients with recurrent disease.

**REFERENCES**

Entries highlighted in bright blue are key references.

43. Lipsett PA, Pitt HA. Acute cholangitis. Front Biosci. 2003;8:S1229-S1239.


This page intentionally left blank
ANATOMY

The pancreas is perhaps the most unforgiving organ in the human body, leading most surgeons to avoid even palpating it unless necessary. Situated deep in the center of the abdomen, the pancreas is surrounded by numerous important structures and major blood vessels. Seemingly minor trauma to the pancreas can result in the release of pancreatic enzymes and cause life-threatening pancreatitis. Therefore, knowledge of the relationships of the pancreas to surrounding structures is critically important for all surgeons to ensure that pancreatic injury is avoided during abdominal surgery.

Gross Anatomy

The pancreas is a retroperitoneal organ that lies in an oblique position, sloping upward from the C-loop of the duodenum to the splenic hilum (Fig. 33-1). In an adult, the pancreas weighs 75 to 100 g and is about 15 to 20 cm long. The fact that the pancreas is situated so deeply in the abdomen and is sealed in the retroperitoneum explains the poorly localized and sometimes ill-defined nature with which pancreatic pathology presents. Patients with pancreatic cancer without bile duct obstruction usually present after months of vague upper abdominal discomfort, or no antecedent symptoms at all. Due to its retroperitoneal location, pain associated with pancreatitis often is characterized as penetrating through to the back.

Regions of the Pancreas

Surgeons typically describe the location of pathology within the pancreas in relation to four regions: the head, neck, body, and tail. The head of the pancreas is nestled in the C-loop of the duodenum and is posterior to the transverse mesocolon. Just posterior to the head of the pancreas lie the vena cava, the right renal artery, and both renal veins. The neck of the pancreas lies directly anterior to the portal vein. At the inferior border of the neck of the pancreas, the superior mesenteric vein joins the splenic vein and then continues toward the porta hepatis as the portal vein. The inferior mesenteric vein often joins the splenic vein near
Key Points

1. Incomplete fusion of the dorsal and ventral pancreatic ducts results in pancreas divisum, but a variety of ductal anomalies can be seen. Magnetic resonance cholangiopancreatography as well as endoscopic retrograde cholangiopancreatography can identify these ductal anomalies, and clarification of the ductal pattern of the pancreas is important before attempts at interventions.

2. The “replaced right hepatic artery” occurs in 15% of patients and needs to be identified preoperatively to prevent inadvertent injury with resulting hepatic necrosis. Anomalous hepatic arterial anatomy can result in hepatic ischemia during dissection of the porta hepatis as well. “Thin cut” multidetector computed tomographic images are usually able to identify the relevant arterial and venous patterns around the pancreas.

3. Regardless of the etiology, the management of the early phase of acute pancreatitis is critical to achieve a successful outcome. Aggressive fluid resuscitation and early enteral feeding both reduce the risk of complications. It is no longer considered appropriate to “rest the pancreas” if the patient can tolerate enteral nutrients.

4. Surgical intervention in acute pancreatitis is reserved for patients with infected collections or infected necrosis only, or to relieve an impacted gallstone in the ampulla if endoscopic or radiologic treatments are unsuccessful. Infection is usually confirmed by a pattern of air in the retroperitoneum on computed tomographic scan or by documentation of bacteria on Gram’s stain or culture from fine-needle aspiration of a suspected infected fluid collection. Fine-needle aspiration of suspicious fluid collections should not be converted to percutaneous drainage unless infection is confirmed and the consensus decision has been made that percutaneous drainage is appropriate for the individual patient.

5. The appearance of chronic pancreatitis on computed tomographic scan varies dramatically, and multiple diagnostic studies are usually needed to establish the extent of disease. Calcific pancreatitis is not a marker of alcoholic pancreatitis alone, and it rarely indicates autoimmune pancreatitis. Endoscopic ultrasound provides a better assessment of the disease than computed tomography and is useful to disclose indolent or unsuspected cancer, which can occur in up to 10% of patients.

6. The nidus of inflammation in chronic pancreatitis due to any cause is the head of the gland. Therefore, treatment approaches that address the disease in the head have the best long-term results. The Whipple procedure, the Beger procedure, and the Frey procedure, with or without longitudinal duct drainage, are the best surgical options, as all three approaches remove all or most of the disease in the head of the gland. Although the limited pancreatic procedures (Beger/Frey) have a lower initial rate of endocrine dysfunction, the long-term risk of diabetes is more related to the progression of the underlying disease than to the effects of operation. Level 1 studies confirm that the duodenum preserving options are associated with a lower immediate morbidity and mortality and therefore, in the absence of a mass or concerns about cancer, are better options than a Whipple procedure for chronic pancreatitis.

7. The precursor lesion that probably leads to most cases of ductal adenocarcinoma is the ductal epithelial hyperplasia/dysplasia process described by the pancreatic intraepithelial neoplasia classification system. Pancreatic intraepithelial neoplasia 2 and pancreatic intraepithelial neoplasia 3 lesions may be associated with other, nonspecific changes in pancreatic morphology seen on imaging studies, or they may only be seen histologically. Resection margins for pancreatic neoplasms should be examined for advanced pancreatic intraepithelial neoplasia stage patterns of ductal hyperplasia to ensure adequate resection status.

8. A low threshold for ordering a computed tomography scan with “pancreatic protocol” should be maintained for older adult patients with unexplained, persistent, although vague, abdominal pain. New-onset diabetes in an older adult patient, especially if combined with vague abdominal pain, should prompt a search for pancreatic cancer.

9. Intraductal papillary mucinous neoplasms are small macroscopic polypoid or plaque-like adenomas that develop in the main pancreatic duct or in side-branch ducts, and secrete mucus. They are often silent symptomatically but cause characteristic appearances of small cyst-like collections of mucus or diffuse dilatation of the main pancreatic duct with mucus. These premalignant lesions may be multifocal or single and can evolve into invasive adenocarcinoma in a similar pattern as with other adenomatous polypoid lesions of the gastrointestinal tract. They have been diagnosed with increasing frequency and account for more than one-third of pancreatic resections at some centers. Main-duct intraductal papillary mucinous neoplasms are an indication for resection; side-branch intraductal papillary mucinous neoplasms have a lower incidence of malignancy and are sometimes followed with serial imaging surveillance.
Figure 33-1. Pancreatic anatomy as seen on computed tomography. Knowledge of the relationship of the pancreas with surrounding structures is important to ensure that injury is avoided during abdominal surgery. IMV = inferior mesenteric vein; SMA = superior mesenteric artery; SMV = superior mesenteric vein.

Figure 33-2. Variations in portal venous anatomy. The superior mesenteric vein joins the splenic vein and then continues toward the porta hepatis as the portal vein. The inferior mesenteric vein often joins the splenic vein near its junction with the portal vein, but sometimes joins the superior mesenteric vein; or the three veins merge as a trifurcation to form the portal vein.
Pancreatic Duct Anatomy

An understanding of embryology is required to appreciate the common variations in pancreatic duct anatomy. The pancreas is formed by the fusion of a ventral and dorsal bud (Fig. 33-3). The duct from the smaller ventral bud, which arises from the hepatic diverticulum, connects directly to the common bile duct. The duct from the larger dorsal bud, which arises from the duodenum, drains directly into the duodenum. The duct of the ventral anlage becomes the duct of Wirsung, and the duct from the dorsal anlage becomes the duct of Santorini. With gut rotation, the ventral anlage rotates to the right and around the posterior side of the duodenum to fuse with the dorsal bud. The ventral anlage becomes the inferior portion of the pancreatic head and the uncinate process, while the dorsal anlage becomes the body and tail of the pancreas. The ducts from each anlage usually fuse together in the pancreatic head such that most of the pancreas drains through the duct of Wirsung, or main pancreatic duct, into the common channel formed from the bile duct and pancreatic duct. The length of the common channel is variable. In about one-third of patients, the bile duct and pancreatic duct remain distinct to the end of the papilla, the two ducts merge at the end of the papilla in another one-third, and in the remaining one-third, a true common channel is present for a distance of several millimeters. Commonly, the duct from the dorsal anlage, the duct of Santorini, persists as the lesser pancreatic duct, and sometimes drains directly into the duodenum through the lesser papilla just proximal to the major papilla. In approximately 30% of patients, the duct of Santorini ends as a blind accessory duct and does not empty into the duodenum. In 10% of patients, the ducts of Wirsung and Santorini fail to fuse. This results in the majority of the pancreas draining through the duct of Santorini and the lesser papilla, while the inferior portion of the pancreatic head and uncinate process drains through the duct of Wirsung and major papilla. This normal anatomic variant, which occurs in one out of 10 patients, is referred to as pancreas divisum (Fig. 33-3). In a minority of these patients, the minor papilla can be inadequate to handle the flow of pancreatic juices from the majority of the gland. This relative outflow obstruction can result in pancreatitis and is sometimes treated by sphincteroplasty of the minor papilla.

The main pancreatic duct is usually only 2 to 3 mm in diameter and runs midway between the superior and inferior borders of the pancreas, usually closer to the posterior than to the anterior surface. Pressure inside the pancreatic duct is about twice that in the common bile duct, which is thought to prevent reflux of bile into the pancreatic duct. The main pancreatic duct joins with the common bile duct and empties at the ampulla of Vater or major papilla, which is located on the medial aspect of the second portion of the duodenum. The muscle fibers around the ampulla form the sphincter of Oddi, which controls the flow of pancreatic and biliary secretions into the duodenum. Contraction and relaxation of the sphincter is regulated by complex neural and hormonal factors. When the accessory pancreatic duct or lesser duct drains into the duodenum, a lesser papilla can be identified approximately 2 cm proximal to the ampulla of Vater.

Figure 33-3. Embryology of pancreas and duct variations. The duct of Wirsung from the ventral bud connects to the bile duct, while the duct of Santorini from the larger dorsal bud connects to the duodenum. With gut rotation, the two ducts fuse in most cases such that the majority of the pancreas drains through the duct of Wirsung to the major papilla. The duct of Santorini can persist as a blind accessory duct or drain through the lesser papilla. In a minority of patients, the ducts remain separate, and the majority of the pancreas drains through the duct of Santorini, a condition referred to as pancreas divisum.
Vascular and Lymphatic Anatomy

The blood supply to the pancreas comes from multiple branches from the celiac and superior mesenteric arteries (Fig. 33-4). The common hepatic artery gives rise to the gastroduodenal artery before continuing toward the porta hepatis as the proper hepatic artery. The right gastric artery branches off the gastroduodenal artery just superior to the duodenum. The gastroduodenal artery also supplies the superior pancreaticoduodenal artery which divides into the anterior and superior pancreaticoduodenal arteries. These travel inferiorly within the pancreaticoduodenal groove giving off small branches to the duodenum and head of the pancreas. The superior pancreaticoduodenal arteries join the inferior pancreaticoduodenal arteries to complete the arcade. The inferior pancreaticoduodenal artery is a branch off the superior mesenteric artery. Therefore, it is impossible to resect the head of the pancreas without devascularizing the duodenum unless a rim of pancreas containing the pancreaticoduodenal arcade is preserved. The inferior pancreaticoduodenal artery needs to be controlled when dissecting the head of the pancreas off the SMA during a Whipple procedure. The gastroduodenal artery travels inferiorly anterior to the neck of the pancreas and posterior to the duodenal bulb. A posterior ulcer in the duodenal bulb can erode into the gastroduodenal artery in this location. At the inferior border of the duodenum, the gastroduodenal artery then gives rise to the right gastroepiploic artery then can continue on to join the inferior pancreaticoduodenal artery.

Variations in the arterial anatomy occur in one out of five patients. The right hepatic artery, common hepatic artery, or gastroduodenal arteries can arise from the superior mesenteric artery. In 15% to 20% of patients, the right hepatic artery will arise from the superior mesenteric artery and travel upwards toward the liver along the posterior aspect of the head of the pancreas (referred to as a replaced right hepatic artery). It is important to look for this variation on preoperative computed tomographic (CT) scans and in the operating room so the replaced hepatic artery is recognized and injury is avoided. The body and tail of the pancreas are supplied by multiple branches of the splenic artery. The splenic artery arises from the celiac trunk and travels along the posterior-superior border of the body and tail of the pancreas toward the spleen. The inferior pancreatic artery usually arises from the superior mesenteric artery and runs to the left along the inferior border of the body and tail of the pancreas, parallel to the splenic artery. Three vessels run perpendicular to the long axis of the pancreatic body and tail and connect the splenic artery and inferior pancreatic artery. They are, from medial to lateral, the dorsal, great, and caudal pancreatic arteries. These arteries form arcades within the body and tail of the pancreas and account for the rich blood supply of the organ.

The venous drainage of the pancreas follows a pattern similar to that of the arterial supply (Fig. 33-5). The veins are usually superficial to the arteries within the parenchyma of the pancreas. There is an anterior and posterior venous arcade within the head of the pancreas. Typically, the superior vein drains directly into the portal vein just above the neck of the pancreas and is often a larger branch of the portal vein which is divided during the Whipple procedure. The posterior inferior arcade drains directly into the inferior mesenteric vein at the inferior border of the neck of the pancreas and this is also divided during a Whipple procedure. The anterior inferior pancreaticoduodenal vein joins the right gastroepiploic vein and the middle colic vein to form a common venous trunk, which enters into the superior mesenteric vein. Traction on the transverse colon during colectomy can tear these fragile veins, which then retract into the parenchyma of the pancreas, making control tedious. There also are numerous small venous branches coming from the pancreatic parenchyma directly into the lateral and posterior aspect of the portal vein. Venous return from the body and tail of the pancreas drains into the splenic vein.

The lymphatic drainage from the pancreas is diffuse and widespread (Fig. 33-6). The profuse network of lymphatic
Figure 33-5. Venous drainage from the pancreas. The venous drainage of the pancreas follows a pattern similar to the arterial supply, with the veins usually superficial to the arteries. Anterior traction on the transverse colon can tear fragile branches along the inferior border of the pancreas, which then retract into the parenchyma of the pancreas. Venous branches draining the pancreatic head and uncinate process enter along the right lateral and posterior sides of the portal vein. There are usually no anterior venous tributaries, and a plane can usually be developed between the neck of the pancreas and the portal and superior mesenteric veins.

Figure 33-6. Lymphatic supply to the pancreas. The lymphatic drainage from the pancreas is diffuse and widespread, which explains the high incidence of lymph node metastases and local recurrence of pancreatic cancer. The pancreatic lymphatics also communicate with lymph nodes in the transverse mesocolon and mesentery of the proximal jejunum. Tumors in the body and tail of the pancreas are often unresectable because they metastasize to these lymph nodes. (Reproduced with permission from Bell RH, Rikkers LF, Mulholland M: Digestive Tract Surgery: A Text and Atlas. Philadelphia, PA: Lippincott Williams & Wilkins; 1996.)
vessels and lymph nodes draining the pancreas provides egress to tumor cells arising from the pancreas. This diffuse lymphatic drainage contributes to the fact that pancreatic cancer often presents with positive lymph nodes and a high incidence of local recurrence after resection. Lymph nodes can be palpated along the distal bile duct and posterior aspect of the head of the pancreas in the pancreaticoduodenal groove, where the mesenteric vein passes under the neck of the pancreas, along the inferior border of the body, at the celiac axis and along the hepatic artery ascending into the porta hepatitis, and along the splenic artery and vein. The pancreatic lymphatics also communicate with lymph nodes in the transverse mesocolon and mesentery of the proximal jejunum. Tumors in the body and tail of the pancreas often metastasize to these nodes and lymph nodes along the splenic vein and in the hilum of the spleen.

**Neuroanatomy**

The pancreas is innervated by the sympathetic and parasympathetic nervous systems. The acinar cells responsible for exocrine secretion, the islet cells responsible for endocrine secretion, and the islet vasculature are innervated by both systems. The parasympathetic system stimulates endocrine and exocrine secretion and the sympathetic system inhibits secretion.² The pancreas is also innervated by neurons that secrete amines and peptides, such as somatostatin, vasoactive intestinal peptide (VIP), calcitonin gene-related peptide (CGRP), and galanin. The exact role of these neurons in pancreatic physiology is uncertain, but they do appear to affect both exocrine and endocrine function. The pancreas also has a rich supply of afferent sensory fibers, which are responsible for the intense pain associated with advanced pancreatic cancer, as well as acute and chronic pancreatitis. These somatic fibers travel superiorly to the celiac ganglia (Fig. 33-7). Interruption of these somatic fibers with a celiac plexus block can stop transmission of pain sensation.

**HISTOLOGY AND PHYSIOLOGY**

The exocrine pancreas accounts for about 85% of the pancreatic mass; 10% of the gland is accounted for by extracellular matrix, and 4% by blood vessels and the major ducts, whereas only 2% of the gland is comprised of endocrine tissue. The endocrine and exocrine pancreas are sometimes thought of as functionally separate, but these different components of the organ are coordinated to allow an elegant regulatory feedback system for digestive enzyme and hormone secretion. This complex system regulates the type of digestion, its rate, and the processing and distribution of absorbed nutrients. This coordination is facilitated by the physical approximation of the islets and the exocrine pancreas, the presence of specific islet hormone receptors on the plasma membranes of pancreatic acinar cells, and the existence of an islet-acinar portal blood system.

Although patients can live without a pancreas when insulin and digestive enzyme replacement are administered, the loss of this islet-acinar coordination leads to impairments in digestive function. Although only approximately 20% of the normal pancreas is required to prevent insufficiency, in many patients undergoing pancreatic resection, the remaining pancreas is not normal, and pancreatic endocrine and exocrine insufficiency can develop with removal of smaller portions of the gland.

**Exocrine Pancreas**

The pancreas secretes approximately 500 to 800 mL per day of colorless, odorless, alkaline, isosmotic pancreatic juice. Pancreatic juice is a combination of acinar cell and duct cell secretions. The acinar cells secrete amylase, proteases, and lipases, enzymes responsible for the digestion of all three food types: carbohydrate, protein, and fat. The acinar cells are pyramid-shaped, with their apices facing the lumen of the acinus. Near the apex of each cell are numerous enzyme-containing zymogen...
granules that fuse with the apical cell membrane (Fig. 33-8). Unlike the endocrine pancreas, where islet cells specialize in the secretion of one hormone type, individual acinar cells secrete all types of enzymes. However, the ratio of the different enzymes released is adjusted to the composition of digested food through nonparallel regulation of secretion.

Pancreatic amylase is secreted in its active form and completes the digestive process already begun by salivary amylase. Amylase is the only pancreatic enzyme secreted in its active form, and it hydrolyzes starch and glycogen to glucose, maltose, maltotriose, and dextrins. These simple sugars are transported across the brush border of the intestinal epithelial cells by active transport mechanisms. Gastric hydrolysis of protein yields peptides that enter the intestine and stimulate intestinal endocrine cells to release cholecystokinin (CCK)-releasing peptide, CCK, and secretin, which then stimulate the pancreas to secrete enzymes and bicarbonate into the intestine.

The proteolytic enzymes are secreted as proenzymes that require activation. Trypsinogen is converted to its active form, trypsin, by another enzyme, enterokinase, which is produced by the duodenal mucosal cells. Trypsin, in turn, activates the other proteolytic enzymes. Trypsinogen activation within the pancreas is prevented by the presence of inhibitors that are also secreted by the acinar cells. A failure to express a normal trypsinogen inhibitor, pancreatic secretory trypsin inhibitor (PSTI), also known as serine protease inhibitor Kazal type 1 (SPINK1), is a cause of familial pancreatitis. Inhibition of trypsinogen activation ensures that the enzymes within the pancreas remain in an inactive precursor state and are activated only within the duodenum. Trypsinogen is expressed in several isoforms, and a missense mutation on the cationic trypsinogen, or PRSS1, results in premature, intrapancreatic activation of trypsinogen. This accounts for about two-thirds of cases of hereditary pancreatitis. Chymotrypsinogen is activated to form chymotrypsin. Elastase, carboxypeptidase A and B, and phospholipase are also activated by trypsin. Trypsin, chymotrypsin, and elastase cleave bonds between amino acids within a target peptide chain, and carboxypeptidase A and B cleave amino acids at the end of peptide chains. Individual amino acids and small dipeptides are then actively transported into the intestinal epithelial cells. Pancreatic lipase hydrolyzes triglycerides to 2-monoglyceride and fatty acid. Pancreatic lipase is secreted in an active form. Colipase is also secreted by the pancreas and binds to lipase, changing its molecular configuration and increasing its activity. Phospholipase A2 is secreted by the pancreas as a proenzyme that becomes activated by trypsin. Phospholipase A2 hydrolyzes phospholipids and, as with all lipases, requires bile salts for its action. Carboxylic ester hydrolase and cholesterol esterase hydrolyze neutral lipid substrates like esters of cholesterol, fat-soluble vitamins, and triglycerides. The hydrolyzed fat is then packaged into micelles for transport into the intestinal epithelial cells, where the fatty acids are reassembled and packaged inside chylomicrons for transport through the lymphatic system into the bloodstream (Table 33-1).

The centroacinar and intercalated duct cells secrete the water and electrolytes present in the pancreatic juice. About 40 acinar cells are arranged into a spherical unit called an acinus. Centroacinar cells are located near the center of the acinus and are responsible for fluid and electrolyte secretion. These cells contain the enzyme carbonic anhydrase, which is needed for bicarbonate secretion. The amount of bicarbonate secreted varies with the pancreatic secretory rate, with greater concentrations of bicarbonate being secreted as the pancreatic secretory rate increases. Chloride secretion varies inversely with bicarbonate secretion such that the sum of these two remains constant. In contrast, sodium and potassium concentrations are kept constant throughout the spectrum of secretory rates (Fig. 33-9). The hormone secretin is released from cells in the duodenal mucosa in response to acidic chyme passing through the pylorus into the duodenum. Secretin is the major stimulant for bicarbonate secretion, which buffers the acidic fluid entering the duodenum from the stomach. CCK also stimulates bicarbonate secretion, but to a much lesser extent than secretin. CCK potentiates secretin-stimulated bicarbonate secretion. Gastrin and acetylcholine, both stimulants of gastric acid secretion, are also weak stimulants of pancreatic bicarbonate secretion.

Truncal vagotomy produces a myriad of complex effects on the downstream digestive tract, but the sum effect on the exocrine pancreas is a reduction in bicarbonate and fluid secretion. The endocrine pancreas also influences the adjacent exocrine pancreatic secretions. Somatostatin, pancreatic polypeptide (PP), and glucagon are all thought to inhibit exocrine secretion.

The acinar cells release pancreatic enzymes from their zymogen granules into the lumen of the acinus, and these proteins combine with the water and bicarbonate secretions of the centroacinar cells. The pancreatic juice then travels into small intercalated ducts. Several small intercalated ducts join to form an interlobular duct. Cells in the interlobular ducts contribute fluid and electrolytes to adjust the final concentrations of the pancreatic fluid. Interlobular ducts then join to form about 20 secondary ducts that empty into the main pancreatic duct.
Table 33-1
Pancreatic enzymes

<table>
<thead>
<tr>
<th>ENZYME</th>
<th>SUBSTRATE</th>
<th>PRODUCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carbohydrate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amylase (active)</td>
<td>Starch, glycogen</td>
<td>Glucose, maltose, maltotriose, dextrins</td>
</tr>
<tr>
<td>Protein</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endopeptidases</td>
<td>Cleave bonds between amino acids</td>
<td>Amino acids, dipeptides</td>
</tr>
<tr>
<td>Trypsinogen (inactive)</td>
<td>Trypsin (active)</td>
<td></td>
</tr>
<tr>
<td>Chymotrypsinogen (inactive)</td>
<td>Chymotrypsin (active)</td>
<td></td>
</tr>
<tr>
<td>Proelastase (inactive)</td>
<td>Elastase (active)</td>
<td></td>
</tr>
<tr>
<td>Exopeptidases</td>
<td>Cleave amino acids from end of peptide chains</td>
<td>—</td>
</tr>
<tr>
<td>Procarboxy peptidase A&amp;B (inactive)</td>
<td>Carboxypeptidase A&amp;B (active)</td>
<td></td>
</tr>
<tr>
<td>Fat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pancreatic lipase (active)</td>
<td>Triglycerides</td>
<td>2-Monoglycerides, fatty acids</td>
</tr>
<tr>
<td>Phospholipase A2 (inactive)</td>
<td>Phospholipase</td>
<td>—</td>
</tr>
<tr>
<td>Phospholipase A2 (active)</td>
<td>Neutral lipids</td>
<td>—</td>
</tr>
<tr>
<td>Cholesterol esterase</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Destruction of the branching ductal tree from recurrent inflammation, scarring, and deposition of stones in chronic pancreatitis eventually contributes to destruction of the exocrine pancreas and exocrine pancreatic insufficiency.

Endocrine Pancreas
There are nearly 1 million islets of Langerhans in the normal adult pancreas. They vary greatly in size from 40 to 900 μm. Larger islets are located closer to the major arterioles and smaller islets are embedded more deeply in the parenchyma of the pancreas. Most islets contain 3000 to 4000 cells of five major types: alpha cells that secrete glucagon, beta-cells that secrete insulin, delta cells that secrete somatostatin, epsilon cells that secrete ghrelin, and PP cells that secrete PP (Table 33-2).

Insulin is the best-studied pancreatic hormone. The discovery of insulin in 1920 by Frederick Banting, an orthopedic

Table 33-2
Pancreatic Islet Peptide Products

<table>
<thead>
<tr>
<th>HORMONES</th>
<th>ISLET CELL</th>
<th>FUNCTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insulin</td>
<td>Beta</td>
<td>Decreases gluconeogenesis, glycogenolysis, fatty acid breakdown, and ketogenesis. Increased glycogenesis, protein synthesis and glucose uptake.</td>
</tr>
<tr>
<td>Glucagon</td>
<td>Alpha</td>
<td>Opposite effects of insulin; increases hepatic glycogenolysis and gluconeogenesis.</td>
</tr>
<tr>
<td>Somatostatin</td>
<td>Delta</td>
<td>Inhibits secretion and action of all pancreatic and gut peptides, inhibits cell growth.</td>
</tr>
<tr>
<td>Pancreatic Polypeptide</td>
<td>PP or F</td>
<td>Inhibits pancreatic exocrine secretion and facilitates hepatic action of insulin.</td>
</tr>
<tr>
<td>Amylin (IAPP)</td>
<td>Beta</td>
<td>Counter-regulates insulin secretion and function.</td>
</tr>
<tr>
<td>Pancreastatin</td>
<td>Beta</td>
<td>Decreases insulin and somatostatin secretion, increases glucagon secretion, decreases exocrine secretion.</td>
</tr>
<tr>
<td>Ghrelin</td>
<td>Epsilon</td>
<td>Decreases insulin secretion and action.</td>
</tr>
<tr>
<td>Peptide YY (PYY)</td>
<td>not known</td>
<td>Increases insulin secretion and beta cell growth.</td>
</tr>
</tbody>
</table>

Figure 33-9. Composition of pancreatic exocrine secretions. Greater concentrations of bicarbonate are secreted at higher secretory rates, and chloride secretion varies inversely with bicarbonate secretion. In contrast, sodium and potassium concentrations are independent of the secretory rate. (Reproduced with permission from Bro-Rasmussen F, Killmann SA, Thaysen JH: The composition of pancreatic juice as compared to sweat, parotid saliva and tears, Acta Physiol Scand. 1956 Sep 26;37(2-3):97-113.)
Part II
Specific Considerations

Insulin secretion by the beta-cell is also influenced by plasma levels of amino acids such as arginine, lysine, leucine, and free fatty acids. Glucagon, GIP, GLP-1, and CCK stimulate insulin release, while somatostatin, amylin, and pancreastatin inhibit insulin release. Cholinergic fibers and alpha-sympathetic fibers stimulate insulin release, while beta-sympathetic fibers inhibit insulin secretion.

Insulin’s glucoregulatory function is to inhibit endogenous (hepatic) glucose production and to facilitate glucose transport into cells, thus lowering plasma glucose levels. Insulin also inhibits glycogenolysis, fatty acid breakdown, and ketone formation, and stimulates protein synthesis. There is a considerable amount of functional reserve in insulin secretory capacity. If the remaining portion of the pancreas is healthy, about 80% of the pancreas can be resected without the patient becoming diabetic. In patients with chronic pancreatitis, or other conditions in which much of the gland is diseased, resection of a smaller fraction of the pancreas can result in pancreateogenic, or type 3c diabetes (Table 33-3).

Insulin receptors are dimeric, tyrosine kinase–containing transmembrane proteins that are located on all cells. Insulin deficiency (seen in type 1 and type 3c diabetes) results in an overexpression or upregulation of insulin receptors, which causes an enhanced sensitivity to insulin in muscle and adipocytes (and therefore increases the risk of insulin-induced hypoglycemia). Type 2 diabetes is associated with a downregulation of insulin receptors and relative hyperinsulinemia, with resulting insulin resistance. Some forms of diabetes are associated with selected impairments of hepatic or peripheral insulin receptors, such as pancreateogenic or type 3c diabetes (T3cDM) or maturity-onset diabetes of the young (MODY).

Glucagon is a 29-amino-acid, single-chain peptide that promotes hepatic glycogenolysis and gluconeogenesis and counteracts the effects of insulin through its hyperglycemic action. Glucose is the primary regulator of glucagon secretion, as it is with insulin, but it has an inhibitory rather than stimulatory effect. Glucagon release is stimulated by hypoglycemia and by the amino acids arginine and alanine. GLP-1 inhibits glucagon secretion in vivo, and insulin and somatostatin inhibit glucagon

---

Table 33-3
Clinical and laboratory findings in types of diabetes mellitus

<table>
<thead>
<tr>
<th>PARAMETER</th>
<th>TYPE 1</th>
<th>TYPE 2</th>
<th>TYPE 3C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IDDDM</td>
<td>NIDDM</td>
<td>Pancreateogen</td>
</tr>
<tr>
<td>Ketoacidosis</td>
<td>Common</td>
<td>Rare</td>
<td>Rare</td>
</tr>
<tr>
<td>Hyperglycemia</td>
<td>Severe</td>
<td>Usually mild</td>
<td>Mild</td>
</tr>
<tr>
<td>Hypoglycemia</td>
<td>Common</td>
<td>Rare</td>
<td>Common</td>
</tr>
<tr>
<td>Peripheral insulin sensitivity</td>
<td>Normal or increased</td>
<td>Decreased</td>
<td>Increased</td>
</tr>
<tr>
<td>Hepatic insulin sensitivity</td>
<td>Normal</td>
<td>Normal or decreased</td>
<td>Decreased</td>
</tr>
<tr>
<td>Insulin levels</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Glucagon levels</td>
<td>Normal or high</td>
<td>Normal or high</td>
<td>Low</td>
</tr>
<tr>
<td>PP levels</td>
<td>Normal or low (late)</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>GIP levels</td>
<td>Normal or low</td>
<td>Normal or low</td>
<td>Low</td>
</tr>
<tr>
<td>GLP-1 levels</td>
<td>Normal</td>
<td>Normal or high</td>
<td>Normal or high</td>
</tr>
<tr>
<td>Typical age of onset</td>
<td>Childhood or adolescence</td>
<td>Adulthood</td>
<td>Any</td>
</tr>
</tbody>
</table>

Abbreviations: IDDDM = insulin dependent diabetes mellitus; NIDDM = non–insulin-dependent diabetes mellitus; PP = pancreatic polypeptide; GIP = glucose-dependent insulino tropic polypeptide; GLP-1 = glucagon-like peptide 1.

secretion in a paracrine fashion within the islet. The same neural impulses that regulate insulin secretion also regulate glucagon secretion, so that the two hormones work together in a balance of actions to maintain glucose levels. Cholinergic and alpha-sympathetic fibers stimulate glucagon release, while beta-sympathetic fibers inhibit glucagon release. In pancreaticogenic or type 3c diabetes, glucagon responsiveness to a fall in blood glucose is lost, thereby increasing the risk for hypoglycemia.

Although originally isolated from the hypothalamus, somatostatin is a peptide that is now known to have a wide anatomic distribution, not only in neurons but also in the pancreas, gut, and other tissues. It is a highly conserved peptide hormone, as it is found in lower vertebrates, and is now realized to be of fundamental importance in regulatory processes throughout the body. One gene encodes for a common precursor that is differentially processed to generate tissue-specific amounts of two bioactive products, somatostatin-14, and somatostatin-28. These peptides inhibit endocrine and exocrine secretion and affect neurotransmission, GI and biliary motility, intestinal absorption, vascular tone, and cell proliferation.

Five different somatostatin receptors (SSTRs) have been cloned and the biologic properties of each are different. The hexapeptide and octapeptide analogues such as octreotide bind only to SSTR2, SSTR3, and SSTR5. These analogues have a longer serum half-life, and their potent inhibitory effect has been used clinically to treat both endocrine and exocrine disorders. For example, octreotide has been shown to decrease fistula output and speed the time it takes for enteric and pancreatic fistulas to close.

Endocrine release of somatostatin occurs during a meal. The major stimulant is probably intraluminal fat. Acidification of the gastric and duodenal mucosa also releases somatostatin in isolated perfused organ preparations. Acetylcholine from the cholinergic neurons inhibits somatostatin release.

Pancreatic polypeptide (PP) is a 36-amino-acid, straight-chain peptide discovered by Kimmel in 1968 during the process of insulin purification. Protein is the most potent enteral stimulator of PP release, closely followed by fat, whereas glucose has a weaker effect. Hypoglycemia, whether or not it is insulin induced, strongly stimulates PP secretion through cholinergic stimulation. Phenylalanine, tryptophan, and fatty acids in the duodenum stimulate PP release, probably by inducing CCK, GIP, and secretin release. Vagal stimulation of the pancreas is the most important regulator of PP secretion. In fact, vagotony eliminates the rise in PP levels usually seen after a meal. This can be used as a test for the completeness of a surgical vagotomy or for the presence of diabetic autonomic neuropathy.

PP has been shown to inhibit cholerasis (bile secretion), gallbladder contraction, and secretion by the exocrine pancreas. However, PP’s most important role is in glucose regulation through its regulation of hepatic insulin receptor gene expression. A deficiency in PP secretion due to proximal pancreatectomy, severe chronic pancreatitis, or cystic fibrosis, is associated with diminished hepatic insulin sensitivity due to reduced hepatic insulin receptor availability. This effect is reversed by PP administration.

Recent studies have shown that a fifth islet peptide, ghrelin, is secreted from a distinct population of islet cells, called epsilon cells. Ghrelin also is present in the gastric fundus in large amounts and stimulates growth hormone secretion via growth hormone releasing hormone release from the pituitary. It is an orexigenic, or appetite-stimulating, peptide the plasma levels of which are increased in obesity. Ghrelin has also been shown to block insulin effects on the liver, and inhibits the beta-cell response to incretin hormones and glucose. Therefore, ghrelin secretion from and within the islet may modulate the responses of other islet cells to nutrient and hormonal stimuli.

In addition to the five main peptides secreted by the pancreas, there are a number of other peptide products of the islet cells, including amylin, peptide YY (PYY), and pancreastatin, as well as neuropeptides such as VIP, galanin, and serotonin. Amylin or islet amyloid polypeptide (IAPP) is a 37-amino-acid polypeptide that is predominantly expressed by the pancreatic beta-cells, where it is stored along with insulin in secretory granules. The function of IAPP seems to be the modulation or counterregulation of insulin secretion and function. Pancreastatin is a recently discovered pancreatic islet peptide product that inhibits insulin, and possibly somatostatin release, and augments glucagon release. In addition to this effect on the endocrine pancreas, pancreastatin inhibits pancreatic exocrine secretion.

PYY is structurally related to PP and was initially found in hormone-secreting “L” cells of the small intestine, where it colocalizes with GLP-1. Recently, PYY has been localized to the islets where it appears to regulate insulin secretion through an autocrine mechanism.

**Islet Distribution**

The beta-cells are generally located in the central portion of each islet and make up about 70% of the total islet cell mass. The other cell types are located predominately in the periphery. The delta cells are least plentiful, making up only 5%; the ε-cells make up 10%, and the PP cells make up 15%. In contrast to the acinar cells that secrete the full gamut of exocrine enzymes, the islet cells seem to specialize in the secretion of predominantly one hormone. However, individual islet cells can secrete multiple hormones. For example, the beta-cells secrete both insulin and amylin, which counter regulates the actions of insulin. In reality, more than 20 different hormones are secreted by the islets, and the exact functions of this milieu are very complex. There is diversity among the islets depending on their location within the pancreas. The beta and delta cells are evenly distributed throughout the pancreas, but islets in the head and uncinate process (ventral anlage) have a higher percentage of PP cells and fewer alpha cells, whereas islets in the body and tail (dorsal anlage) contain the majority of alpha cells and few PP cells. This is clinically significant because pancreateoduodenectomy removes 95% of the PP cells in the pancreas. This may partially explain the higher incidence of glucose intolerance after the Whipple procedure compared to a distal pancreatectomy with an equivalent amount of tissue resected. In addition, chronic pancreatitis, which disproportionately affects the pancreatic head, is associated with PP deficiency and pancreaticogenic diabetes.

The relative preponderance of alpha cells in the body and tail of the pancreas explains the typical location of glucagonomas.

**ACUTE PANCREATITIS**

**Definition, Incidence, and Epidemiology**

Acute pancreatitis is an inflammatory disorder of the pancreas that is characterized by edema and, when severe, necrosis. It is a common and challenging disease that can develop local and systemic complications. As such, it ranges from a mild, self-limiting inflammation of the pancreas to severe and critical disease characterized by infected pancreatic necrosis, multiple...
organ failure, and a high mortality. The traditional view is that acute pancreatitis completely resolves with no morphological, functional, or symptomatic sequelae. But necrotizing pancreatitis can leave significant scarring, strictures, and impairment of exocrine and endocrine pancreatic function. The overall clinical outcome has improved over recent decades, even in the absence of specific treatments that target outcome-determining pathophysiology, probably because of a more standardized approach to diagnosis, monitoring, and management.

Acute pancreatitis is the most common inpatient principal gastrointestinal discharge diagnosis in the United States (274,119 in 2009), with an increasing incidence (30% since 2000) and is associated with the highest aggregate inpatient costs at 2.6 billion dollars per year. The crude mortality rate of 1 per 100,000 population ranks it as the 14th most common overall and the 9th most common noncancer cause of gastrointestinal deaths. Worldwide, the incidence of acute pancreatitis ranges from 5 to 80 per 100,000 population, with the highest incidence recorded in Finland and the United States. The incidence of acute pancreatitis also shows significant variation related to the prevalence of etiological factors and ethnicity. The annual incidence of acute pancreatitis in Native Americans is 4 per 100,000 population; in whites, it is 5.7; and in blacks it is 20.7. Smoking is an independent risk factor for acute pancreatitis.

### Etiology

Many factors are causally related to the onset of acute pancreatitis, but the mechanism is often poorly understood. The most common causes are gallstones and alcohol (Table 33-4), accounting for up to 80% of cases, but it is not uncommon to diagnose acute pancreatitis in the absence of these etiological factors (“idiopathic acute pancreatitis”), and it is important that a systematic approach is taken to the identification of other, less common and potentially modifiable factors. The median age at index presentation of acute pancreatitis varies with etiology; with alcohol- and drug-induced pancreatitis presenting in the third or fourth decade compared with gallstone and trauma in the sixth decade. The gender difference is probably more related to etiology: in males alcohol is more often the cause while in females it is gallstones.

#### Gallstones

Evidence that passage of a gallstone is related to the onset of acute pancreatitis comes from the characteristic transient derangement of liver function tests and the high retrieval rate of gallstones from feces within 10 days of an attack of acute pancreatitis compared with those without acute pancreatitis (88% vs. 11%). The mechanism by which small gallstones cause acute pancreatitis in migrating through to the duodenum is not clear. Opie made the seminal observation of a gallstone impacted in the sphincter of Oddi in two fatal cases of acute pancreatitis, which lead to the “common channel” hypothesis. It was proposed that this allowed bile to reflux into the pancreatic duct, but this cannot be reliably reproduced in experimental models. Another proposal was that transient incompetence caused by the passage of a stone through the sphincter might allow duodenal fluid and bile to reflux into the pancreatic duct, but this is not supported by the failure of this to commonly occur after endoscopic sphincterotomy. A third possibility is that acute pancreatitis is due to the gallstone obstructing the pancreatic duct and leading to ductal hypertension. It has been postulated that this backpressure might lead to minor ductal disruption, extravasation of pancreatic juice into the less alkaline interstitium of the pancreas, and promotion of enzyme activation. When gallstones and other etiological factors cannot be identified, there is still the possibility of finding microolithiasis, seen as birefringent crystals, on bile microscopy. This occult microolithiasis is probably responsible for up to half of those with idiopathic acute pancreatitis.

#### Alcohol

Alcohol ingestion is associated with acute pancreatitis, and sustained alcohol ingestion is associated with recurrent acute pancreatitis and development of chronic pancreatitis in susceptible individuals who have been drinking for more than a decade. The type of alcohol consumed is less important than the amount consumed (typically 100–150 grams per day) and the pattern of drinking. It is common for patients with alcohol-associated acute pancreatitis to have a history of excess alcohol consumption prior to the first attack. There are several mechanisms by which ethanol causes acute pancreatitis by acting on the acinar and stellate cells. The acinar cell metabolizes ethanol by oxidative and nonoxidative pathways, and exhibits changes that predispose the cell to autodigestive injury, necroinflammation, and cell death. The stellate cells are activated on exposure to ethanol to a myofibroblast phenotype, stimulating synthesis of proinflammatory mediators and cytokines. Ethanol causes a brief secretory increase followed by inhibition. The secretory burst coupled with ethanol induced spasm of the sphincter of Oddi probably incites acute pancreatitis. Ethanol also induces ductal permeability, which allows prematurely activated enzymes to cause damage to the pancreatic parenchyma. Ethanol also increases the protein content of pancreatic juice and decreases bicarbonate levels and trypsin inhibitor concentration. The formation of protein plugs may also contribute by causing an obstructive element to pancreatic outflow, more often seen in chronic pancreatitis.

<table>
<thead>
<tr>
<th>Table 33-4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Etiologies of acute pancreatitis</strong></td>
</tr>
<tr>
<td>Alcohol</td>
</tr>
<tr>
<td>Biliary tract disease</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
</tr>
<tr>
<td>Hereditary</td>
</tr>
<tr>
<td>Hypercalcemia</td>
</tr>
<tr>
<td>Trauma</td>
</tr>
<tr>
<td>External</td>
</tr>
<tr>
<td>Surgical</td>
</tr>
<tr>
<td>Endoscopic retrograde cholangiopancreatography</td>
</tr>
<tr>
<td>Ischemia</td>
</tr>
<tr>
<td>Hypoperfusion</td>
</tr>
<tr>
<td>Atheroembolic</td>
</tr>
<tr>
<td>Vasculitis</td>
</tr>
<tr>
<td>Pancreatic duct obstruction</td>
</tr>
<tr>
<td>Neoplasms</td>
</tr>
<tr>
<td>Pancreas divisum</td>
</tr>
<tr>
<td>Ampullary and duodenal lesions</td>
</tr>
<tr>
<td>Infections</td>
</tr>
<tr>
<td>Venom</td>
</tr>
<tr>
<td>Drugs</td>
</tr>
<tr>
<td>Idiopathic</td>
</tr>
</tbody>
</table>

Iatrogenic
Acute pancreatitis can occur due to trauma to the ducts or parenchyma after surgical procedures, including biopsy, bile duct exploration, distal gastrectomy, and splenectomy. As the pancreas is susceptible to ischemia, it can also occur secondary to splanchnic hypoperfusion with cardiopulmonary bypass, cardiac transplant, hemorrhagic shock, and major trauma. The most common iatrogenic cause is ERCP in which acute pancreatitis occurs after about 5% to 10% of procedures, and in many series, it is the third most common identified etiological factor. The risk of post-ERCP acute pancreatitis is increased if the contrast agent is infused repeatedly under high pressure by the endoscopist and in patients with sphincter of Oddi dysfunction. Recent evidence demonstrates that the risk can be decreased with prophylactic rectal nonsteroidal drugs, and this may be a better strategy than prophylactic pancreatic duct stenting.

Hereditary Pancreatitis
Hereditary pancreatitis is an autosomal dominant disorder usually related to mutations of the cationic trypsinogen gene (PRSS1). Mutations in this gene cause premature activation of trypsinogen to trypsin and causes abnormalities of ductal secretion, both of which promote acute pancreatitis. Mutations in the SPINK1 protein, which blocks the active binding site of trypsin, is likely to also have a role in predisposing to acute pancreatitis. Variations in penetration and phenotype are common, and there are many other mutations that have become implicated. Mutant enzymes activated within acinar cells can overwhelm the first line of defense (pancreatic secretory trypsin inhibitor) and resist backup defenses (e.g., proteolytic degradation, enzyme Y, and trypsin itself) allowing activated mutant cationic trypsin to trigger the entire zymogen activation cascade.

Tumors
A pancreatic or periampullary tumor should be considered in patients with idiopathic acute pancreatitis, especially in those over 50 years old. Approximately 1% to 2% of patients with acute pancreatitis have a pancreatic tumor, and an episode of acute pancreatitis can be the first clinical indicator. Cross-sectional imaging after the resolution of the acute pancreatitis is required.

Hyperlipidemia
Patients with types I and V hyperlipoproteinemia can experience episodes of abdominal pain, and these often occur in association with marked hypertriglyceridaemia. Lipase is thought to liberate toxic fatty acids into the pancreatic microcirculation, leading to microcirculatory impairment and ischemia. Dietary modifications and drug treatment are used to lower triglycerides.

Drugs
Isolated cases of acute pancreatitis have been associated with exposure to certain drugs, such as thiazide diuretics, furosemide, estrogen replacement therapy, and steroid therapy in children. In addition, certain chemotherapy agents and anti-immune drugs have been associated with acute pancreatitis, and lipid-based drugs or solutions, such as propofol, have been shown to cause acute pancreatitis.

Pathophysiology
Acute pancreatitis occurs in various degrees of severity, the determinants of which are multifactorial. The generally prevalent belief today is that pancreatitis begins with the activation of digestive zymogens inside acinar cells, which cause acinar cell injury. Studies suggest that the ultimate severity of the resulting pancreatitis may be determined by the events that occur subsequent to acinar cell injury. These include inflammatory cell recruitment and activation, as well as generation and release of cytokines and other chemical mediators that cause systemic inflammation and multiple organ dysfunction/failure (Figure 33-10).

Precipitating Events
In 1896, Chiari proposed that pancreatitis was due to the premature, intrapancreatic activation of digestive enzymes, resulting in “auto-digestion” of the organ. Since then the

---

**Figure 33-10.** Schema of key loco-regional pathophysiological events in the pancreas and intestine and how they interact to drive the severity and outcome of acute pancreatitis. (Adapted with permission from Flint RS, Windsor JA: The role of the intestine in the pathophysiology and management of severe acute pancreatitis, HPB (Oxford). 2003;5(2):69-85.)
Intra-acinar activation of zymogens has been demonstrated consistently in multiple animal models of acute pancreatitis and is considered a key precipitating event.\textsuperscript{37,38} The key role of trypsin activation in acute pancreatitis has gained additional support from recent studies showing that mice lacking trypsinogen-7 (the isoform of trypsinogen that is activated during acute pancreatitis in mice) have significantly less pancreatic injury during acute pancreatitis\textsuperscript{39} and that intra-acinar expression of active trypsin causes pancreatitis in mice.\textsuperscript{40} The role of trypsin activation in the pathophysiology of acute pancreatitis has also been suggested by clinical studies. Hereditary pancreatitis is associated with mutations that lead to elevated intracellular trypsin activation,\textsuperscript{41} and activation of trypsinogen causes clinical pancreatitis.\textsuperscript{42}

Significant progress has been made in understanding the mechanisms by which injurious stimuli lead to intra-acinar activation of trypsinogen and autodigestion of the gland (Figure 33-11). Under physiologic conditions, several protective mechanisms have evolved to prevent autodigestion of the pancreas by these enzymes. This includes synthesis of enzymes as inactive precursors, separation of the site of production, and activation of the enzymes and presence of trypsin inhibitors in the pancreas. It is thought that acute pancreatitis occurs when these protective mechanisms are overwhelmed by erroneously activated enzymes, causing injury. It has been shown that intra-acinar activation of trypsinogen goes hand-in-hand with inhibition of acinar secretion.\textsuperscript{43,44} Furthermore, with injurious stimuli the zymogens responsible for initiating the disease are not secreted outside, but colocalize with cytoplasmic vacuoles that contain lysosomal enzymes such as cathepsin B\textsuperscript{45} that activate trypsinogen. Thus, inhibition of cathepsin B by pharmacological inhibitors\textsuperscript{46} or by genetic deletion of cathepsin B eliminates trypsin activation and decreases the severity of pancreatitis in animal models.\textsuperscript{47} What leads to the colocalization of zymogens and lysosomal hydrolases is unclear, but injurious stimuli leading to sustained cytosolic calcium increase have been indicted. Blocking this calcium increase prevents colocalization and activation of trypsin, and it decreases injury due to pancreatitis.\textsuperscript{48} Based on these data, pre-ERCP supplementation of magnesium, a natural antagonist of calcium, is currently being evaluated as a strategy to decrease post-ERCP pancreatitis.\textsuperscript{49} Recent work has led to the novel hypothesis that the lysosomal hydrolase cathepsin B activates trypsinogen to trypsin within the colocalization vacuoles. Trypsin then permeabilizes these colocalization vacuoles causing the release of cathepsin B into the cytosol. Once in the cytosol, cathepsin B initiates apoptotic cell death by permeabilizing mitochondrial membranes, which allows cytochrome C to be released into the cytosol. This initiates the apoptotic cascade and results in the apoptotic death of the acinar cells\textsuperscript{50} (see Figure 33-11).

### Intrapancreatic Events

Although intra-acinar events initiate acute pancreatitis, events occurring subsequent to acinar cell injury determine the severity of pancreatitis. Activated neutrophils are attracted and activated in the pancreas releasing superoxide (the respiratory burst) and proteolytic enzymes (cathepsins, elastase, and collagenase) that cause further pancreatic injury. In addition, macrophages release cytokines (including tumor necrosis factor-alpha (TNF-$\alpha$) and interleukins (IL-1, IL-2, IL-6, and IL-8) that mediate local and systemic inflammation.\textsuperscript{38}

These inflammatory mediators cause an increased pancreatic vascular permeability, leading to hemorrhage, edema, and

![Figure 33-11](image-url). Schematic representation of the acinar cell events in acute pancreatitis. When acinar cells are pathologically stimulated, their lysosomal (L) and zymogen (Z) contents colocalize, and consequently trypsinogen is activated to trypsin by cathepsin B. Increased cytosolic calcium is required for colocalization. Once trypsin has permeabilized the contents of the cytosol, cathepsin B and other contents of these colocalized organelles are released. Once in the cytosol, cathepsin B activates apoptosis by causing cytochrome C to be released from the mitochondria. Activation of PKC results in a sudden activation of nuclear factor kappa beta (NFkB), which in turn triggers the release of cytokines that attract inflammatory response cells that mediate local and systemic inflammation cascades.
microthrombi. Fluid may collect in and around the pancreas. The failure of the pancreatic microcirculation, a feature of more severe acute pancreatitis, results in pancreatic hypoperfusion and necrosis. Acute inflammation of the pancreatic parenchyma and peripancreatic tissues, but with no recognizable necrosis, is termed interstitial edematous pancreatitis.25 When necrosis is present, as evidenced by pancreatic hypoperfusion with contrast CT, it is termed necrotizing pancreatitis (Figure 33-12). The updated morphological definitions and the contrast enhanced CT criteria for the diagnosis of the local complications of acute pancreatitis are in the revised Atlanta statement51 and summarized in Table 33-5.52

**Systemic Events**

An important aspect of the pathophysiology of acute pancreatitis is the mechanism by which events occurring in the pancreas induce systemic inflammation and multiorgan failure. The NFκB-dependent inflammatory pathway is an important mechanism (see Figure 33-11). Activation of NFκB parallels trypsin activation in acute pancreatitis but appears to be independent of it, as it occurs in trypsin knockout mice.50 Sustained calcium increase, which leads to trypsinogen activation, is critical for NFκB activation since attenuation of cytosolic calcium abrogates NFκB activation.53 Once activated, NFκB regulates synthesis of multiple cytokines and chemokines, leading to recruitment of various inflammatory cells that then magnify and propagate systemic inflammation.54 The infiltrating neutrophils can also further augment the pancreatic injury.55,56 Inhibition of many of these cytokines have led to reduced local and systemic injury in animal models of acute pancreatitis. But these results did not translate to clinical improvement.57 New strategies are required to reduce the systemic inflammatory response.58-60

Organ failure can develop at any stage of acute pancreatitis, associated with an overwhelming proinflammatory response early, or later secondary to the development of infected local complications. The drivers of the systemic response are poorly understood, although factors include the elaboration of proinflammatory cytokines, and it appears that mesenteric lymph, bypassing the liver and containing these constituents, may contribute to the development of organ failure64 (see Figure 33-10). The development of pancreatic necrosis, the breakdown of the intestinal barrier, and the suppression of the immune response through the compensatory inflammatory response contribute to the development of infected pancreatic necrosis, the incidence of which peaks in the third to fourth week. This is usually associated with deterioration in the patient’s condition and may be associated with the late development of the systemic inflammatory response syndrome (SIRS) and multiorgan dysfunction syndrome/failure (MODS/F).

Organ failure is scored using the Marshall or Sequential Organ Failure Assessment (SOFA) systems (Table 33-6). The three organ systems most frequently involved are cardiovascular, respiratory, and renal. Multiple organ failure is defined as two or more organs registering two or more points on these scoring systems.51 Monitoring organ failure over time and in response to treatment is important in the care and timing of intervention in these patients.

**Management of Acute Pancreatitis**

**General Considerations.** The management of acute pancreatitis covers a wide spectrum of severity. All patients with suspected acute pancreatitis should be admitted to hospital. Those with mild acute pancreatitis usually remain in hospital for less than a week, while those with severe and critical acute pancreatitis may require many weeks or months of intensive treatment. The risk of mortality reflects this spectrum of severity. The risk is less than 1% for those with mild disease, increasing to around 10% for those with moderate disease, but for severe (20–40%) and critical (>50%) disease the mortality risk is much higher. The earlier identification of these high-risk categories and their transfer to specialized centers is an important priority of management.52

The management of acute pancreatitis is multidisciplinary, and it is important that this is coordinated care plan is carefully supervised.63 The essential requirements of the management

---

**Table 33-5**

**Local complications of acute pancreatitis**

<table>
<thead>
<tr>
<th>CONTENT</th>
<th>ACUTE (≤4 WEEKS, NO DEFINED WALL)</th>
<th>CHRONIC (≥4 WEEKS, DEFINED WALL)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NO INFECTION</td>
<td>INFECTION</td>
</tr>
<tr>
<td>Fluid</td>
<td>Acute pancreatic fluid collection (APFC)</td>
<td>Infected APFC</td>
</tr>
<tr>
<td>Solid ± fluid</td>
<td>Acute necrotic collection (ANC)</td>
<td>Infected ANC</td>
</tr>
</tbody>
</table>

of acute pancreatitis are accurate diagnosis, appropriate triage, high-quality supportive care, and monitoring for and treatment of complications.64

Diagnosis
The diagnosis of acute pancreatitis requires the patient to present with abdominal pain consistent with acute pancreatitis (acute onset of a severe constant epigastric pain that often radiates through to the mid back) and the elevation of serum amylase or lipase (>3 times upper limit of normal). Imaging (usually by contrast-enhanced CT scanning) is only required for the diagnosis of acute pancreatitis when these diagnostic criteria are not met.51 Because of the many causes of hyperamylasemia, it is important to use either the pancreatic isoenzyme of amylase or lipase.63

The serum amylase concentration increases almost immediately with the onset of disease and peaks within several hours and remains elevated for 3 to 5 days.64 There is no correlation between the extent of serum amylase elevation and severity of pancreatitis; in fact, a milder form of acute pancreatitis is often associated with higher levels of serum amylase compared with that in a more severe form of the disease. It is important to note that hyperamylasemia can also occur in association with other diseases. For example, it can occur in a patient with small bowel obstruction, perforated duodenal ulcer, or other intra-abdominal inflammatory conditions. In contrast, a patient with acute pancreatitis can have a normal serum amylase level, for several different reasons. In patients with hyperlipidemia, values might appear to be normal because of interference with chemical determination of serum amylase by lipids. In some cases, urinary clearance of pancreatic enzymes from the circulation increases during pancreatitis, meaning that urinary levels may be more sensitive than serum levels. For these reasons, it is recommended that amylase concentrations also be measured in the urine. Urinary amylase levels usually remain elevated for several days after serum levels have returned to normal. In patients with severe pancreatitis associated with significant necrosis, the pancreas may not have the capacity to release large amounts of enzymes into the circulation. With more severe disease, there is also hemoconcentration from third space fluid loss, and this can affect the serum concentration of amylase.

The clinical signs of acute pancreatitis include abdominal tenderness, often with signs of peritonitis in the upper abdomen. Rarely, pancreatic fluid and bleeding from the pancreas into the retroperitoneum may result in a bruise-like discoloration around the umbilicus (Cullen’s sign) or in the flanks (Grey Turner’s sign). Another rare sign is tetany as a result of hypocalcaemia. In addition to hemoconcentration, patients with acute pancreatitis often have azotemia with elevated blood urea nitrogen and creatinine levels, hyperglycemia, and hypoalbuminemia.

Pain Management
Pain is the cardinal symptom of acute pancreatitis, and its relief is a clinical priority. There is a lack of high-quality evidence to guide the choice of analgesic. Because of unpredictable absorption, analgesia should be administered intravenously, at least at the outset and before oral intake has been established. Those with mild pain can usually be managed with a nonsteroidal anti-inflammatory drugs (e.g., metamizole 2 g/8 h IV), while those with more severe pain are best managed with opioid analgesia (e.g., buprenorphine 0.3 mg/4 h IV). Administration of buprenorphine, pentazocine, procaine hydrochloride, and meperidine are all of value in controlling abdominal pain. Morphine is to be avoided because of its potential to cause sphincter of Oddi spasm.

Predicting Severity
Whereas classification relates to the present or past severity of acute pancreatitis, prediction is about the future and ultimate severity and outcome of the patient. Accurately predicting acute

| Table 33-6
Sequential organ failure assessment (SOFA) score in acute pancreatitis |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Respiration</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(PaO₂/FIO₂) (mmHg)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;400</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤400</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤200 with respiratory support</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤100 with respiratory support</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Coagulation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Platelets (x10⁹ per μL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;150</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤150</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Liver</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilirubin (μmol/L)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20–32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33–101</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>102–204</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;204</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cardiovascular</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypotension</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No hypotension</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAP &lt;70 mmHg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dopamine ≤5 or dobutamine (any dose)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dopamine &gt;5 or epi ≤0.1 or norepi ≤0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dopamine &gt;15 or epi &gt;0.1 or norepi &gt;0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Central nervous system</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glasgow coma score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13–14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10–12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6–9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Kidney</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creatinine (μmol/L) or urine output</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;110</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>110–170</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>171–299</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>300–440 or &lt;500 mL/day</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;440 or &lt;200 mL/day</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

 Shall be read as "Table 33-6".

Note: Adrenergic agents administered for at least 1 h (doses given in μg/kg per min). A score of 2 or more in any two systems indicates the presence of multiple organ failure.

Abbreviations: MAP = mean arterial pressure; Epi = epinephrine; Norepi = norepinephrine.
Table 33-7

Ranson’s prognostic signs of pancreatitis

<table>
<thead>
<tr>
<th>Criteria for acute pancreatitis not due to gallstones</th>
<th>At admission</th>
<th>During the initial 48 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age &gt;55 y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WBC &gt;16,000/mm³</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blood glucose &gt;200 mg/dL</td>
<td>Serum LDH  &gt;350 IU/L</td>
<td>Arterial PO₂  &lt;60 mmHg</td>
</tr>
<tr>
<td>Serum AST &gt;250 U/dL</td>
<td>Serum calcium &lt;8 mg/dL</td>
<td>Blood glucose &gt;220 mg/dL</td>
</tr>
<tr>
<td>Serum AST &gt;250 U/dL</td>
<td>WBC &gt;18,000/mm³</td>
<td>Base deficit &gt;4 mEq/L</td>
</tr>
<tr>
<td>Estimated fluid sequestration &gt;6 L</td>
<td>Estimated fluid sequestration &gt;6 L</td>
<td></td>
</tr>
</tbody>
</table>

Criteria for acute gallstone pancreatitis

<table>
<thead>
<tr>
<th>At admission</th>
<th>During the initial 48 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age &gt;70 y</td>
<td>Hematocrit fall &gt;10 points</td>
</tr>
<tr>
<td>WBC &gt;18,000/mm³</td>
<td>BUN elevation &gt;2 mg/dL</td>
</tr>
<tr>
<td>Blood glucose &gt;220 mg/dL</td>
<td>Serum calcium &lt;8 mg/dL</td>
</tr>
<tr>
<td>Serum LDH &gt;400 IU/L</td>
<td>Base deficit &gt;5 mEq/L</td>
</tr>
<tr>
<td>Serum AST &gt;250 U/dL</td>
<td>Estimated fluid sequestration &gt;4 L</td>
</tr>
</tbody>
</table>

Note: Fewer than three positive criteria predict mild, uncomplicated disease, whereas more than six positive criteria predict severe disease with a mortality risk of 50%.

Abbreviations: AST = aspartate transaminase; BUN = blood urea nitrogen; LDH = lactate dehydrogenase; PO₂ = partial pressure of oxygen; WBC = white blood cell count.


pancreatitis severity is important in making triage decisions about whether a patient should be transferred to a tertiary hospital or an intensive care unit and in making decisions about fluid therapy and whether an ERCP is indicated, as well as other issues. There is a very long history of attempts to find prognostic or predictive markers that accurately stratify the risk, with the most widely used being the Ranson’s criteria (Table 33-7) or modified Glasgow criteria. Both use clinical and biochemical parameters scored over the first 48 hours of admission. When there are three or more positive criteria, the disease is considered “predicted severe.” There are many other approaches to predicting severity. At 24 hours after admission an APACHE II score of 8 or more or a serum C-reactive protein level of >150 mg/dL has a similar accuracy in predicting severity as Ranson’s criteria. The more recently proposed Bedside Index for Severity of Acute Pancreatitis (BISAP) is calculated from blood urea nitrogen (> 25 mg/dL), impaired mental status (GCS <15), presence of systemic inflammatory response syndrome, age >60 years, and pleural effusion. Although it has the advantage of simplicity and can be performed within the first 24 hours of admission, it performed no better than other predictors. The presence of SIRS also has prognostic significance. There remains some controversy as to how important obesity is as a risk factor for severe and critical acute pancreatitis. Another approach has been taken in seeking to predict those with the “harmless acute pancreatitis score” using three factors that can be determined on admission: absence of rebound tenderness or guarding, normal hematocrit, and normal serum creatinine. The accuracy of this approach appears to be over 90% and triages most patients away from intensive care.

Unfortunately, these and many other single and combined predictors of severity have an accuracy of around 70%. This means that there is misclassification error of 30% that limits the value in predicting the severity of acute pancreatitis in individual patients. In the absence of any new biomarkers of pancreatitis severity, making better use of existing predictors through sequencing tests, combining tests, or using artificial neural network methodologies has shown some promise. Scoring systems should augment clinical judgment, but not replace it. Over the first 2 to 3 days the clinician must be alert to patients with an elevated BUN or creatinine and/or persistent SIRS after adequate fluid resuscitation because these patients are at risk of developing severe acute pancreatitis.

Classification of Severity

Accurately classifying or staging acute pancreatitis severity is important for clinical decision-making, communication, and enrolment into trials. The wide spectrum of pancreatitis severity was not captured in the previous binary classifications (mild or severe). The key determinants of severity are local complications (absent, sterile, or infected) and systemic complications (absent, transient organ failure, persistent organ failure). Two classification systems have recently been proposed: the three grades (mild, moderately severe, and severe) of the Revised Atlanta Criteria (RAC) and the four categories (mild, moderate, severe, critical) of the Determinants Based Classification (DBC) (Table 33-8). The DBC

Table 33-8

Definitions for the classification of acute pancreatitis severity according to Revised Atlanta Classification and the Determinants Based Classification. (Transient organ failure has a duration of <48 hours, persistent organ failure has duration of >48 hours.)

<table>
<thead>
<tr>
<th>Determinant Based Classification (2012)</th>
<th>Revised Atlanta Classification (2013)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild</td>
<td>No local complication</td>
</tr>
<tr>
<td>No systemic complication</td>
<td>No local complication</td>
</tr>
<tr>
<td>No systemic complication</td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>Sterile local complication and/or</td>
</tr>
<tr>
<td>Transient organ failure</td>
<td>Local or systemic complications without persistent organ failure or exacerbation of preexisting comorbidity</td>
</tr>
<tr>
<td>Severe</td>
<td>infected local complication or</td>
</tr>
<tr>
<td>Persistent organ failure</td>
<td>Persistent organ failure (single or multiple)</td>
</tr>
<tr>
<td>Critical</td>
<td>Infected local complication and</td>
</tr>
<tr>
<td>Persistent organ failure</td>
<td></td>
</tr>
</tbody>
</table>
was developed on the principle of casual inference, derived by meta-analysis, and refined by an international multidisciplinary process, and both classifications have been independently validated, which suggests that they are broadly equivalent. More recently a prospective multicenter study has modified the DBC to address the ongoing issue that patients considered to have severe acute pancreatitis represent several subgroups with different morbidity, mortality, and intervention profiles (Table 33-9).

The classification of patient severity is helpful in tracking the clinical trajectory of a patient, and it can be applied on a daily or more frequent basis. It can also be used in retrospect for audit purposes.

### Determining Etiology

The history of alcohol ingestion must be ascertained and preferably confirmed with blood ethanol levels. Gallstones should be investigated by ultrasonography. A gallstone etiology is more likely in females over the age of 50 with an elevation of alkaline phosphatase (>300 iu/L), alanine transferase (>100 iu/L), and amylase (>4000 iu/L). In the absence of gallstones and alcohol, a systematic approach to the identification of another factor will include taking a history of drugs, trauma, ERCP, infection, and measuring serum triglycerides, calcium, and others (see Table 33-4).

### Fluid Resuscitation

Fluid therapy to restore and maintain circulating blood volume is the most important intervention in the early management of acute pancreatitis. However, a recent systematic review has shown that the evidence base for fluid therapy is scant, and most recommendations are based on expert opinion. It is not known which fluid to give, how aggressively to administer it, or what goal to use to guide and monitor the response to it. While there are proponents for vigorous fluid therapy (5–10 mL per kilogram per hour), especially in the first 24 hours, and for specific resuscitation goals, it is probably best to resuscitate with a balanced crystalloid and aim to restore normal blood volume, blood pressure, and urine output. In one study, lactated Ringer’s solution was superior to normal saline in reducing the systemic inflammatory response. Caution needs to be exercised in those with cardiac and renal disease and in the elderly, where the risks of over-resuscitation are greater.

### Nutritional Support

In contrast to analgesia and fluid therapy, there is a sound evidence base for nutritional support in acute pancreatitis. It is no longer acceptable to “rest the pancreas” by avoiding enteral nutrition, now the mainstay of nutritional support. Parenteral nutrition is now known to be more expensive, riskier, and not more effective than enteral nutrition and should only be offered if the patient’s calculated nutritional requirements cannot be achieved by the enteral route. Early initiation of enteral nutrition (within the first 24 hours of admission) is not superior to delaying an oral diet until 72 hours. If this is not tolerated over 48 to 72 hours, then nasogastric tube feedings can be started and increased in step-wise fashion over 2 to 3 days. The tube can be advanced to the jejunum, by endoscopy or fluoroscopy, if there is evidence of feeding intolerance. A delay in commencing enteral nutrition may contribute to the development of intestinal ileus and feeding intolerance, but aggressive early enteral feeding, particularly before adequate resuscitation, may put the patient at risk of nonocclusive mesenteric ischemia. There is no evidence to support the use of elemental or immune-enhancing formulas over standard polymeric formulas. In predicted mild acute pancreatitis the recommencement of oral fluids and then food was delayed until resolution of abdominal pain and normalization of serum levels of amylase, but it appears safe to allow patients to resume intake ad libitum (i.e., patient-controlled nutrition). If after 3 to 5 days there is evidence of feeding intolerance, tube feeding should be commenced.

### Cross-Sectional Imaging

It may be necessary to perform a CT scan to diagnose acute pancreatitis in patients who are severely ill or in those presenting with undifferentiated abdominal pain. But there is no advantage in using CT scanning to predict the severity of acute pancreatitis. The primary purpose of cross-sectional imaging is the diagnosis of local complications; in particular, the development and extent of pancreatic necrosis and the different collections (see Table 33-5). CT scanning is also important to guide the insertion of percutaneous drains, now assuming a greater role in the management of the local complications (discussed later). Magnetic resonance imaging (MRI) is superior to CT scanning in detecting any solid content within collections (Figure 33-13). And when a bleed is suspected, in association with a local complication, an arterial phase CT scan (CTa) is useful in detecting a pseudoaneurysm, active bleeding, and/or hematoma.

### Table 33-9

The Modified Determinant Based Classification (MDBC) (Acevedo) of acute pancreatitis severity compared with the Revised Atlanta Classification (RAC) and Determinants Based Classification (DBC). Note that the DBC has a narrower definition for local complications than RAC, leading to a slightly broader range of mild acute pancreatitis in this table.

<table>
<thead>
<tr>
<th>RAC</th>
<th>DBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>(No OF, No LC)</td>
<td>(TOF or IN)</td>
</tr>
<tr>
<td></td>
<td>(POF and SN)</td>
</tr>
<tr>
<td>(TOF and/or LC)</td>
<td>(POF)</td>
</tr>
<tr>
<td></td>
<td>Mild</td>
</tr>
<tr>
<td>(TOF and/or LC)</td>
<td>Severe</td>
</tr>
<tr>
<td>(POF)</td>
<td></td>
</tr>
<tr>
<td>DBC Excluded</td>
<td>(No OF, No LC)</td>
</tr>
<tr>
<td>(TOF and/or SN)</td>
<td>(POF or IN)</td>
</tr>
<tr>
<td>(IN without POF)</td>
<td>(POF without IN)</td>
</tr>
<tr>
<td>(POF without IN)</td>
<td>(POF and IN)</td>
</tr>
<tr>
<td>Group 1</td>
<td>Group 2</td>
</tr>
<tr>
<td>Group 2</td>
<td>Group 3</td>
</tr>
<tr>
<td>Group 3</td>
<td>Group 4</td>
</tr>
</tbody>
</table>

Abbreviations: OF = organ failure; LC = local complication; TOF = transient organ failure; POF = persistent organ failure; SN = sterile necrosis; IN = infected necrosis.

---

...
Therapeutic Endoscopic Retrograde Cholangiopancreatography

Randomized trials have demonstrated that early ERCP (within 24 or 48 hours of admission) reduce complications, but not mortality, in patients with predicted severe gallstone associated acute pancreatitis. While the benefits of this invasive modality are clear (e.g., treatment of cholangitis and release of impacted stone), this can be offset by the risks of increasing the severity of pancreatitis, bleeding, cholangitis, and perforation. More recent evidence has suggested that early ERCP confers no benefit in the absence of concomitant cholangitis, as the offending common duct stone usually passes before ERCP can be performed. This may be evidenced by improvement in the liver function tests over the first 2 to 3 days. If there is persistent cholestasis, an MRCP can be used to detect a common duct stone and can be used as a prerequisite for attempting an ERCP. Persistent cholestasis without cholangitis may require an ERCP but not usually in the acute setting.

Antibiotics

Although the use of broad-spectrum antibiotics to treat established infection in acute pancreatitis is a well-established practice, there has been considerable controversy surrounding the use of prophylactic antibiotics. The overuse of antibiotics has been associated with a documented rise in fungal infections and resistant organisms. Overall, it appears that the most recent and generally better designed studies do not support the use of prophylactic antibiotics to reduce the frequency of pancreatic infectious complications, surgical intervention, and death.

Managing Local Complications

Vigilance is required for the timely and accurate diagnosis of local complications. The decisions regarding how and when to intervene are often difficult. While guided by the information gained by cross-sectional imaging, the decision to intervene is based on the clinical status and trajectory of the patient and the poor response to maximal intensive care support. This means close monitoring of the patient by serial examination, supplemented by regular measurement of inflammatory markers (e.g., C-reactive protein) and a pancreatic protocol CT scan if a local complication is suspected and intervention considered warranted. In practice, intervention is delayed in order to allow demarcation and to reduce the risk of bleeding, disseminated infection, and collateral damage to adjacent organs by an intervention (Figure 33-14). Appreciation of this has resulted in a notable trend toward delayed intervention.

Figure 33-13. Corresponding computed tomography (CT) (A) and MR (B) images of a patient with a symptomatic pseudocyst. CT image reveals a well circumscribed homogenous collection (arrows) exerting mass effect on antrum of stomach. The T2-weighted MR image clearly distinguishes necrotic pancreas (black arrows) from fluid (white arrows). (Reproduced with permission from Bollen TL: Imaging of acute pancreatitis: update of the revised Atlanta classification, Radiol Clin North Am. 2012 May;50(3):429-445.)

Figure 33-14. Operative view of infected acute pancreatitis. Peripancreatic infection, characterized by mucopurulent exudate, extends far beyond the boundaries of the pancreas in the retroperitoneum.
intervention, now uncommon before 3 to 4 weeks from the onset of symptoms. An important emerging approach is the increasing use of percutaneous catheter drainage in patients with suspected infected collections. Fine-needle aspiration is now rarely used to confirm infection because the insertion of a needle at the time of planned drainage allows confirmation of the suspected infection. Preemptive drainage with one or more catheters often produces improvement or stabilization of the patient’s overall clinical status. In this way, drainage “buys time” and allows the lesion to become more walled off and safer to treat. Recent data suggests that primary percutaneous catheter drainage may be the only intervention required in a third to a half of patients and that this proportion might increase further if there were a policy of regular catheter exchange, upsizing, and irrigation. A proportion of patients do, however, require further treatment when they fail to respond and there is a wide array of minimally invasive options to choose from. These interventions can be classified on the basis of the method of visualization, route taken to the lesion, and the purpose of the intervention. In practice, the approach taken will depend on local expertise and equipment as well as the location and type of the specific local complication. A large Dutch randomized trial has shown that open surgical techniques should only be considered in those who fail to respond to the step-up approach, that is, prior percutaneous drainage and minimally invasive intervention. The exception is to be found in the rare situation where an abdominal compartment syndrome requires open decompression, but this is usually earlier than the optimal time to intervene for local complications. A recent landmark randomized trial has compared two minimally invasive techniques, endoscopic transgastric drainage, and the videoscope-assisted retroperitoneal debridement through a flank incision (see Figure 33-15). The data shows that the former approach is superior, although the latter has a role when the walled-off necrosis is remote from the stomach or duodenum, as in the left flank.

The management of an acute noninfected pseudocyst is usually conservative, as about half of these will resolve spontaneously. When symptoms of pain or the inability to eat persist or infection occurs, intervention is required. The indications for intervention are therefore no longer based on size and duration alone. Pseudocysts persist because of communication with the main pancreatic duct and/or distal ductal stenosis. Percutaneous drainage should be avoided in this situation because of the risk of external pancreatic fistula. EUS-guided internal drainage into stomach or duodenum or transpapillary stenting is the preferred approach.

Managing Organ Failure

The specific management of multiple organ failure is beyond the scope of this chapter. The early identification of organ dysfunction and failure is important because it is a key determinant of severity and outcome and to facilitate the timely transfer of the patient to an intensive care unit to optimize management, provide organ support and allow more intensive monitoring. The severity of organ failure can be scored (see Table 33-6). The responsiveness of organ failure to resuscitation over the first 48 hours is an important prognostic clue; those that respond have transient organ failure and have a better outlook than those who do not respond and have persistent organ failure. Organ failure that develops later in the disease course is usually

secondary to infection of a local complication and should be managed accordingly (see Figure 33-9).

**Cholecystectomy**

While it is widely accepted that cholecystectomy is essential to prevent recurrent gallstone associated pancreatitis, the question relates to the timing of it. Index cholecystectomy, done in the same admission and prior to discharge, appears safe and can almost always be accomplished laparoscopically. But index cholecystectomy is not suitable for all patients, particularly some who have had local pancreatic complications, which includes a large inflammatory mass that extends into the porta hepatis. These patients may require an interval cholecystectomy after resolution of the inflammatory process. If surgery is required for the management of local complications, then a cholecystectomy is often performed at that time.

**Diabetes**

Recent evidence indicates that prediabetes and diabetes are common after acute pancreatitis and occur in nearly 40% of patients after hospital discharge. The prevalence of newly diagnosed diabetes is much higher after acute pancreatitis (23%) than the prevalence of diabetes in the general population (4-9%). The risk of diabetes increases by at least twofold after 5 years as compared with 12 months. Interestingly, the severity of acute pancreatitis appears to have minimal effect on risk of diabetes. The implication is that patients recovered from an attack of acute pancreatitis may need follow-up and screening for glucose intolerance.

The management of acute pancreatitis remains a formidable challenge (Table 33-10) due to the variety and severity of the many associated complications (Table 33-11), and continues to evolve. Although specific treatments for acute pancreatitis remain elusive, progress has been made in the management of pain, fluid resuscitation, antibiotic prophylaxis, enteral nutrition, therapeutic ERCP, and cholecystectomy. Progress has also been made in the intensive care management of systemic complications and in the development of less invasive interventions for the treatment of local complications, particularly infected pancreatic necrosis.

---

**Table 33-10**

**Algorithm for the evaluation and management of acute pancreatitis**

<table>
<thead>
<tr>
<th align="left">1. Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td align="left">- History of abdominal pain consistent with acute pancreatitis</td>
</tr>
<tr>
<td align="left">- &gt;3x elevation of pancreatic enzymes</td>
</tr>
<tr>
<td align="left">- CT scan if required to confirm diagnosis</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th align="left">2. Initial assessment/management (first 4 hrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td align="left">- Analgesia</td>
</tr>
<tr>
<td align="left">- Fluid resuscitation</td>
</tr>
<tr>
<td align="left">- Predict severity of pancreatitis</td>
</tr>
<tr>
<td align="left">- Ranson’s criteria</td>
</tr>
<tr>
<td align="left">- HAPS score</td>
</tr>
<tr>
<td align="left">- Assess systemic response</td>
</tr>
<tr>
<td align="left">- SIRS score</td>
</tr>
<tr>
<td align="left">- SOFA (organ failure)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th align="left">3. Reassessment/management (4 to 6 hrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td align="left">- Assess response to fluid resuscitation</td>
</tr>
<tr>
<td align="left">- mean arterial pressure</td>
</tr>
<tr>
<td align="left">- heart rate</td>
</tr>
<tr>
<td align="left">- urine output</td>
</tr>
<tr>
<td align="left">- hematocrit</td>
</tr>
<tr>
<td align="left">- Determine etiology</td>
</tr>
<tr>
<td align="left">- Ultrasound for gallstones/sludge</td>
</tr>
<tr>
<td align="left">- History of alcohol consumption</td>
</tr>
<tr>
<td align="left">- Laboratory evaluation of other causes</td>
</tr>
<tr>
<td align="left">- MRCP and/or Urgent ERCP if concomitant cholangitis is present</td>
</tr>
<tr>
<td align="left">- not for cholestasis or predicted severe disease per se</td>
</tr>
<tr>
<td align="left">- Transfer to ICU or specialist center as needed</td>
</tr>
<tr>
<td align="left">- Deterioration or failure to respond to initial management</td>
</tr>
<tr>
<td align="left">- Intensive support for persistent organ failure</td>
</tr>
<tr>
<td align="left">- Commence enteral nutrition</td>
</tr>
<tr>
<td align="left">- Once normovolemia restored (usually after 6 hours)</td>
</tr>
<tr>
<td align="left">- Commence via NG tube if no gastric stasis</td>
</tr>
<tr>
<td align="left">- No prophylactic antibiotics or probiotics</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th align="left">4. Conservative management and monitoring (at least daily)</th>
</tr>
</thead>
<tbody>
<tr>
<td align="left">- Clinical evaluation</td>
</tr>
<tr>
<td align="left">- Assess cardiovascular, respiratory, and renal function</td>
</tr>
<tr>
<td align="left">- Detect peritonitis and abdominal compartment syndrome</td>
</tr>
<tr>
<td align="left">- Daily C-reactive protein</td>
</tr>
<tr>
<td align="left">- Classify severity (mild, moderate, severe, critical)</td>
</tr>
<tr>
<td align="left">- Detect intolerance of NG EN</td>
</tr>
<tr>
<td align="left">- Advance tube for NJ feeding if needed</td>
</tr>
<tr>
<td align="left">- Consider supplemental parenteral nutrition by day 4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th align="left">5. Indications for “pancreatic protocol CT scan” (rarely in first week)</th>
</tr>
</thead>
<tbody>
<tr>
<td align="left">- For significant clinical deterioration and elevated CRP</td>
</tr>
<tr>
<td align="left">- For suspicion of local pancreatic complications</td>
</tr>
<tr>
<td align="left">- For suspected bowel ischemia</td>
</tr>
<tr>
<td align="left">- For acute bleeding (CTa) (if stable enough and consider embolization)</td>
</tr>
<tr>
<td align="left">- For abdominal compartment syndrome</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th align="left">6. Invasive intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td align="left">- For deteriorating patient with suspected infected local complication</td>
</tr>
<tr>
<td align="left">- “Step up approach” with initial drain guided by current CT scan (percutaneous or endoscopic drainage)</td>
</tr>
<tr>
<td align="left">- Delay for 3 to 4 weeks with intensive care support, if possible</td>
</tr>
<tr>
<td align="left">- If failure to respond or secondary deterioration, repeat CT scan, and select appropriate minimally invasive technique based on available expertise and equipment</td>
</tr>
<tr>
<td align="left">- Video-assisted retroperitoneal debridement or percutaneous nephroscopic debridement</td>
</tr>
<tr>
<td align="left">- Endoscopic transluminal debridement</td>
</tr>
<tr>
<td align="left">- Ongoing large bore drainage and irrigation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th align="left">7. Indication for laparotomy</th>
</tr>
</thead>
<tbody>
<tr>
<td align="left">- Failed “step-up approach” for further debridement/drainage</td>
</tr>
<tr>
<td align="left">- Acute abdomen (perforation or ischemia)</td>
</tr>
<tr>
<td align="left">- Severe abdominal compartment syndrome (rarely)</td>
</tr>
</tbody>
</table>
Table 33-11

Complications of acute pancreatitis

I. Local
A. Pancreatic phlegmon
B. Pancreatic abscess
C. Pancreatic pseudocyst
D. Pancreatic ascites
E. Involvement of adjacent organs, with hemorrhage, thrombosis, bowel infarction, obstructive jaundice, fistula formation, or mechanical obstruction

II. Systemic
A. Pulmonary
1. Pneumonia, atelectasis
2. Acute respiratory distress syndrome
3. Pleural effusion
B. Cardiovascular
1. Hypotension
2. Hypovolemia
3. Sudden death
4. Nonspecific ST-T wave changes
5. Pericardial effusion
C. Hematologic
1. Hemocoagulation
2. Disseminated intravascular coagulopathy
D. GI hemorrhage
1. Peptic ulcer
2. Erosive gastritis
3. Portal vein or splenic vein thrombosis with varices
E. Renal
1. Oliguria
2. Azotemia
3. Renal artery/vein thrombosis
F. Metabolic
1. Hyperglycemia
2. Hypocalcemia
3. Hypertriglyceridemia
4. Encephalopathy
5. Sudden blindness (Purtscher’s retinopathy)
G. Central nervous system
1. Psychosis
2. Fat emboli
3. Alcohol withdrawal syndrome
H. Fat necrosis
1. Intra-abdominal saponification
2. Subcutaneous tissue necrosis


CHRONIC PANCREATITIS

Definition, Incidence, and Prevalence
Chronic pancreatitis is an incurable, chronic inflammatory condition that is multifactorial in its etiology, highly variable in its presentation, and a challenge to treat successfully. Autopsy studies indicate that evidence of chronic inflammation, such as fibrosis, duct ectasia, and acinar atrophy is seen in up to 5% of the population, although these data are difficult to interpret because many of these changes are also present in asymptomatic elderly patients. Population studies suggest a prevalence that ranges from 5 to 40 persons per 100,000 population, with considerable geographic variation. Differences in diagnostic criteria, regional nutrition, alcohol consumption, and medical access account for variations in the frequency of the diagnosis, but the overall incidence of the disease has risen progressively over the past 50 years.

Etiology
There are multiple etiologies of chronic pancreatitis, including genetic mutations, alcohol exposure, duct obstruction due to trauma, gallstones, and tumors, metabolic diseases such as hyperlipidemia and hyperparathyroidism, and auto-immune disease. In addition, nutritional causes include so-called tropical pancreatitis, which has been thought to result from ingestion of certain starches. A significant number of patients have no discernible cause of the disease despite extensive testing, and are said to have “idiopathic” chronic pancreatitis.

Genetic Causes
In 1952, Comfort and Steinberg reported a kindred of “hereditary chronic relapsing pancreatitis” after treating the proband, a 24-year-old woman, at the Mayo Clinic. Subsequently, familial patterns of chronic, nonalcoholic pancreatitis have been described worldwide, and a familiar pattern has emerged. Typically, patients first present in childhood or adolescence with abdominal pain and are found to have chronic calcific pancreatitis on imaging studies. Progressive pancreatic dysfunction is common, and many patients present with symptoms due to pancreatic duct obstruction. The risk of subsequent carcinoma formation is increased, reaching a prevalence, in some series, of 40%, but the age of onset for carcinoma is typically >50 years old. The disorder is characterized by an autosomal dominant pattern of inheritance, with 80% penetrance and variable expression. The incidence is equal in both sexes.

Whitcomb and colleagues, and separately LeBodic and associates, performed gene-linkage analysis and identified a linkage for hereditary pancreatitis to chromosome 7q35. Subsequently, the region was sequenced and revealed eight trypsinogen genes. Mutational analysis revealed a missense mutation resulting in an Arg-to-His substitution at position 122 of the anionic trypsinogen gene, or PRSS1, one of the primary sites for proteolysis of trypsin. This gain-of-function mutation results in an excess production of trypsinogen, which results in persistent and uncontrolled proteolytic activity and autodestruction within the pancreas. The position mutation of PRSS1 and an additional mutation, now known collectively as the R122H and N291 mutations of PRSS1, account for about two-thirds of cases of hereditary pancreatitis. Masson and associates described a gain-of-function mutation in the anionic trypsinogen gene, PRSS2, that is also present in some cases.

Similarly, SPINK1, an inflammation-induced trypsin inhibitor secreted in acinar cells, has been found to have a role in hereditary pancreatitis. SPINK1 specifically inhibits trypsin action by competitively blocking the active site of the enzyme. Witt and colleagues investigated unrelated children with chronic pancreatitis in Germany and found a variety of SPINK1 mutations in 23% of the patients. Several studies have now confirmed an association of loss-of-function SPINK1 mutations with familial and idiopathic forms of chronic pancreatitis, as well as so-called tropical pancreatitis. SPINK1 mutations
are common in the general population as well, and the frequency of these mutations varies in different cohorts of idiopathic chronic pancreatitis, from 6.4% in France to 25.8% in the United States. Thus, hereditary pancreatitis results from one or more mutational defects that incapacitate an auto-protective process that normally prevents proteolysis within the pancreas.

Cystic fibrosis, originally termed cystic fibrosis of the pancreas, results from a variety of mutations of the cystic fibrosis transmembrane receptor (CFTR). The CFTR is present in pancreatic duct cells and controls the amount of chloride and bicarbonate secreted into the normally alkaline pancreatic juice. The CFTR gene contains over 4300 nucleotides, divided into 24 exons, which encode a 1480-amino acid protein. Over 1000 polymorphisms have been reported, and many are common. The CFTR mutation associated with the classic pulmonary disease, F508, is rarely observed in chronic pancreatitis. But other CFTR mutations have been noted to be associated with chronic idiopathic pancreatitis, auto-immune pancreatitis, and pancreas divisum, in which the pulmonary, intestinal, and cutaneous manifestations of the disease are silent.

Many studies have been undertaken to determine whether specific genetic abnormalities are associated with alcoholic chronic pancreatitis and which might confer susceptibility to the disease. In 2012, a landmark study by Whitcomb and associates demonstrated a likely genetic cause of the predisposition to alcohol-induced chronic pancreatitis in men. In a genome-wide association study of more than 2000 patients, these researchers discovered that a common DNA variant on the X chromosome is present in 26% of men without pancreatitis, but jumps to nearly 50% of men diagnosed with alcoholic pancreatitis. The variant involves the claudin 2 (CLDN2) gene, which encodes a tight junction protein normally present in ductal cells. In cases of chronic pancreatitis, the CLDN2 protein is abnormally expressed in acinar cells and may alter the secretory dynamics of enzyme release. The abnormality does not appear to cause pancreatitis, but if pancreatitis occurs for any reason in a person with the CLDN2 variant, it is more likely that the person will develop chronic pancreatitis; the risk is increased even further among alcohol users. Only 10% of women have the X chromosome–linked variant on both X chromosomes, and most women with the CLDN2 variant on one X chromosome appear to be protected from alcoholic chronic pancreatitis by the other X chromosome, if it is normal. Men, with only one X chromosome, have no protection if they inherit a CLDN2 mutation. This helps to explain the high prevalence of alcoholic chronic pancreatitis among men, although the mechanism remains unclear. This study does not demonstrate a genetic cause for all cases of alcohol-related chronic pancreatitis, but it shows that a genetic element contributes to many patients with the disease (Fig. 33-16).

Alcohol
In 1878, Friedreich proposed that “a general chronic interstitial pancreatitis may result from excessive alcoholism (drunkard’s pancreas).” Since that observation, numerous studies have shown that a causal relationship exists between alcohol and chronic pancreatitis, but the prevalence of alcohol as the etiology of the disease in Western countries ranges widely, from 38% to 94% (Fig. 33-17).

There is a linear relationship between exposure to alcohol and the development of chronic pancreatitis. The risk of disease is present in patients with even a low or occasional exposure to alcohol (1 to 20 g/d), perhaps due to the CLDN2 gene mutation described previously, so there is no threshold level of alcohol exposure below which there is no risk of developing chronic pancreatitis. Furthermore, although the risk of disease is dose related and highest in heavy (>150 g/d, or about 11 1 oz shots, or 12 beers per day) drinkers, the prevalence of chronic pancreatitis among confirmed alcohol abusers is only 5% to 15%. However, the duration of alcohol consumption is definitely associated with the development of pancreatic disease. The onset of disease typically occurs between ages 35 to 40 years, after 16 to 20 years of heavy alcohol consumption. Recurrent episodes of acute pancreatitis are typically followed by chronic symptoms after 4 or 5 years.

In their 1946 classic study, Comfort, Gambrill, and Baggenstoss proposed that chronic pancreatitis was the result of multiple episodes of acute inflammation, with residual and
Progressively increasing chronic inflammation.\textsuperscript{117} Subsequently, Kondo and associates showed that other, additional factors were necessary for repeated exposure to alcohol to cause chronic pancreatitis.\textsuperscript{118} Regardless of the requirement for other predisposing or facilitative factors, the concept that multiple episodes (or a prolonged course) of pancreatic injury ultimately leads to chronic disease is widely accepted as the pathophysiologic sequence\textsuperscript{119} (Fig. 33-18).

Although direct alcohol exposure to the pancreatic ductal system, or elevated levels of alcohol in the bloodstream, has been shown to alter the integrity and function of pancreatic ducts and acini directly,\textsuperscript{120} most investigators believe that alcohol metabolites such as acetaldehyde, combined with oxidant injury, result in local parenchymal injury that is preferentially targeted to the pancreas in predisposed individuals. Repeated or severe episodes of toxin-induced injury activate a cascade of cytokines, which, in turn, induces pancreatic stellate cells (PSCs) to produce collagen and cause fibrosis (Fig. 33-19).

Hyperlipidemia

In addition to the risk of acute pancreatitis, hyperlipidemia and hypertriglyceridemia predispose women to chronic pancreatitis when they receive estrogen replacement therapy.\textsuperscript{127} Fasting triglyceride levels <300 mg/dL are below the threshold for this to occur, and the mechanism of estrogen potentiation of hyperlipidemia-induced chronic pancreatitis is unknown. It is assumed that chronic changes occur after repeated subclinical episodes of acute inflammation. Aggressive therapy of hyperlipidemia is therefore important in peri- or postmenopausal patients who are candidates for estrogen therapy.

Classification

A major impediment to a better understanding of the etiology, frequency, and severity of chronic pancreatitis has been the difficulty with which investigators and clinicians have struggled to identify a useful classification system. Multiple classification systems have been proposed. The TIGAR-O scheme categorizes chronic pancreatitis according to risk factors and etiologies, such as (a) toxic-metabolic, (b) idiopathic, (c) genetic, (d) autoimmune, (e) recurrent and severe acute pancreatitis, or (f) obstructive.\textsuperscript{128} A recent classification system based on histopathology as well as etiology was delineated by Singer and Chari.\textsuperscript{129}

Chronic Calcific (Lithogenic) Pancreatitis

This type is the largest subgroup in the classification scheme proposed by Singer and Chari and includes patients with calcific pancreatitis of most etiologies. Although the majority of patients with calcific pancreatitis have a history of alcohol abuse, stone formation and parenchymal calcification can develop in a variety of etiologic subgroups; hereditary pancreatitis and tropical pancreatitis are particularly noteworthy for the formation of stone disease. The clinician should therefore avoid the assumption that calcific pancreatitis confirms the diagnosis of alcohol abuse.
Normal pancreas
Alcohol
Metabolic / oxidative stress
Sentinel event
Acinar cell injury
(+ / – Necrosis)
Inflammatory response:
proinflammatory (Early)
Anti-inflammatory (Later)

Chronic Obstructive Pancreatitis
This refers to chronic inflammatory changes that are caused by the compression or occlusion of the proximal ductal system by tumor, gallstone, posttraumatic scar, or inadequate duct caliber (as in pancreas divisum). Obstruction of the main pancreatic duct by inflammatory (posttraumatic) or neoplastic processes can result in diffuse fibrosis, dilated main and secondary pancreatic ducts, and acinar atrophy. The patient may have little in

NORMAL PANCREAS

Absence of duct of Santorini

Normal pancreas with duct of Santorini

Pancreas divisum

Small duct of Wirsung

Pancreas divisum
No duct of Wirsung

Figure 33-20. Pancreas divisum. Normal pancreatic duct anatomy and the variations of partial or complete pancreas divisum are shown. (Reproduced with permission from Beger HG: The pancreas: an integrated textbook of basic science, medicine, and surgery. London: Blackwell-Science; 1998.)

Chronic Inflammatory Pancreatitis

Chronic inflammatory pancreatitis is characterized by diffuse fibrosis and a loss of acinar elements with a predominant mononuclear cell infiltration throughout the gland.

A variant of chronic pancreatitis is a nonobstructive, diffusely infiltrative disease associated with fibrosis, a mononuclear cell (lymphocyte, plasma cell, or eosinophil) infiltrate, and an increased titer of one or more autoantibodies. This type, referred to as autoimmune pancreatitis (AIP), is associated with a variety of illnesses with suspected or proven autoimmune etiology, such as Sjögren’s syndrome, rheumatoid arthritis, and type 1 diabetes mellitus. AIP has been characterized as either type I, with accompanying systemic or multiorgan dysfunction, or type II, which is restricted to the pancreas.

Compressive stenosis of the intrapancreatic portion of the common bile duct is frequently seen in both types of AIP, along with symptoms of obstructive jaundice. Increased levels of serum β-globulin or immunoglobulin G4 are also present. Steroid therapy is uniformly successful in ameliorating the disease, including any associated bile duct compression. CFTR mutations that result in dislocation of the transmembrane protein have been found in AIP, and steroid therapy results in a normalization of the CFTR localization and a resumption of normal chloride and bicarbonate secretion. The differential diagnosis includes lymphoma, plasmacytoma (“pseudotumor” of the pancreas) and diffuse infiltrative carcinoma. Although the diagnosis is confirmed on pancreatic biopsy, presumptive treatment with steroids is usually undertaken, especially when clinical and laboratory findings, such as an elevation in IgG4 levels, support the diagnosis. Failure to obtain a cytologic specimen may lead to an unnecessary resectional procedure, and an untreated inflammatory component may cause sclerosis of the extrahepatic or intrahepatic bile ducts, with eventual liver failure.
**Tropical (Nutritional) Pancreatitis**

Chronic pancreatitis is highly prevalent among adolescents and young adults in Indonesia, southern India, and tropical Africa. Abdominal pain develops in adolescence, followed by the development of a brittle form of pancreateogenic diabetes. Parenchymal and intraductal calcifications are seen, and the pancreatic duct stones may be quite large. Many of the patients appear malnourished, some present with extreme emaciation, and a characteristic cyanotic coloration of the lips may be seen. In addition to protein-calorie malnutrition, toxic products of some indigenous foodstuffs have also been thought to contribute to the disease. Because of the geographic concentration of this early-onset form of chronic pancreatitis, it has been termed tropical pancreatitis, although the exact etiology remains unclear.

Clinically, tropical pancreatitis presents much like hereditary pancreatitis, and a familial pattern among cases is not unusual. *SPINK1* mutations have been documented in 20% to 55% of patients with tropical pancreatitis, and *CFTR* mutations have been reported as well. The accelerated deterioration of endocrine and exocrine function, the chronic pain due to obstructive disease, and the recurrence of symptoms despite decompressive procedures characterize the course of disease. As immigrants from the tropical regions increasingly find their way to all parts of the world, an awareness of this severe form of chronic pancreatitis is helpful for those who treat patients with pancreatic disease.

**Asymptomatic Pancreatic Fibrosis**

Pancreatic fibrosis is seen in some asymptomatic elderly patients, in tropical populations, or in asymptomatic alcohol users. There is diffuse perilobular fibrosis and a loss of acinar cell mass, but there is not a main ductular component. In addition, the presence of fibrosis and decreased exocrine function in patients with diabetes has raised the question of whether long-standing diabetes is a cause of chronic pancreatitis. Patients with this entity are usually asymptomatic in terms of typical pancreatic pain, and a recent histopathologic study of patients with typical chronic pancreatitis and “diabetic exocrine pancreatopathy” reveals significant differences in morphology, including a virtual absence of duct distortion or obstruction (Fig. 33-21).

It remains unknown whether this form of chronic inflammation precedes or contributes to the roughly twofold increase in the risk of pancreatic cancer in patients with long-standing diabetes.

**Idiopathic Pancreatitis**

When a definable cause for chronic pancreatitis is lacking, the term idiopathic is used to categorize the illness. Classically, the idiopathic group includes young adults and adolescents who lack a family history of pancreatitis but who may represent individuals with spontaneous gene mutations encoding regulatory proteins in the pancreas. A variable percentage of *SPINK1* and *CFTR* mutations have been described in various studies. In addition, the idiopathic group has included a large number of older patients for whom no obvious cause of recurrent or chronic pancreatitis can be found. However, because the prevalence of biliary calculi increases steadily with age, it is not surprising that, as methods of biliary stone detection have improved, many elderly “idiopathic” pancreatitis patients are found to have biliary tract disease.

An increasing number of mutations of the *SPINK1* and *CFTR* genes have been identified in association with various forms of chronic pancreatitis. However, the role of genetic analysis in the management of these patients remains unclear, as guidelines have yet to be developed to allow physicians to use the data consistently. Although the clinical management of patients who harbor a minor *CFTR* mutation and chronic pancreatitis, for example, is still dictated by the clinical manifestations of the pancreatitis, recent data suggest that the etiology of chronic pancreatitis, rather than the morphology, may determine the response to surgical treatment.

A shortcoming of these clinical classification systems is the lack of histologic criteria of chronic inflammation due to the usual absence of a biopsy specimen. The differentiation of recurrent acute pancreatitis from chronic pancreatitis with exacerbations of pain can be difficult to establish and is not facilitated by the current system. Similarly, cystic fibrosis is known to cause fibrosis and acinar dysfunction but is not included in the classification despite increasing evidence for its possible role in idiopathic chronic pancreatitis. Therefore, further refinements in the classification system for chronic pancreatitis...
are needed to allow a better prediction of its clinical course and a more accurate diagnosis of a likely etiologic agent.

**Pathology**

**Histology.** In early chronic pancreatitis, the histologic changes are unevenly distributed and are characterized by induration, nodular scarring, and lobular regions of fibrosis (Fig. 33-22). As the disease progresses, there is a loss of normal lobulation, with thicker sheets of fibrosis surrounding a reduced acinar cell mass and dilatation of ductular structures (Fig. 33-23). The ductular epithelium is usually atypical and may display features of dysplasia, as evidenced by cuboidal cells with hyperplastic features, accompanied by areas of mononuclear cell infiltrates or patchy areas of necrosis. Cystic changes may be seen, but areas of relatively intact acinar elements and normal-appearing islets persist. In severe chronic pancreatitis, there is considerable replacement of acinar tissue by broad, coalescing areas of fibrosis, and the islet size and number are reduced (Fig. 33-24). Small arteries appear thickened, and neural trunks become prominent.¹⁴⁷

Tropical pancreatitis and hereditary pancreatitis are histologically indistinguishable from chronic alcoholic pancreatitis. In obstructive chronic pancreatitis, calculi are absent, although periacinar fibrosis and dilated ductular structures are prominent. In pancreatic lobular fibrosis seen in elderly subjects, small ducts are dilated, sometimes with small calculi trapped within. Hypertrophy of ductular epithelia is thought to cause this small-duct disease, which is accompanied by perilobular fibrosis.¹⁴⁸

**Fibrosis.** A common feature of all forms of chronic pancreatitis is the perilobular fibrosis that forms surrounding individual acini, then propagates to surround small lobules, and eventually coalesces to replace larger areas of acinar tissue. The pathogenesis of this process involves the activation of pancreatic stellate cells (PSCs) that are found adjacent to acini and small arteries.¹⁴⁹ The extended cytoplasmic processes of PSCs encircle the acini but appear quiescent in the normal gland, where they contain lipid vacuoles and cytoskeletal proteins. In response to pancreatic injury, the PSCs become activated and proliferate (similarly to hepatic stellate cells), lose their lipid vesicles, and transform into myofibroblast-like cells. These cells respond to proliferative factors such as transforming growth factor β, platelet-derived growth factor, and proinflammatory cytokines that synthesize and secrete type I and III collagen and fibronectin. Studies indicate that vitamin A metabolites, similar to those present in quiescent PSCs, can inhibit the collagen production of activated cultured PSCs.¹⁵⁰ This raises the possibility that early intervention may be possible to interrupt or prevent the fibrosis resulting from ongoing activation of PSCs.

The overall pathogenic sequence proposed by Schneider and Whitcomb¹⁵¹ whereby alcohol induces acute pancreatitis and, with ongoing exposure, promotes the development of chronic fibrosis, is summarized in Fig. 33-19. PSCs surrounding the acinus are activated in acute pancreatitis but may be inactivated by anti-inflammatory cytokines and, in the absence of further injury, may revert to a quiescent state. The role of proinflammatory macrophages, cytokines, and PSCs in models of acute and chronic pancreatitis represents an important area of current research.

**Stone Formation.** Pancreatic stones are composed largely of calcium carbonate crystals trapped in a matrix of fibrillar and other material. The fibrillar center of most stones contains no
calcium but rather a mixture of other metals. This suggests that stones form from an initial noncalcified protein precipitate, which serves as a focus for layered calcium carbonate precipitation. The same low molecular weight protein is present in stones and protein plugs and was initially named pancreatic stone protein, or PSP.\textsuperscript{152} PSP was found to be a potent inhibitor of calcium carbonate crystal growth and has subsequently been renamed lithostathine.\textsuperscript{153} Independently, a 15-kDa fibrillar protein isolated from the pancreas was named pancreatic thread protein, and it has been shown to be homologous with lithostathine. Finally, a protein product of the \textit{reg} gene, so named because it is expressed in association with regenerating islets in models of pancreatic injury, was isolated and called \textit{reg} protein and was subsequently found to be homologous with lithostathine.\textsuperscript{154} The PSP/pancreatic thread protein/\textit{reg}/lithostathine gene encodes for a 166-amino acid product that undergoes posttranslational modification to produce isoforms present in pancreatic juice. The protein is expressed in all rodents and mammals, both in the pancreas as well as in brain tissue, where it is found in particularly high concentrations in pyramidal neurons in Alzheimer’s disease and Down syndrome. It is also found in the renal tubules, which is consistent with its biologic action of preventing calcium carbonate precipitation.

Calcium and bicarbonate ions are normally present in pancreatic juice in high concentrations, and the solubility product of calcium carbonate is greatly exceeded under normal conditions. Microcrystals of calcium carbonate can be seen in normal pancreatic juice but are usually clinically silent. Lithostathine is a potent inhibitor of calcium carbonate crystal formation, at a concentration of only 0.1 \(\mu\text{mol/L}\). However, lithostathine concentrations in normal pancreatic juice are in the range of 20 to 25 \(\mu\text{mol/L}\), so a constant suppression of calcium carbonate crystal formation is present in the normal pancreas.

In alcoholics and in patients with alcoholic chronic pancreatitis, lithostathine expression and secretion are dramatically inhibited\textsuperscript{155} (Fig. 33-25). In addition, elevated levels of precipitated lithostathine in the duct fluid in chronic pancreatitis patients suggests that the availability of the protein may be further reduced by the action of increased proteases and other proteins present in the duct fluid of alcoholic patients. Increased pancreatic juice protein levels in alcoholic men are reversible by abstinence from alcohol,\textsuperscript{156} so the availability and effectiveness of lithostathine may be restored in patients with early-stage disease by timely intervention. Nevertheless, calcific stone formation represents an advanced stage of disease, which can further promote injury or symptoms due to mechanical damage to duct epithelium or obstruction of the ductular network.

**Duct Distortion.** Although calcific stone disease is normally a marker for an advanced stage of disease, parenchymal and ductular calcifications do not always correlate with symptoms. Obstructing main duct stones are commonly observed and are thought to be an indication for endoscopic or surgical removal. The ball-valve effect of a stone in a secreting system produces inevitable episodes of duct obstruction, usually accompanied by pain. But some patients with complete duct obstruction have prolonged periods of painlessness. Ductular hypertension has been documented in patients with proximal stenosis of the main pancreatic duct, and prolonged ductular distention after secretin administration is taken as a sign of ductular obstruction.\textsuperscript{157} Although calculus disease and duct enlargement appear together as late stages of chronic pancreatitis, controversy persists over whether they are associated, are independent events, or are causally related.

**Radiology.** Radiologic imaging of chronic pancreatitis assists in four areas: (a) diagnosis, (b) the evaluation of severity of disease, (c) detection of complications, and (d) assistance in determining treatment options.\textsuperscript{158} With the advent of cross-sectional imaging techniques such as CT and MRI, the contour, content, ductal pattern, calcifications, calculi, and cystic disease of the pancreas are all readily discernible. Transabdominal ultrasonography is frequently used as a screening method for patients with abdominal symptoms or trauma, and the extension of ultrasonic imaging to include endoscopic ultrasound (EUS) and laparoscopic US have resulted in the highest-resolution images that are capable of detecting very small (<1 cm) abnormalities in the pancreas. EUS is now frequently used as a preliminary step in the evaluation of patients with pancreatic disease, and magnetic resonance cholangiopancreatography (MRCP) is increasingly being used to select patients who are candidates for the most invasive imaging method, ERCP. The staging of disease is important in the care of patients, and a combination of imaging methods is usually used (Table 33-12).

Ultrasonography is frequently used as an initial imaging method in patients with abdominal symptoms, and changes consistent with pancreatic duct dilatation, intraductal filling defects, cystic changes, and a heterogeneous texture are seen in chronic pancreatitis (Fig. 33-26). The sensitivity of transabdominal ultrasonography ranges from 48\% to 96\%, and it is operator dependent.\textsuperscript{159} However, the contour, texture, and ductal pattern are usually quite discernible, and it is a reliable method for periodic reexamination to determine the efficacy of treatment.

EUS has heavily impacted the evaluation and management of patients with chronic pancreatitis. Although it is more operator dependent than transabdominal ultrasonography, EUS provides not only imaging capability but also adds the capacity to obtain cytologic and chemical samples of tissue and fluid aspirated with linear array monitoring (Fig. 33-27). EUS images obtained through a high-frequency (7.5- to 12.5-mHz) transducer are able to evaluate subtle changes in 2- to 3-mm

![Figure 33-25. Lithostathine levels in chronic calcific pancreatitis (CCP) patients, patients with alcohol abuse (Alc.), patients with other pancreatic disease (OPD), and controls. (Reproduced with permission from Beger HG: The pancreas: an integrated textbook of basic science, medicine, and surgery. London: Blackwell-Science; 1998.)](image-url)
Table 33-12
Cambridge classification of pancreatic morphology in chronic pancreatitis

<table>
<thead>
<tr>
<th>CLASSIFICATION</th>
<th>ERCP FINDINGS</th>
<th>CT AND US FINDINGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>No abnormal SBDs</td>
<td>Normal gland size, shape; homogeneous parenchyma</td>
</tr>
<tr>
<td>Equivocal</td>
<td>MPD normal</td>
<td>One of the following: less than three abnormal SBDs; MPD 2–4 mm; gland enlarged more than two times normal size; heterogeneous parenchyma</td>
</tr>
<tr>
<td>Mild</td>
<td>MPD normal</td>
<td>Two or more of the following: less than three abnormal SBDs; MPD 2–4 mm; slight gland enlargement; heterogeneous parenchyma</td>
</tr>
<tr>
<td>Moderate</td>
<td>MPD changes</td>
<td>Small cysts &lt;10 mm; MPD irregularity</td>
</tr>
<tr>
<td></td>
<td>SBD changes</td>
<td>Focal acute pancreatitis; increased echogenicity of MPD walls; gland-contour irregularity</td>
</tr>
<tr>
<td>Severe</td>
<td>Any of the above changes plus one or more of the following: cysts &lt;10 mm; intraductal filling defects; calculi; MPD obstruction or stricture; severe MPD irregularity; contiguous organ invasion</td>
<td>—</td>
</tr>
</tbody>
</table>

Abbreviations: CT = computed tomography; ERCP = endoscopic retrograde cholangiopancreatography; MPD = main pancreatic duct; SBD = side-branch duct; US = ultrasound.


Structures within the pancreas and can detect indolent neoplasms in the setting of chronic inflammation. Small intraductal lesions, intraductal mucous cystic lesions, and sub ductular abnormalities are recognizable by EUS (Table 33-13). This allows ERCP to be reserved for these patients who require therapeutic maneuvers, or for the evaluation of more complex problems. EUS is comparable to ERCP in the detection of advanced changes in chronic pancreatitis and may be more sensitive than ERCP in the detection of mild disease.160

CT scanning has affected the diagnosis of pancreatic disease more broadly than any other method. With the advent of faster helical CT scanning and CT angiography, visualization of the nature, extent, location, and relative relationships of pancreatic structures and lesions is possible with great clarity. Duct dilatation, calculous disease, cystic changes, inflammatory events, and anomalies are all detectable with a resolution of 3 to 4 mm (Fig. 33-28). CT scanning has a false-negative rate of <10% for chronic pancreatitis, but early or mild chronic disease may go undetected by CT imaging. The earliest changes are dilatation of secondary ducts and heterogeneous parenchymal changes, which are detectable by EUS and ERCP. Another drawback of CT scanning is its lower sensitivity for detecting small neoplasms, which are seen with increased frequency in chronic pancreatitis and may be invisible to all modalities except EUS.

An MRI, in both the cross-sectional mode and the coronally oriented heavily weighted T2 or high spin ratio imaging structures within the pancreas and can detect indolent neoplasms in the setting of chronic inflammation. Small intraductal lesions, intraductal mucous cystic lesions, and sub ductular abnormalities are recognizable by EUS (Table 33-13). This allows ERCP to be reserved for these patients who require therapeutic maneuvers, or for the evaluation of more complex problems. EUS is comparable to ERCP in the detection of advanced changes in chronic pancreatitis and may be more sensitive than ERCP in the detection of mild disease.160


Figure 33-27. Endoscopic ultrasound of chronic pancreatitis. The endoscopic ultrasound appearance of the parenchyma is heterogeneous, and dilated ducts are seen, indicating early obstructive pancreatitis. (Used with permission from Mark Topazian, Division of Digestive Diseases, Department of Medicine, Mayo Clinic.)
Table 33-13

Endoscopic ultrasound features of chronic pancreatitis

<table>
<thead>
<tr>
<th>ENDOSCOPIC ULTRASOUND FEATURE</th>
<th>IMPLICATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ductal changes</td>
<td></td>
</tr>
<tr>
<td>Duct size &gt;3 mm</td>
<td>Ductal dilation</td>
</tr>
<tr>
<td>Tortuous pancreatic duct</td>
<td>Ductal irregularity</td>
</tr>
<tr>
<td>Intraductal echogenic foci</td>
<td>Stones or calcification</td>
</tr>
<tr>
<td>Echogenic duct wall</td>
<td>Ductal fibrosis</td>
</tr>
<tr>
<td>Side-branch ectasia</td>
<td>Periductal fibrosis</td>
</tr>
<tr>
<td>Parenchymal changes</td>
<td></td>
</tr>
<tr>
<td>Inhomogeneous echo pattern</td>
<td>Edema</td>
</tr>
<tr>
<td>Reduced echogenic foci (1–3 mm)</td>
<td>Edema</td>
</tr>
<tr>
<td>Enhanced echogenic foci</td>
<td>Calcifications</td>
</tr>
<tr>
<td>Prominent interlobular septae</td>
<td>Fibrosis</td>
</tr>
<tr>
<td>Lobular outer gland margin</td>
<td>Fibrosis, glandular atrophy</td>
</tr>
<tr>
<td>Large, echo-poor cavities (&gt;5 mm)</td>
<td>Pseudocyst</td>
</tr>
</tbody>
</table>


(MRCP) that can disclose fluid-filled ducts and cystic lesions, has added greatly to the imaging options for chronic pancreatitis (Fig. 33-29). The resolution of cross-sectional MRI scanning is now approaching that of CT scanning, although the availability of MRI scanners and the complexity of the images produced have limited their large-scale use for routine imaging of the pancreas. MRCP has been shown to be an effective screening technique for disclosing ductal abnormalities that correlate closely with the contrast-filled ducts imaged by ERCP. The advantages of MRCP include its noninvasive methodology and its ability to image obstructed ducts that are not opacified by ERCP injection. It is therefore a useful screening study to detect duct abnormalities and to confirm the need for interventional procedures. Oral, IV, and intraductal contrast are unnecessary for MRCP, and its lack of ionizing radiation makes this the safest method to image the ductal system in high-risk patients.

For the diagnosis and staging of chronic pancreatitis, ERCP is considered to be the gold standard. It also serves as a vehicle that enables other diagnostic and therapeutic maneuvers, such as biopsy or brushing for cytology, or the use of stents to relieve obstruction or drain a pseudocyst (Fig. 33-30).
Unfortunately, ERCP also carries a risk of procedure-induced pancreatitis that occurs in approximately 5% of patients.\textsuperscript{161} Patients at increased risk include those with sphincter of Oddi dysfunction and those with a previous history of post-ERCP pancreatitis. Post-ERCP pancreatitis occurs after uncomplicated procedures, as well as after those that require prolonged manipulation. Severe pancreatitis and deaths have occurred after ERCP. It should be reserved for patients in whom the diagnosis is unclear despite the use of other imaging methods, or in whom a diagnostic or therapeutic maneuver is specifically indicated.

**Presentation, Natural History, and Complications**

**Presenting Signs and Symptoms.** Pain is the most common symptom of chronic pancreatitis. It is usually midepigastric in location but may localize or involve either the left or right upper quadrant of the abdomen. Occasionally, it is perceived in the lower midabdomen but is frequently described as penetrating through to the back (Fig. 33-31). The pain is typically steady and boring, but not colicky. It persists for hours or days and may be chronic with exacerbations caused by eating or drinking alcohol. Chronic alcoholics also describe a steady, constant pain that is temporarily relieved by alcohol, followed by a more severe recurrence hours later.

Patients with chronic pancreatic pain typically flex their abdomen and either sit or lie with their hips flexed, or lie on their side in a fetal position. Unlike ureteral stone pain or biliary colic, the pain causes the patient to be still. Nausea or vomiting may accompany the pain, but anorexia is the most common associated symptom.

Pain from chronic pancreatitis has been ascribed to multiple etiologies. Ductal hypertension, due to strictures or stones, may predispose to pain that is initiated or exacerbated by eating. Chronic pain without exacerbation may be related to parenchymal disease or retroperitoneal inflammation with persistent neural involvement. Acute exacerbations of pain in the setting of chronic pain may be due to acute increases in duct pressure or recurrent episodes of acute inflammation in the setting of chronic parenchymal disease. Nealon and Matin have described these various pain syndromes as being predictive of the response to various surgical procedures.\textsuperscript{162} Pain that is found in association with ductal hypertension is most readily relieved by pancreatic duct decompression, through endoscopic stenting or surgical decompression.

The surgical relief of pain due to obstructive pancreatopathy may be dependent on the degree of underlying fibrosis and the etiology of the disease rather than the presence of ductal obstruction, per se, according to a recent studies from Johns Hopkins. Cooper et al studied 35 patients with chronic pain associated with evidence of duct obstruction who were treated with local resection of the pancreatic head and longitudinal pancreaticojejunostomy (LR-LPJ), or Frey procedure.\textsuperscript{163} The degree of pain resolution after surgery was compared to the degree of underlying parenchymal fibrosis. After a follow-up that averaged 22 months, patients with more than 80% fibrosis had 100% pain relief, whereas only 60% patients with less than 10% fibrosis experienced substantial or complete pain relief. Subsequently, this group studied 60 patients who had undergone either the Frey procedure or the Whipple procedure for refractory pain due to chronic pancreatitis.\textsuperscript{164} In addition to histopathologic findings, they also analyzed the etiology of the disease. Of patients with “toxic” etiologies of acquired disease (i.e., a history of alcohol or tobacco abuse), 89% experienced prolonged pain relief, whereas only 39% of those with hereditary or idiopathic disease achieved this result. Further, these results were independent of the degree of pancreatic fibrosis. These findings suggest that the etiology of the disease may be the most important predictor of the benefit of a resectional or hybrid procedure, and that patients with idiopathic or hereditary disease might be considered for an alternative approach, such as total pancreatectomy with islet autotransplantation (see following section).

The pain of chronic pancreatitis may decrease or disappear completely over a period of years, as symptoms of exocrine and endocrine deficiency become apparent.\textsuperscript{164} This is referred to as burned out pancreatitis and correlates with the progression of disease from a mild or moderate stage to severe destruction of the pancreas. Although this evolution of painful to nonpainful disease is sometimes used as a justification to avoid intervention in painful chronic pancreatitis, noninterventional approaches to the treatment of chronic pancreatitis are inevitably accompanied by the development of narcotic addiction, inability to work, and the sequelae of chronic illness.

Although increased ductal pressure, and therefore parenchymal pressure, has been thought to be the cause of pain in chronic obstructive pancreatitis, the role of chronic inflammation per se, and the development of actual nerve damage in the diseased gland, are also thought to contribute to pain.\textsuperscript{165} Chronic inflammation results in the infiltration of tissue by macrophages, which secrete prostaglandins and other nociceptive agents that cause chronic stimulation of afferent neural fibers. Inflammatory damage to the perineural layers surrounding the unmyelinated pancreatic nerves and a focal infiltration of inflammatory cells around nerves suggest that neural fibers are a target for the cellular response to inflammation in the pancreas.\textsuperscript{166}

Strategies to relieve pain are therefore based on three approaches: (a) reducing secretion and/or decompress the secretory compartment, (b) resecting the focus of chronic inflammatory change, or (c) interrupting the transmission of afferent neural impulses through neural ablative procedures. A trial of antisecretory therapy or endoscopic duct drainage may select those patients who will benefit preferentially from a decompressive procedure.

---

**Figure 33-31.** Pain location in chronic pancreatitis. (Reproduced with permission from Greenfield LJ, Mulholland MW, Oldham KT, et al: Surgery, Scientific Principles and Practice, 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2001.)
Patients with chronic pain due to chronic pancreatitis are often treated for years with escalating doses of narcotics, which imposes the additional problem of opioid dependency on this population. Further, it is now apparent that the pain pattern of many patients may be described as being either “visceral pain,” caused by inflammation in and around the pancreas, or “central sensitization of pain,” which is a syndrome of increased pain perception and sensitivity caused by the prolonged presence of pain\(^\text{167}\) (Fig. 33-32). Recent studies have suggested a method to differentiate these pain processes and further suggest that drugs that decrease or ameliorate the central sensitization of pain, such as pregabalin and tramadol, may be especially useful in chronic pancreatitis patients.\(^\text{168}\) These methods of differential diagnosis of pain syndromes now appear to offer the means by which it is possible to further identify and predict which patients will benefit from a surgical approach to chronic pancreatitis.\(^\text{169}\)

**Malabsorption and Weight Loss.** When pancreatic exocrine capacity falls below 10% of normal, diarrhea and steatorrhea develop\(^\text{170}\) (Fig. 33-33). Patients describe a bulky, foul-smelling, loose (but not watery) stool that may be pale in color and float on the surface of toilet water. Frequently, patients will describe a greasy or oily appearance to the stool, or may describe an “oil slick” on the water’s surface. In severe steatorrhea, an orange,
oily stool is often reported. As exocrine deficiency increases, symptoms of steatorrhea are often accompanied by weight loss. Patients may describe a good appetite despite weight loss or diminished food intake due to abdominal pain.

In severe symptomatic chronic pancreatitis, anorexia or nausea may occur with or separate from abdominal pain. The combination of decreased food intake and malabsorption of nutrients usually results in chronic weight loss. As a result, many patients with severe chronic pancreatitis are below ideal body weight.

Lipase deficiency tends to manifest itself before trypsin deficiency, so the presence of steatorrhea may be the first functional sign of pancreatic insufficiency. As pancreatic exocrine function deteriorates further, the secretion of bicarbonate into the duodenum is reduced, which causes duodenal acidification and further impairs nutrient absorption. Pancreatic exocrine insufficiency is frequently asymptomatic, however, and pancreatic exocrine function is difficult to measure, so a diagnosis of chronic pancreatitis is sufficient to justify a trial of pancreatic enzyme supplements. Each meal should be followed by 90,000 United States Pharmacopeia units of lipase, and the metabolic and symptomatic status of the patients should be followed.

Pancreatogenic Diabetes. The islets comprise only 2% of the mass of the pancreas, but they are preferentially conserved when pancreatic inflammation occurs. In chronic pancreatitis, acinar tissue loss and replacement by fibrosis is greater than the degree of loss of islet tissue. Islets are typically smaller than normal and may be isolated from their surrounding vascular network by the fibrosis. With progressive destruction of the gland, endocrine insufficiency commonly occurs. Frank diabetes is seen initially in about 20% of patients with chronic pancreatitis, and impaired glucose metabolism can be detected in up to 70% of patients. In a study of 500 patients with predominantly alcoholic chronic pancreatitis, diabetes developed in 83% within 25 years of the clinical onset of chronic pancreatitis, and more than half of the diabetic patients required insulin treatment. Ketoacidosis and diabetic nephropathy are relatively uncommon in pancreatogenic diabetes (see Table 33-3), but retinopathy and neuropathy are seen to occur with a similar frequency as in type 1 and type 2 diabetes.

Pancreatogenic diabetes is most common in cases of chronic pancreatitis, and it is often seen after surgical resection for benign or malignant disease (Fig. 33-34). Distal pancreatectomy and Whipple procedures have a higher incidence of diabetes than do drainage procedures, and the severity of diabetes is usually worse after subtotal or total pancreatectomy. Pancreatogenic, or type 3c diabetes (T3cDM), is seen in cystic fibrosis, in association with pancreatic cancer, and in cases of severe hemochromatosis.

The etiology and pathophysiology of pancreatogenic diabetes is distinct from that of either autoimmune (type 1) or obesity-related (type 2) diabetes. In type 3c diabetes, the loss of functioning pancreatic tissue by disease or surgical removal results in a global deficiency of all three glucoregulatory islet cell hormones: insulin, glucagon, and PP. In addition, there is a paradoxical combination of enhanced peripheral sensitivity to insulin and decreased hepatic sensitivity to insulin. As a result, insulin therapy is frequently difficult; patients are hyperglycemic when insulin replacement is insufficient (due to unsuppressed hepatic glucose production) or hypoglycemic when insulin replacement is barely excessive (due to enhanced peripheral insulin sensitivity and a deficiency of pancreatic glucagon secretion to counteract the hypoglycemia). This form of diabetes is referred to as brittle diabetes and requires special attention.

PP deficiency correlates with the severity of chronic pancreatitis, and impairments in the hepatic action of insulin are reversed in PP-deficient chronic pancreatitis patients by administration of PP. In addition, a study of type 1 and type 3c diabetic patients treated with insulin pump therapy revealed that the addition of a continuous subcutaneous infusion of PP reduced the insulin requirements needed for glycemic control. Studies are currently underway to identify a clinically suitable PP analog or PP receptor agonist.

Laboratory Studies. The diagnosis of chronic pancreatitis depends on the clinical presentation, a limited number of indirect measurements that correlate with pancreatic function, and selected imaging studies (Table 33-14). The direct measurement of pancreatic enzymes (e.g., lipase and amylase) by blood test is highly sensitive and fairly specific in acute pancreatitis but is seldom helpful in the diagnosis of chronic pancreatitis. The pancreatic endocrine product that correlates most strongly with chronic pancreatitis is the PP response to a test meal (Fig. 33-35). Severe chronic pancreatitis is associated with a blunted or absent PP response to feeding but, as with many other tests, a normal PP response does not rule out the presence of early disease.

---

**Figure 33-34.** Distribution of types of diabetes (A) and causes of type 3c (pancreatogenic) diabetes (B) based on studies of 1922 diabetic patients referred to an academic medical center as reported by Hardt et al. (Data from Hardt PD, Brendel MD, Kloor HU et al: Is pancreatic diabetes (type 3c diabetes) underdiagnosed and misdiagnosed? Diabetes Care. 2008 Feb;31 Suppl 2:S165-S169.)
Table 33-14

Tests for chronic pancreatitis

I. Measurement of pancreatic products in blood
   A. Enzymes
   B. Pancreatic polypeptide

II. Measurement of pancreatic exocrine secretion
   A. Direct measurements
      1. Enzymes
      2. Bicarbonate
   B. Indirect measurement
      1. Bentiromide test
      2. Schilling test
      3. Fecal fat, chymotrypsin, or elastase concentration
      4. [14C]-olein absorption

III. Imaging techniques
   A. Plain film radiography of abdomen
   B. Ultrasonography
   C. Computed tomography
   D. Endoscopic retrograde cholangiopancreatography
   E. Magnetic resonance cholangiopancreatography
   F. Endoscopic ultrasonography

The measurement of pancreatic exocrine secretion requires aspiration of pancreatic juice from the duodenum after nutrient (Lundh test meal) or hormonal (CCK or secretin) stimulation. Direct aspiration of pancreatic juice by endoscopic cannulation of the duct is performed in some centers, but it is not risk free, comfortable for the patient, or more sensitive than luminal intubation methods.

Indirect tests of pancreatic exocrine function are based on the measurement of metabolites of compounds that are altered (“digested”) by pancreatic exocrine products and can be quantified by serum or urine measurements. A commonly used indirect test is the bentiromide test, in which pancreatic polypeptide (PP) response to a test meal.

Figure 33-35. Pancreatic polypeptide (PP) response to a test meal. Immunoreactive PP (IR-PP) responses in control subjects (NL, n = 6) and patients with severe chronic pancreatitis (CP) accompanied by PP deficiency (CP, n = 5) are shown. A test meal was administered at 0 minutes. Means ± standard error of the mean are shown. (Reproduced with permission from Brunicardi FC, Chaiken RL, Ryan AS, et al. Pancreatic polypeptide administration improves abnormal glucose metabolism in patients with chronic pancreatitis, J Clin Endocrinol Metab. 1996 Oct;81(10):3566-3572.)

Table 33-15

Cambridge classification of chronic pancreatitis by endoscopic retrograde cholangiopancreatography

<table>
<thead>
<tr>
<th>GRADE</th>
<th>MAIN PANCREATIC DUCT</th>
<th>SIDE BRANCHES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>Suggestive</td>
<td>Normal</td>
<td>&lt;3 Abnormal</td>
</tr>
<tr>
<td>Mild</td>
<td>Normal</td>
<td>≥3 Abnormal</td>
</tr>
<tr>
<td>Moderate</td>
<td>Abnormal</td>
<td>&gt;3 Abnormal</td>
</tr>
<tr>
<td>Severe</td>
<td>Abnormal plus at least one of the following: large cavity, duct obstruction, dilation or duct irregularity, intraductal filing defects</td>
<td></td>
</tr>
</tbody>
</table>

EUS findings may be inconclusive in mild or “minimal change” pancreatitis, however, and improved criteria for an EUS-based diagnosis are still a work in progress. Most importantly, EUS is highly reliable in ruling out pancreatic carcinoma when CT findings are normal or equivocal.

Prognosis and Natural History. The prognosis for patients with chronic pancreatitis is dependent on the etiology of the disease, the development of complications, and on the age and socioeconomic status of the patient. The influence of treatment is less evident in long-term studies, although the general absence of randomized, prospective trials clouds the issue of whether specific forms of therapy alter the long-term outlook for patients with the disease.

Several studies have demonstrated that, although symptoms of pain decrease over time in about half of the patients, this decline is also accompanied by a progression of exocrine and endocrine insufficiency. In general, the likelihood of eventual pain relief is dependent upon the stage of disease at diagnosis, and the persistence of alcohol use in patients with alcoholic chronic pancreatitis. Miyake and colleagues found that pain relief was achieved in 60% of alcoholic patients who successfully discontinued drinking, but in only 26% who did not.

The long-term survival of patients with chronic pancreatitis is less than for patients without pancreatitis. In an international multicenter study of >2000 patients, Lowenfels and colleagues found that the 10- and 20-year survival rates for patients with chronic pancreatitis were 70% and 45%, respectively, compared to 93% and 65% for patients without pancreatitis. The mortality risk was found to be 1.6-fold higher in patients who continued to abuse alcohol, compared to those who did not. Continued alcohol abuse has a similar effect on the response to surgical treatment (Fig. 33-36), and results in a twofold increase in mortality over a 10- to 14-year follow-up period.

In addition to progressive endocrine and exocrine dysfunction, and the risk of the specific complications outlined here and in Table 33-16, the other significant long-term risk for the patient with chronic pancreatitis is the development of pancreatic carcinoma. There is a progressive, cumulative increased risk of carcinoma development in patients with chronic pancreatitis, which continues throughout the subsequent lifetime of the patient (Fig. 33-37). The incidence of carcinoma in patients with chronic pancreatitis ranges from 1.5% to 6%, which is at least 10-fold greater than that of patients of similar age seen in a hospital setting. In patients with chronic pancreatitis accompanied by diabetes, the risk of carcinoma has been found to be increased 12- to 33-fold compared to healthy, comparably aged controls. In patients with advanced chronic pancreatitis referred for surgical therapy, indolent, undiagnosed carcinoma can be seen in as many as 10% of patients.

The development of carcinoma in the setting of chronic pancreatitis is no doubt related to the dysregulation of cellular proliferation and tissue repair processes in the setting of chronic inflammation, as is seen throughout the alimentary tract and elsewhere. In the setting of chronic pancreatitis, carcinoma development can be especially cryptic, and the diagnosis of early-stage tumors is particularly difficult. Awareness of this risk justifies close surveillance for cancer in patients with chronic pancreatitis. Periodic measurement of tumor markers such as CA19-9, and periodic imaging of the pancreas with CT scan and EUS seem logical in order to detect the development

<table>
<thead>
<tr>
<th>Complications of chronic pancreatitis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intrapancreatic complications</strong></td>
</tr>
<tr>
<td>Pseudocysts</td>
</tr>
<tr>
<td>Duodenal or gastric obstruction</td>
</tr>
<tr>
<td>Thrombosis of splenic vein</td>
</tr>
<tr>
<td>Abscess</td>
</tr>
<tr>
<td>Perforation</td>
</tr>
<tr>
<td>Erosion into visceral artery</td>
</tr>
<tr>
<td>Inflammatory mass in head of pancreas</td>
</tr>
<tr>
<td>Bile duct stenosis</td>
</tr>
<tr>
<td>Portal vein thrombosis</td>
</tr>
<tr>
<td>Duodenal obstruction</td>
</tr>
<tr>
<td>Duct strictures and/or stones</td>
</tr>
<tr>
<td>Ductal hypertension and dilatation</td>
</tr>
<tr>
<td>Pancreatic carcinoma</td>
</tr>
<tr>
<td><strong>Extrapancreatic complications</strong></td>
</tr>
<tr>
<td>Pancreatic duct leak with ascites or fistula</td>
</tr>
<tr>
<td>Pseudocyst extension beyond lesser sac into mediastinum, retroperitoneum, lateral pericolic spaces, pelvis, or adjacent viscera</td>
</tr>
</tbody>
</table>

![Figure 33-36](image-url)
of carcinoma in the patient with chronic pancreatitis, although no evidence exists to indicate that this alters the outcome of patients who develop pancreatic cancer. Surgical procedures, particularly drainage procedures performed for presumed chronic pancreatitis, should always include biopsy of the tissue to exclude the diagnosis of malignancy.

**Complications**

**Pseudocyst.** A chronic collection of pancreatic fluid surrounded by a nonepithelialized wall of granulation tissue and fibrosis is referred to as a pseudocyst. Pseudocysts occur in up to 10% of patients with acute pancreatitis, and in 20% to 38% of patients with chronic pancreatitis, and thus, they comprise the most common complication of chronic pancreatitis. The identification and treatment of pseudocysts requires definition of the various forms of pancreatic fluid collections that occur (Table 33-17). In chronic pancreatitis, a pancreatic duct leak with extravasation of pancreatic juice results in a peripancreatic fluid collection (PPFC). Over a period of 3 to 4 weeks, the PPFC is sealed by an inflammatory reaction that leads to development of a wall of acute granulation tissue without much fibrosis. This is referred to as an acute pseudocyst. Acute pseudocysts may resolve spontaneously in up to 50% of cases, over a course of 6 weeks or longer. Pseudocysts >6 cm resolve less frequently than smaller ones but may regress over a period of weeks to months. Pseudocysts are multiple in 17% of patients, or they may be multiloculated. They may occur intrapancreatically or extend beyond the region of the pancreas into other cavities or compartments (Fig. 33-38).

Pseudocysts may become secondarily infected, in which case they become abscesses. They can compress or obstruct adjacent organs or structures, leading to superior mesenteric-portal vein thrombosis or splenic vein thrombosis. They can erode into visceral arteries and cause intracystic hemorrhage or pseudoaneurysms (Fig. 33-39). They also can perforate and cause peritonitis or intraperitoneal bleeding.

### Table 33-17

**Definitions of pancreatic fluid collections**

<table>
<thead>
<tr>
<th>TERM</th>
<th>DEFINITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peripancreatic fluid collection</td>
<td>A collection of enzyme-rich pancreatic juice that occurs early in the course of acute pancreatitis, or that forms after a pancreatic duct leak; located in or near the pancreas; it lacks a well-organized wall of granulation or fibrous tissue</td>
</tr>
<tr>
<td>Early pancreatic (sterile) necrosis</td>
<td>A focal or diffuse area of nonviable pancreatic parenchyma, typically occupying &gt;30% of the gland and containing liquefied debris and fluid</td>
</tr>
<tr>
<td>Late pancreatic (sterile) necrosis</td>
<td>An organized collection of sterile necrotic debris and fluid with a well-defined margin or wall within the normal domain of the pancreas</td>
</tr>
<tr>
<td>Acute pseudocyst</td>
<td>A collection of pancreatic juice enclosed within a perimeter of early granulation tissue, usually as a consequence of acute pancreatitis that has occurred within the preceding 3–4 wk</td>
</tr>
<tr>
<td>Chronic pseudocyst</td>
<td>A collection of pancreatic fluid surrounded by a wall of normal granulation and fibrous tissue, usually persisting for &gt;6 wk</td>
</tr>
<tr>
<td>Pancreatic abscess</td>
<td>Any of the above in which gross purulence (pus) is present, with bacterial or fungal organisms documented to be present</td>
</tr>
</tbody>
</table>

Pseudocysts usually cause symptoms of pain, fullness, or early satiety. Asymptomatic pseudocysts can be managed expectantly and may resolve spontaneously or persist without complication. Symptomatic or enlarging pseudocysts require treatment, and any presumed pseudocyst without a documented antecedent episode of acute pancreatitis requires investigation to determine the etiology.
A pseudocyst can erode into an adjacent artery, which results in contained hemorrhage otherwise known as a pseudoaneurysm. A contrast-injected computed tomographic scan reveals active bleeding (area marked B) into a pseudocyst (arrows) as a result of this process. (Reproduced with permission from Balthazar EJ: CT diagnosis and staging of acute pancreatitis, Radiol Clin North Am. 1989 Jan;27(1):19-37.)

The timing and method of treatment requires careful consideration. Pitfalls in the management of pseudocysts result from the incorrect (presumptive) diagnosis of a cystic neoplasm masquerading as a pseudocyst, a failure to appreciate the solid or debris-filled contents of a pseudocyst that appears to be fluid filled on CT scan, and a failure to document true adherence with an adjacent portion of the stomach before attempting transgastric internal drainage.

If the pseudocyst has failed to resolve with conservative therapy and symptoms persist, internal drainage is usually preferred to external drainage to avoid the complication of a pancreaticocutaneous fistula. Pseudocysts communicate with the pancreatic ductal system in up to 80% of cases, so external drainage creates a pathway for pancreatic duct leakage to and through the catheter exit site. Internal drainage may be performed with either endoscopic methods (transgastric or transduodenal puncture and multiple stent placements, with or without a nasocystic irrigation catheter), or surgical methods (a true cystoenterostomy, biopsy of cyst wall, and evacuation of all debris and contents). Surgical options include a cystogastrostomy (Fig. 33-40), a Roux-en-Y cystojejunostomy, or a cystoduodenostomy. Cystojejunostomy is the most versatile method, and it can be applied to pseudocysts that penetrate into the transverse mesocolon, the paracolic gutters, or the lesser sac. Cystogastrostomy can be performed endoscopically (Fig. 33-41), laparoscopically, or by a combined laparoscopic-endoscopic method.

Because pseudocysts often communicate with the pancreatic ductal system, two newer approaches to pseudocyst management are based on main duct drainage, rather than pseudocyst drainage per se. Transpapillary stents inserted at the time of the lesion, including a cystic neoplasm. Although pseudocysts comprise roughly two-thirds of all pancreatic cystic lesions, they resemble cystadenomas and cystadenocarcinoma radiographically. An incidentally discovered cystic lesion should be examined by EUS and aspirated to determine whether it is a true pancreatic cystic neoplasm or a pseudocyst.

Endoscopic ultrasound (EUS)-guided transgastric puncture of pancreatic pseudocyst. A. Transgastric stents placed across fused posterior wall of stomach and anterior wall of pseudocyst. (Reproduced with permission from Chauhan SS, Forsmark CE. Evidence-based treatment of pancreatic pseudocysts, Gastroenterology. 2013 Sep;145(3):511.)
large-bore catheters or multiple stents and an aggressive approach to management for success to be achieved. Failure of nonsurgical therapy, with subsequent salvage procedures to remove infected debris and establish complete drainage, is associated with increased risks for complications and death.\textsuperscript{211,212}

The most experienced therapeutic endoscopists report a complication rate of 17% to 19% for the treatment of sterile pseudocysts, and deaths as a result of endoscopic therapy have occurred.\textsuperscript{212} Therefore, the use of endoscopic methods to treat sterile or infected pancreatic necrosis has a higher complication rate and is limited to specialized centers.

Resection of a pseudocyst is sometimes indicated for cysts located in the pancreatic tail or when a midpancreatic duct disruption has resulted in a distally located pseudocyst. Distal pancreatectomy for removal of a pseudocyst, with or without splenectomy, can be a challenging procedure in the setting of prior pancreatitis. An internal drainage procedure of the communicating duct or of the pseudocyst itself should be considered when distal resection is being contemplated.

**Pancreatic Ascites.** When a disrupted pancreatic duct leads to pancreatic fluid extravasation that does not become sequestered as a pseudocyst, but drains freely into the peritoneal cavity, pancreatic ascites occurs. Occasionally, the pancreatic fluid tracks superiorly into the thorax, and a pancreatic pleural effusion occurs. Referred to as \textit{internal pancreatic fistulae}, both complications are seen more often in patients with chronic pancreatitis rather than after acute pancreatitis. Pancreatic ascites and pleural effusion occur together in 14% of patients, and 18% have a pancreatic pleural effusion alone.\textsuperscript{213}

Patients demonstrate the general demographics of chronic pancreatitis and usually present with a subacute or recent history of progressive abdominal swelling despite weight loss. Pain and nausea are rarely present. The abdominal CT scan discloses ascites and the presence of chronic pancreatitis or a partially collapsed pseudocyst (Fig. 33-43). Paracentesis or thoracentesis reveals noninfected fluid with a protein level >25 g/L and a markedly elevated amylase level. Serum amylase may also be elevated, presumably from reabsorption across the parietal membrane. Serum albumin may be low, and patients may have

---

\textbf{Figure 33-42.} Transpapillary drainage of a pancreatic pseudocyst. \textbf{A.} Endoscopic passage of a flexible wire through the major papilla, through the pancreatic duct, and into a communicating pseudocyst. \textbf{B.} Placement of a stent over the wire into the pseudocyst with transpapillary drainage. (Reproduced with permission from Kozarek RA, Brayko CM, Harlan J, et al. Endoscopic drainage of pancreatic pseudocysts, Gastrointest Endosc. 1985 Oct;31(5):322-327.)

\textbf{Figure 33-43.} Pancreatic ascites. Computed tomographic scan of a patient with a ruptured pancreatic pseudocyst resulting in intra-peritoneal pancreatic fluid. (Reproduced with permission from Cameron JL, Cameron AM: Current Surgical Therapy, 11th ed. Philadelphia, PA: Elsevier; 2014.)
coexisting liver disease. Paracentesis is therefore critical to differentiate pancreatic from hepatic ascites.

ERCP is most helpful to delineate the location of the pancreatic duct leak and to elucidate the underlying pancreatic ductal anatomy. Pancreatic duct stenting may be considered at the time of ERCP, but if nonsurgical therapy is undertaken and then abandoned, repeat imaging of the pancreatic duct is appropriate to guide surgical treatment.

Antisecretory therapy with the somatostatin analogue octreotide acetate, together with bowel rest and parenteral nutrition, is successful in more than half of patients. Reappraisal of serosal surfaces to facilitate closure of the leak is considered a part of therapy, and this is accomplished by complete paracentesis. For pleural effusions, a period of chest tube drainage may facilitate closure of the internal fistula. Surgical therapy is reserved for those who fail to respond to medical treatment. If the leak originates from the central region of the pancreas, a Roux-en-Y pancreaticojejunostomy is performed to the site of duct leakage (Fig. 33-44). If the leak is in the tail, a distal pancreatectomy may be considered, or an internal drainage procedure can be performed. The results of surgical treatment are usually favorable if the ductal anatomy has been carefully delineated preoperatively.

Pancreatic-Enteric Fistula. The erosion of a pancreatic pseudocyst into an adjacent hollow viscus can result in a pancreatic-enteric fistula. The most common site of communication is the transverse colon or splenic flexure. The fistula usually presents with evidence of GI or colonic bleeding and sepsis. If the fistula communicates with the stomach or duodenum, it may close spontaneously or persist as a pancreatic-enteric fistula. When the fistula involves the colon, operative correction is usually required.

**Figure 33-44.** Internal drainage for leaking pancreatic duct. A Roux-en-Y pancreaticojejunostomy is performed at the site of duct rupture to accomplish internal drainage of the pancreatic duct leak. (Reproduced with permission from Beger HG: The Pancreas. London: Blackwell-Science; 1998.)

### Table 33-18

<table>
<thead>
<tr>
<th>Signs and Symptoms</th>
<th>With Head Enlargement (n = 138) (%)</th>
<th>Without Head Enlargement (n = 141) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily severe pain</td>
<td>67</td>
<td>40</td>
</tr>
<tr>
<td>Cholelithiasis</td>
<td>46</td>
<td>11</td>
</tr>
<tr>
<td>Duodenal obstruction</td>
<td>30</td>
<td>7</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>18</td>
<td>30</td>
</tr>
<tr>
<td>Vascular involvement</td>
<td>15</td>
<td>8</td>
</tr>
</tbody>
</table>


**Head-of-pancreas Mass.** In up to 30% of patients with advanced chronic pancreatitis, an inflammatory mass develops in the head of the pancreas. The clinical presentation includes severe pain and frequently includes stenosis of the distal common bile duct, duodenal stenosis, compression of the portal vein, and stenosis of the proximal main pancreatic duct (Table 33-18). Mutations and polymorphisms of *p53* have been found in these patients, and a focus of ductular carcinoma was found in 3.7% of patients with pancreatic head enlargement in one series. It was concluded that an accelerated transformation from hyperplasia to dysplasia exists in patients with pancreatic head enlargement, although the etiology for this process remains unclear.

**Splenomegaly**. Vascular complications of chronic pancreatitis are fortunately infrequent because they are difficult to treat successfully. Portal vein compression and occlusion can occur as a consequence of an inflammatory mass in the head of the pancreas, and splenic vein thrombosis occurs in association with chronic pancreatitis in 1% to 8% of cases. Variceal formation can occur as a consequence of either portal or splenic venous occlusion, and splenic vein thrombosis with gastric variceal formation is referred to as left-sided or sinistral portal hypertension. Although bleeding complications are infrequent, the mortality risk of bleeding is >20%. When gastroesophageal varices are caused by splenic vein thrombosis, the addition of splenectomy to prevent variceal hemorrhage is prudent when surgery is otherwise indicated to correct other problems.

**Treatment**

**Medical Therapy.** The medical treatment of chronic or recurrent pain in chronic pancreatitis requires the use of analgesics, a cessation of alcohol use, oral enzyme therapy, and the selective use of antisecretory therapy. Interventional procedures to block visceral afferent nerve conduction or to treat obstructions of the main pancreatic duct are also an adjunct to medical treatment.

**Analgesia.** Oral analgesics are prescribed as needed, alone or with analgesia-enhancing agents such as gabapentin. Adequate pain control usually requires the use of narcotics, but these should be titrated to achieve pain relief with the lowest effective dose. Opioid addiction is common, and the use of long-acting analgesics by transdermal patch together with oral agents for pain exacerbations slightly reduces the sedative effects of high-dose oral narcotics.
It is essential for patients to abstain from alcohol. In addition to removing the causative agent, alcohol abstinence results in pain reduction or relief in 60% to 75% of patients with chronic pancreatitis. Despite this benefit, roughly half of alcoholic chronic pancreatitis patients continue to abuse alcohol.

**Enzyme Therapy.** Pancreatic enzyme administration serves to reverse the effects of pancreatic exocrine insufficiency. Adequate pancreatic enzyme replacement reverses the exocrine insufficiency seen in most patients, and it prevents secondary complications such as metabolic bone disease due to inadequate absorption of the fat-soluble vitamins A, D, E, and K. In addition, pancreatic enzyme replacement may reduce or alleviate the pain experienced by patients. The choice of enzyme supplement and the dose should be selected based on whether malabsorption or pain (or both) are the indications for therapy (Table 33-19). Conventional (nonenteric-coated) enzyme preparations are partially degraded by gastric acid but are available within the duodenal and jejunal regions to bind to CCK-releasing peptide and downregulate the release of CCK. This theoretically reduces the enteric signal for pancreatic exocrine secretion, which reduces the pressure within a partially or completely obstructed pancreatic duct. Enteric-coated preparations result in little to no pain relief, presumably due to their reduced bioavailability in the proximal gut. Due to the loss of pancreatic enzymes by acid hydrolysis and proteolysis, relatively large doses are required to achieve effective levels of enzyme within the proximal small bowel. Enteric-coated preparations are protected from acid degradation but are presumably not released in the critical proximal gut in sufficient quantity to inhibit the stimulus for endogenous pancreatic enzyme secretion. Nonalcoholic patients may experience more effective pain relief than alcoholic patients, but it is recommended that all patients with chronic pancreatitis pain begin a trial of nonenteric-coated enzyme supplements together with an acid-suppressive medication for 1 month. If pain relief is achieved, therapy is continued. If enzyme therapy fails, further investigation of the pancreatic ductal system by ERCP guides the therapy based on specific anatomical findings (Fig. 33-45).

**Antisecretory Therapy.** Somatostatin administration has been shown to inhibit pancreatic exocrine secretion and CCK release. The somatostatin analogue octreotide acetate has therefore been investigated for pain relief in patients with chronic pancreatitis. In a double-blind, prospective, randomized 4-week trial, 65% of patients who received 200 μg of octreotide acetate subcutaneously three times daily reported pain relief, compared with 35% of placebo-treated subjects. Patients who had the best results were patients with chronic abdominal pain, suggestive of obstructive pancreatopathy. However, in another trial that used a 3-day duration of treatment, no significant pain relief was observed. Anecdotal reports suggest that severe pain exacerbations in chronic pancreatitis can benefit from a combination of octreotide therapy and TPN, and a pilot study of the effectiveness of the sustained-release form of octreotide suggested that it was as effective as three-times-per-day administration of the short-acting form of the drug.

**Neurolytic Therapy.** Celiac plexus neurolysis with alcohol injection has been an effective form of analgesic treatment in patients with pancreatic carcinoma. However, the use of radiologically or endoscopically guided celiac plexus blockade in chronic pancreatitis has been disappointing. Due to the risk of alcohol injury and the need for repeated injections, celiac plexus blockade in chronic pancreatitis has used short-acting analgesics or other drugs rather than 50% alcohol. A trial of EUS-guided celiac plexus blockade revealed successful pain relief in 55% of patients, but the benefit lasted beyond 6 months in only 10% of patients. The procedure therefore appears safe, but the effect is short lived in those patients who obtain pain relief.

**Endoscopic Management.** The techniques of endoscopic treatment of pancreatic duct obstruction, stone disease, pseudocyst formation, pancreatic duct leak, and for the diagnosis and management of associated pancreatic tumors have expanded greatly over the past 20 years. Newer endoscopes with expanded therapeutic capabilities have been introduced, and the role of EUS-guided needle and catheter insertion has expanded the ability of the therapeutic endoscopist in the diagnosis and treatment of chronic pancreatitis and its complications.

Pancreatic duct stenting is used for treatment of proximal pancreatic duct stenosis, decompression of a pancreatic duct leak, and for drainage of pancreatic pseudocysts that can be catheterized through the main pancreatic duct. Pancreatic duct stents can induce an inflammatory response within the

<table>
<thead>
<tr>
<th>PRODUCT</th>
<th>FORMULATION</th>
<th>MANUFACTURER</th>
<th>LIPASE CONTENT (USP)/PILL OR CAPSULE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zenpep</td>
<td>Enteric-coated porcine</td>
<td>Aptalis</td>
<td>3000, 5000, 10,000, 15,000, 20,000</td>
</tr>
<tr>
<td>Creon</td>
<td>Enteric-coated porcine</td>
<td>Abbott</td>
<td>3000, 6000, 12,000, 24,000</td>
</tr>
<tr>
<td>Pancreaze</td>
<td>Enteric-coated porcine</td>
<td>Ortho-McNeil-Janssen</td>
<td>4200, 10,500, 16,800, 21,000</td>
</tr>
<tr>
<td>Pertzye</td>
<td>Enteric-coated porcine mixed with bicarbonate granules</td>
<td>Digestive Care</td>
<td>8000, 16,000</td>
</tr>
<tr>
<td>Ultrasa</td>
<td>Enteric-coated porcine</td>
<td>Aptalis</td>
<td>13,800, 20,700, 23,000</td>
</tr>
<tr>
<td>Viokace</td>
<td>Tablet non-enteric-coated porcine</td>
<td>Aptalis</td>
<td>10,440, 20,880</td>
</tr>
</tbody>
</table>

Make a correct diagnosis
• Appropriate history
• Corroborating imaging tests
  • MRI/MRCP
  • EUS
  • CT
• Functional tests if imaging tests equivocal
  • Tube-based secretin test
  • Endoscopic-based secretin test
• Assess for alternative diseases and complications and treat if present
  • Pancreatic cancer or IPMN
  • Pseudocyst
  • Bile duct obstruction
  • Duodenal obstruction

Medical therapy
• Measure pain severity, character, and impact on QOL
• Refer for formal structured smoking and alcohol cessation programs
• Counsel on good nutrition and initiate supplementation with vitamin D and calcium
  • Baseline bone mineral density testing
• Provide information on local and national support groups
• Initiate analgesics (starting with Tramadol)
  • Increase dose and potency slowly as required
• Initiate adjunctive agents in those with persistent pain or requiring higher dosages or potency of narcotics
  • Pregabalin, Gabapentin
  • SSRI
  • SSNRI
  • Tricyclic antidepressants
• Assess for evidence of coexistent exocrine or endocrine insufficiency and treat if present
  • Fecal elastase or serum trypsin
  • HbA1C or GTT
• Initiate steroids if autoimmune pancreatitis

Figure 33-45. Management algorithm for chronic pancreatitis. IPMN = intraductal papillary mucinous neoplasm; QOL = quality of life; SSRI = selective serotonin reuptake inhibitor; SSNRI = selective serotonin-norepinephrine reuptake inhibitor; GTT = glucose tolerance test. (Reproduced with permission from Forsmark CE. Management of chronic pancreatitis, Gastroenterology. 2013 Jun;144(6):1282-1291.)
duct, so prolonged stenting is usually avoided. Patients with sphincter of Oddi dyskinesia are at high risk for developing post-ERCP pancreatitis after biliary sphincterotomy, and the prophylactic placement of a pancreatic duct stent or the administration of rectal indomethacin reduces the amylase level and development of pancreatitis after biliary sphincterotomy.\textsuperscript{33,228} Pancreatic duct leaks are seen in 37\% of patients with acute pancreatitis, and pancreatic duct stenting appears to facilitate the resolution of the leak.\textsuperscript{229} Similarly, pancreatic duct stenting has been used to treat postsurgical pancreatic duct leaks and posttraumatic leaks.\textsuperscript{229-231}

Pancreas divisum (see Fig. 33-3) is thought to cause pain and chronic pancreatitis due to functional or mechanical obstruction of the dorsal duct draining exclusively, or predominantly, through the lesser papilla. A study from Marseille reported good long-term results in 24 patients after minor papilla sphincterotomy and dorsal duct stenting.\textsuperscript{232} The number of patients with chronic pain decreased from 83\% before stenting to 29\% after stenting, but pancreatitis or recurrent papillary stenosis occurred in 38\%. Patients that responded best were those with intermittent pain, and this subset may be preferentially treated with endoscopic therapy. Patients with recurrent pain and a dilated dorsal duct may be candidates for internal drainage of the obstructed duct with either an extended Puestow procedure or a Frey procedure (see later in this section).

Idiopathic pancreatitis patients have been treated with endoscopic stenting, pancreatic duct sphincterotomy, and endoscopic stone removal with good results. In a prospective randomized trial, 53\% of idiopathic recurrent pancreatitis patients in the control group experienced continued episodes of pancreatitis, although only 11\% of the treated patients had continued symptoms.\textsuperscript{233}

Extracorporeal shock wave lithotripsy (ESWL) has been used for pancreatic duct stones, together with endoscopic stenting and stone removal.\textsuperscript{234} A single ESWL session was used in 35 patients with pancreatic duct stones, together with 86 ERCP sessions to complete the stone removal process. After 2.4 years, 80\% of patients had significant relief of symptoms (Fig. 33-46). Also, endoscopic intraductal lithotripsy can now be performed in some specialized centers. However, due to the tendency for recurrent stone formation, the use of ESWL or endoscopic lithotripsy for long-term management of calcific pancreatitis remains uncertain.

**Surgical Therapy**

**Indications and History** The traditional approach to surgical treatment of chronic pancreatitis and its complications has maintained that surgery should be considered only when the medical therapy of symptoms has failed. Nealon and Thompson published a landmark study in 1993, however, that showed that the progression of chronic obstructive pancreatitis could be delayed or prevented by pancreatic duct decompression.\textsuperscript{235} No other therapy has been shown to prevent the progression of chronic pancreatitis, and this study demonstrated the role of surgery in the early management of the disease (Table 33-20). Small-duct disease or “minimal change chronic pancreatitis” are causes for uncertainty over the choice of operation, however. Major resections have a high complication rate, both early and late, in chronic alcoholic pancreatitis, and lesser procedures often result in symptomatic recurrence. Therefore, the choice of operation and the timing of surgery are based on each patient’s pancreatic anatomy, the likelihood (or lack thereof) that further

**Table 33-20**

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>24-Month Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operated (n = 47)</td>
<td>Mild to moderate 48 (87%); severe 6 (13%)</td>
</tr>
<tr>
<td>Nonoperated (n = 36)</td>
<td>Mild to moderate 8 (22%); severe 28 (78%)</td>
</tr>
</tbody>
</table>

Eighty-three patients with chronic pancreatitis were evaluated by exocrine, endocrine, nutritional, and endoscopic retrograde cholangiopancreatography studies, and all had mild to moderate disease and dilated pancreatic ducts. A Puestow-type duct decompression procedure was performed in 47 patients, and all subjects were restaged by the same methods 24 months later.

medical and endoscopic therapy will halt the symptoms of the disease, and the chance that a good result will be obtained with the lowest risk of morbidity and mortality. Finally, preparation for surgery should include restoration of protein-caloric homeostasis, abstinence from alcohol and tobacco, and a detailed review of the risks and likely outcomes to establish a bond of trust and commitment between the patient and the surgeon.

Historically, the surgery for chronic pancreatitis before the second half of the 20th century was a true demonstration of trial and error. Obtaining good surgical outcomes before the availability of CT scans and ERCP was either the result of serendipity or due to the skill and creativity of the surgeon. In 1911, Link described an operation he devised on the spot, when a laparotomy in a young woman with abdominal pain revealed a fluctuant, obstructed pancreatic duct. After performing a dochotomy and evacuating multiple stones, he inserted a rubber tube, and exteriorized the pancreatostomy just above her navel. He later described the operation as having been a success for the next 30 years of the patient’s life, during which the patient managed the care of the drainage tube without apparent problems.

With the demonstration in 1942 by Priestley that total pancreatectomy was technically feasible, and the report in 1946 by Whipple that proximal pancreatic resection was beneficial in (three) patients with chronic pancreatitis, the option of surgical resection as treatment for chronic pancreatitis was established. By the mid 1950s, however, growing disappointment with the high risk of resection and the lack of long-term benefit overshadowed the surgical treatment of chronic pancreatitis. The choice of resection vs. drainage was largely based on surgeon preference until the 1970s, when the widespread adoption of ERCP and CT scans provided the ability to preoperatively diagnose obstructive and sclerotic disease, and this resulted in the rational selection of operative procedures. During this period, the major drawbacks to surgical therapy remained the recurrence of symptoms despite surgery, the corresponding development of an inflammatory (or malignant) mass in the undrained pancreatic head (Fig. 33-47), or the high morbidity and mortality of major resectional procedures that predisposed patients to a cascade of metabolic problems.

**Sphincteroplasty** The sphincter of Oddi and the pancreatic duct sphincter serve as gatekeepers for the passage of pancreatic juice into the duodenum (Fig. 33-48). Stenosis of either sphincter (sclerosing papillitis), due to scarring from pancreatitis or from the passage of gallstones, may result in obstruction of the pancreatic duct and chronic pain. As gallstone pancreatitis became a popular diagnosis in the 1940s and 1950s, attention was focused on the ampullary region as a possible cause of chronic symptoms, and surgical sphincteroplasty was advocated. Although endoscopic techniques are now used routinely to perform sphincterotomy of either the common bile duct or pancreatic duct, a true (permanent)
sphincteroplasty can only be performed surgically. Transduodenal sphincteroplasty with incision of the septum between the pancreatic duct and common bile duct may offer significant relief for the rare patient with a focal obstruction and inflammation isolated to this region (Fig. 33-49).

**Drainage Procedures** After the early reports of success with pancreatostomy for the relief of symptoms of chronic pancreatitis, Cattell described pancreaticojejunostomy for relief of pain in unresectable pancreatic carcinoma. Shortly thereafter, Duval and, separately, Zollinger and associates described the caudal Roux-en-Y pancreaticojejunostomy for the treatment of chronic pancreatitis in 1954 (Fig. 33-50). The so-called Duval procedure was used for decades by some surgeons, but it almost invariably failed due to restenosis and segmental obstruction of the pancreas due to progressive scarring. In 1958, Puestow and Gillesby described these segmental narrowings and dilatations of the ductal system as a “chain of lakes,” and proposed a longitudinal decompression of the body and tail of the pancreas into a Roux limb of jejunum (Fig. 33-51). Four of Puestow and Gillesby’s 21 initial cases were side-to-side anastomoses, and 2 years after their report, Partington and Rochelle described a much simpler version of the longitudinal, or side-to-side Roux-en-Y pancreaticojejunostomy that became universally known as the Puestow procedure (Fig. 33-52).

Successful pain relief after the Puestow-type decompression procedure has been reported in 75% to 85% of patients for the first few years after surgery, but pain recurs in >20% of patients after 5 years due to progressive disease even in patients who are abstinent from alcohol.

With the advent of therapeutic endoscopy and techniques for transluminal stone removal and lithotripsy, multiple series have reported the successful endoscopic treatment of pancreatic duct calculi, although the long-term outcomes of these efforts has been uneven. Endoscopic removal of pancreatic duct stones is usually coupled to prolonged pancreatic duct stenting,
Figure 33-52. Longitudinal dochotomy in obstructing calcific pancreatitis. A longitudinal pancreatotomy typically discloses segmental stenosis of the pancreatic duct and the presence of intraductal calculi in a patient with chronic calcific pancreatitis (A). Following mobilization of a Roux limb of jejunum, a longitudinal pancreaticojejunostomy is performed to permit extensive drainage of the pancreatic duct system (B). This technique, described by Partington and Rochelle, is the typical method used for the Puestow procedure.

which carries the risk of further inflammation.252,253 Despite the risk of perioperative complications, the surgical management of pancreatic duct stones and stenosis has been shown to be superior to endoscopic treatment in randomized clinical trials in which the long, side-to-side technique of pancreaticojejunostomy is used.254-256

Resectional Procedures

Distal Pancreatectomy For patients with focal inflammatory changes localized to the body and tail, or in whom no significant ductal dilatation exists, the technique of partial (40–80%) distal pancreatectomy has been advocated (Fig. 33-53). Although distal pancreatectomy is less morbid than more extensive resectional procedures, the operation leaves untreated a major portion of the gland, and is therefore associated with a significant risk of symptomatic recurrence. It has been a more popular operation in British centers, where its success seems to be greater, perhaps due to the lower incidence of alcoholic chronic pancreatitis.257 However, long-term outcomes reveal good pain relief in only 60% of patients, with completion pancreatectomy required for pain relief in 13% of patients.

Laparoscopic distal pancreatectomy has been shown to be feasible for the removal of focal lesions of the distal pancreas,258 but it is more difficult in the setting of chronic pancreatitis.

Ninety-Five Percent Distal Pancreatectomy In 1965, Fry and Child proposed the more radical 95% distal pancreatectomy, which was intended for patients with sclerotic (small duct) disease and which attempted to avoid the morbidity of total pancreatectomy by preserving the rim of pancreas in the pancreaticoduodenal groove, along with its associated blood vessels and distal common bile duct.241 The operation was found to be associated with pain relief in 60% to 77% of patients long term, but it is accompanied by a high risk of brittle diabetes, hypoglycemic coma, and malnutrition. Although the operation was the first attempt to resect the pancreatic head while preserving the duodenum and distal bile duct, the extensive degree of metabolic complications led to its failure as viable treatment for the symptoms of pancreatic sclerosis.

Proximal Pancreatectomy In 1946, Whipple reported a series of five patients treated with either pancreaticoduodenectomy or total pancreatectomy for symptomatic chronic pancreatitis, with one operative death.240 Subsequently, proximal pancreatectomy or pancreaticoduodenectomy, with or without pylorus preservation (Fig. 33-54), has been widely used for the treatment of chronic pancreatitis.259 In the three largest modern (circa 2000) series of the treatment of chronic pancreatitis by the Whipple

Figure 33-53. Distal (spleen-sparing) pancreatectomy. A distal pancreatectomy for chronic pancreatitis is usually performed with en bloc splenectomy, using either an open or laparoscopic technique. In the presence of minimal inflammation, a spleen-sparing version can be performed, as shown here.
procedure, pain relief 4 to 6 years after operation was found in 71% to 89% of patients. However, mortality ranged from 1.5% to 3%, and major complications occurred in 25% to 38% of patients at the Johns Hopkins Hospital,260 the Mayo Clinic,261 and the Massachusetts General Hospital.262 In follow-up, 25% to 48% of patients developed diabetes, and about the same percentage required exocrine therapy. Advocates of the Whipple procedure as treatment for chronic pancreatitis suggest that the high rate of symptomatic relief outweighs the metabolic consequences and the mortality risk of the procedure, but increasingly this approach is being reserved for those patients with suspected occult malignancy.

Total Pancreatectomy Priestley and associates first described successful total pancreatectomy in 1944 in a patient with hyperinsulinism.239 and two of Whipple’s original five cases of chronic pancreatitis reported in 1946 were treated with total pancreatectomy.240 Subsequently, surgeons who used total pancreatectomy found that the operation produces no better pain relief for their patients than pancreaticoduodenectomy (about 80–85%). Moreover, the metabolic consequences of total pancreatectomy in the absence of islet cell transplantation can be profound and life-threatening. The patients have a “brittle” form of diabetes in which avoidance of hyper- and hypoglycemia is problematic.263 In addition, lethal episodes of hypoglycemia are common in severe apancreatic diabetes. These are due to hypoglycemic unresponsiveness, due to the absence of pancreatic glucagon, and to hypoglycemia unawareness, despite an ongoing need to treat with exogenous insulin.277 In a series of >100 patients treated with total pancreatectomy, Gall and colleagues showed that half of all the late deaths after this operation were due to (iatrogenic) hypoglycemia.264 Despite newer forms of insulin, insulin delivery systems, and continuous blood glucose monitoring systems, severe pancreateogenic diabetes remains an adverse outcome, as complete prevention of the physiologic consequences of total pancreatectomy remains an unfulfilled goal. Even with the growing acceptance of islet auto-transplantation as an adjunct to the procedure (see later in this section), total pancreatectomy itself is now used only rarely for the treatment of refractory chronic pancreatitis.

Hybrid Procedures In 1980, Beger and associates described the Duodenum-preserving Pancreatic Head Resection or DPPHR265 (Fig. 33-55), and they published long-term results with DPPHR for the treatment of chronic pancreatitis in 1985266 and again in 1999.267 In 388 patients who were followed for an average of 6 years after DPPHR, pain relief was reportedly maintained in 91%, mortality was <1%, and diabetes developed in 21%, with 11% demonstrating a reversal of their preoperative diabetic status. These authors also compared the DPPHR procedure with the pylorus-sparing Whipple procedure in a randomized trial of 40 patients with chronic pancreatitis.268 The mortality was reportedly zero in both groups, and the morbidity was also comparable. Pain relief (over 6 months) was seen in 94% of DPPHR patients, but in only 67% of Whipple patients. Furthermore, the insulin secretory capacity and glucose tolerance were noted to deteriorate in the Whipple group, but they actually improved in the DPPHR patients.

The DPPHR requires the careful dissection of the gastroduodenal artery and the creation of two anastomoses (Fig. 33-56), and it carries a similar complication risk as the Whipple procedure due to the risk of pancreatic leakage and intra-abdominal fluid collections.

In 1987, Frey and Smith described the local resection of the pancreatic head with longitudinal pancreaticojunostomy (LR-LPJ), which included excavation of the pancreatic head, including the ductal structures in continuity with a long ductotomy of the dorsal duct269 (Fig. 33-57). The Frey procedure provides thorough decompression of the pancreatic head as well as the body and tail of the gland, and a long-term follow-up suggested that improved outcomes are associated with this more extensive decompressive procedure. Frey and Amikura reported their results in 50 patients followed for >7 years, and they found complete or substantial pain relief in 87% of patients. There was no operative mortality, but 22% of patients developed postoperative complications.270

Figure 33-54. The pancreaticoduodenectomy (Whipple procedure) can be performed either with the standard technique, which includes distal gastrectomy (A), or with preservation of the pylorus (B). The pylorus-sparing version of the procedure is used most commonly. (Reproduced from Wu GY, Aziz K, Whalen GF: An Internist’s Illustrated Guide to Gastrointestinal Surgery. Totowa: Humana Press; 2003.)
Figure 33-55. The duodenum-preserving pancreatic head resection described by Beger and colleagues. A. The completed resection after transection of the pancreatic neck, and subtotal removal of the pancreatic head, with preservation of the distal common bile duct and duodenum. B. Completion of the reconstruction with anastomosis to the distal pancreas and to the proximal pancreatic rim by the same Roux limb of jejunum. (Reproduced with permission from Bell RH, Rikkers LF, Mulholland M: Digestive Tract Surgery: A Text and Atlas. Philadelphia, PA: Lippincott Williams & Wilkins; 1996.)

Key steps in the performance of the LR-LPJ include preservation of the pancreatic neck as well as the capsule of the posterior pancreatic head. In the pancreaticoduodenectomy and the DPPHR, the pancreatic neck is freed up from the portal and superior mesenteric vein confluence and divided. In the LR-LPJ, the neck of the pancreas is preserved intact as are the body and tail of the pancreas. Not having to divide the pancreatic neck, as in the pancreaticoduodenectomy or DPPHR, reduces the risk of the operation because it avoids intraoperative problems with the venous structures lying posterior to the gland. To reduce the risk of penetrating the posterior capsule of the head, Frey recommended in his 1994 report that the posterior limit of resection be the back wall of the opened duct of Wirsung and duct to the uncinate (Fig. 33-58).

Subsequent to Frey’s own modification of the technique, other surgeons have described modifications of the extent or technique of the LR-LPJ. Andersen and Topazian advocated performing the LR-LPJ as it was originally described, in which the entirety of the ducts are excised from the head (Fig. 33-59), and described the use of the ultrasonic aspirator and dissector for this purpose. This device permits precise removal of the ducts and adjacent tissue with good visualization and without complications. There is little pancreatic tissue behind these ducts, and the pancreatic capsule is continuously palpated as the dissection proceeds to ensure a safe margin of resection. The intrapancreatic portion of the common bile duct is usually exposed, and avoiding injury to it is enhanced by the ultrasonic aspirator. The majority of the parenchyma of the uncinate process is spared, and the excavation of the pancreatic head is made contiguous with a generous dochotomy of the dorsal duct. Whether merely unroofing as opposed to removal of the proximal ducts contributes to better pain relief is not known and awaits a randomized trial to compare the two versions of the LR-LPJ. Izbicki and colleagues at the University of Hamburg also recommend a more extensive excavation of the pancreatic head, and they use a technique that they refer to as the Hamburg modification of the LR-LPJ. This wider excavation of the pancreatic head is created in continuity with the dorsal dochotomy, and it is followed by a single, side-to-side pancreaticojejunostomy.

In 2001, Ho and Frey subsequently described merely excavating the core of the pancreatic head and draining the excava-

Figure 33-56. Intraoperative view of the Beger procedure. The gastroduodenal artery is encircled by a vessel loop. Just below, the intrapancreatic portion of the common bile duct is exposed as it courses toward the ampulla. A rim of well-vascularized pancreatic tissue remains in the duodenal C-loop. Preservation of the posterior branch of the gastroduodenal artery is essential to preserve viability of these structures.
Figure 33-57. Frey procedure. The local resection of the pancreatic head with longitudinal pancreaticojejunostomy (LR-LPJ) provides complete decompression of the entire pancreatic ductal system. Reconstruction is performed with a side-to-side Roux-en-Y pancreaticojejunostomy. (Reproduced with permission from Bell RH, Rikkers LF, Mulholland M: Digestive Tract Surgery: A Text and Atlas. Philadelphia, PA: Lippincott Williams & Wilkins; 1996.)

Figure 33-58. Operative view of excavated head of the pancreas during the Frey procedure. The main pancreatic duct is opened widely down to the level of the ampulla, and the head of the pancreas is excavated in a conical fashion so as to allow complete decompression of the chronically obstructed and inflamed pancreatic ducts. (Reproduced with permission from Aspelund G et al. Improved outcomes for benign disease with limited pancreatic head resection, J Gastrointest Surg. 2005 Mar;9(3):400-409.)

Figure 33-59. Complete excavation of the pancreatic head and distal pancreatic dochotomy. A true excavation and removal of the proximal ductal system is combined with a distal pancreatic dochotomy. Reconstruction is performed with a single side-to-side Roux-en-Y pancreaticojejunostomy. (Reproduced with permission from Andersen DK, Topazian MD. Pancreatic head excavation: a variation on the theme of duodenum-preserving pancreatic head resection, Arch Surg. 2004 Apr;139(4):375-379.)
Farkas and colleagues described a similar excavation of the central portion of the pancreatic head without any effort to include the duct of the body in the lateral pancreaticojejunostomy, and they reported excellent results with what they termed an organ-preserving pancreatic head resection (OPPHR) in a randomized comparison to the pylorus-preserving pancreaticoduodenectomy (PPPD).

This approach was advocated by Gloor and associates in Bern as an alternative to the DPPHR procedure in patients with portal hypertension and was described as the Berne modification of the DPPHR (Fig. 33-62). Köninger and colleagues in Heidelberg subsequently published a randomized, controlled trial of the “Berne” version of the excavation method compared to the “classic” Beger procedure. Operative times and length of stay were shorter in the group undergoing excavation of the pancreatic head, while long-term outcomes and quality-of-life scores were identical over 2 years postoperatively.

The common element of these variations on the theme of LR-LPJ remains the excavation or “coring out” of the central portion of the pancreatic head. It remains uncertain, however, whether and to what degree the dochotomy needs to be extended into the body and tail. The logical conclusion of all of these efforts is that the head of the pancreas is the nidus of the chronic inflammatory process in chronic pancreatitis and that removal of the central portion of the head of the gland is the key to the successful resolution of pain in the long term.

Complications Initial and long-term results of the LR-LPJ demonstrate pain relief that is equivalent to that of pancreaticoduodenectomy and the DPPHR. The observed mortality rate has been virtually zero, and therefore, less than with the Whipple procedure. Major complications were less with the LR-LPJ (16%) than with pancreaticoduodenectomy (40%) or DPPHR (25%) in one single-site series, and the incidence of new postoperative diabetes after LR-LPJ was 8% with an average follow-up of 3 years.

Comparisons of the Three Operative Procedures: Pancreaticoduodenectomy (Whipple procedure), DPPHR (Beger procedure), and LR-LPJ (Frey procedure). There has been considerable interest to apply evidence-based methods to the study of the three operations currently advocated for the treatment of chronic pancreatitis. The best studies, or level 1 data by the Strength of Recommendation Taxonomy, are prospective, randomized controlled trials comparing two or more operations from a single or multi-institutional study. Retrospective, cohort-based studies are regarded as level 2 data by the Strength of Recommendation Taxonomy criteria.

To date, ten published level 1 studies and three level 2 studies have examined various comparisons between these three operations. In the level 1 study of Klempp and colleagues and that of Buchler et al, DPPHR patients had a shorter hospital stay, greater weight gain, less postoperative diabetes, and exocrine dysfunction than standard Whipple patients over a 3- to 5-year follow-up. Pain control was similar between the two procedures. Similar results were observed in a
of new diabetes (8%) for both DPPHR and LR-LPJ compared to
the Whipple procedure (25%), but no significant differences in
outcomes or pain relief between DPPHR and LR-LPJ.\textsuperscript{199} Finally,
level 2 data support the efficacy of both DPPHR and LR-LPJ in
patients with dilated as well as nondilated ducts.\textsuperscript{286,289,289}

Long-term exocrine and/or endocrine insufficiency in
chronic pancreatitis patients treated surgically is a product of the
surgical intervention as well as the progression of the underlying
disease. Although the short-term (3-year) incidence of new diabe-
tes after operation appears less with the LR-LPJ and DPPHR than
with the PPPD, the late incidence of diabetes appears similar in all
groups. After an average of 7 years of follow-up after LR-LPJ or
PPPd, survival, pain relief, and pancreatic function were similar
in both groups. The rate of diabetes was slightly lower after LR-
LPJ (61\%) than after PPPD (65\%), but these had both more than
doubled from their preoperative status.\textsuperscript{275-281} Therefore, although
the limited pancreatic procedures of DPPHR and LR-LPJ have
a lower initial rate of endocrine dysfunction, the long-term risk
of diabetes is more related to the progression of the underlying
disease than to the effects of operation.

The level 1 studies confirm that the duodenum preserving
options are associated with a lower immediate morbidity and
mortality and therefore, in the absence of a mass or concerns
about cancer, are better options than a Whipple procedure for
chronic pancreatitis. The choice of LR-LPJ, DPPHR, or OPPHR
depends largely on surgeon experience, and the LR-LPJ is most
common in the United States.

**Total Pancreatectomy With Islet Auto-Transplantation** Islet
cell transplantation for the treatment of diabetes is an attractive
adjunct to pancreatic surgery in the treatment of benign pancre-
atic disease. Despite the difficulties in recovering islets from a
chronically inflamed gland, Najarian and associates demonstrated
the utility of autotransplantation of islets in patients with chronic
pancreatitis in 1980.\textsuperscript{281} Subsequently, through refinements in
the methods of harvesting and gland preservation, and through
standardization of the methods by which islets are infused into
the portal venous circuit for intrahepatic engraftment, the suc-
cess of total pancreatectomy combined with islet autotransplant-
lation has steadily increased to achieve insulin independence in
the majority of patients treated in recent series.\textsuperscript{282,283} Although
2 to 3 million islets are required for successful engraftment in
an allogeneic recipient, the auto-transplant recipient can usually
achieve long-term, insulin-independent status after engraftment
of only 300,000 to 400,000 islets (about one-third to one-half of
the number of islets in the normal pancreas).\textsuperscript{284}

The ability to recover a sufficient quantity of islets from
a sclerotic gland is dependent on the degree of fibrotic dis-
ease present, so the selection of patients as candidates for
autologous islet transplantation is important. The impressive
improvement in quality of life measures and pain relief seen
after total pancreatectomy with islet auto-transplantation (TP-
IAT) indicate that it is a highly successful form of therapy for
some patients (Fig. 33-63).\textsuperscript{295} The outcomes of TP-IAT are sig-
nificantly better in pediatric patients than in adults, largely due
to the prevalence of hereditary and idiopathic causes of chronic
pancreatitis in pediatric patients.\textsuperscript{296} These studies suggest that
further definition is needed regarding criteria for considering
TP-IAT vs. hybrid or resectional procedures for patients with
persistent symptoms. With the emerging evidence that hybrid
procedures seem to offer better outcomes in patients with toxic
etiologies of chronic pancreatitis, and that TP-IAT appears to

---

**Figure 33-62.** The Berne modification of the local resection of the
pancreatic head with longitudinal pancreaticojejunostomy.
preferentially benefit patients with hereditary and idiopathic forms of the disease, the importance of the careful definition of the etiology of chronic pancreatitis, including genetic testing, is steadily increasing.

**PANCREATIC NEOPLASMS**

**Neoplasms of the Endocrine Pancreas**

Neoplasms of the endocrine pancreas are relatively uncommon but do occur with enough frequency (five cases per million population) that most surgeons will encounter them in an urban practice. The cells of the endocrine pancreas, or islet cells, originate from neural crest cells, also referred to as amine precursor uptake and decarboxylation cells. Multiple endocrine neoplasia (MEN) syndromes occur when these cells cause tumors in multiple sites. The MEN1 syndrome involves pituitary tumors, parathyroid hyperplasia, and pancreatic neoplasms. Although most pancreatic endocrine tumors are nonfunctional, some are functional, secreting peptide products that produce interesting clinical presentations. Neoplasms of the endocrine pancreas that are not associated with excess hormone levels and a recognizable clinical syndrome are considered nonfunctional. Special immunohistochemical stains allow pathologists to confirm the peptide products being produced within the cells of a pancreatic endocrine tumor. However, the histologic characteristics of these neoplasms do not predict their clinical behavior, and malignancy is usually determined by the presence of local invasion and lymph node or hepatic metastases. Unfortunately, most pancreatic endocrine tumors are malignant, but the course of the disease is far more favorable than that seen with pancreatic exocrine cancer. The key to diagnosing these rare tumors is recognition of the classic clinical syndrome; confirmation is achieved by measuring serum levels of the elevated hormone. Localization of the tumor can be a challenging step, but once accomplished, the surgery is relatively straightforward. The goals of surgery range from complete resection, often accomplished with insulinomas, to controlling symptoms with debulking procedures. Unresectable disease in the liver is often addressed with chemoembolization.

As with pancreatic exocrine tumors, the initial diagnostic imaging test of choice for pancreatic endocrine tumors is a multidetector CT scan with four phases of contrast and fine cuts through the pancreas and liver. Neuroendocrine tumors of the pancreas often enhance with contrast. EUS can be superior to CT in localizing these tumors, which can produce dramatic symptoms despite their small (<1 cm) size. In contrast to pancreatic exocrine tumors, many of the endocrine tumors have somatostatin receptors (SSTRs) that allow them to be detected by a radiolabeled octreotide scan. A radioactive somatostatin analogue is injected intravenously, followed by whole-body radionuclide scanning (Fig. 33-64). The success of this modality in localizing tumors and detecting metastases has decreased the use of older techniques such as angiography and selective venous sampling.

**Insulinoma**

Insulinomas are the most common functional pancreatic endocrine neoplasms and present with a typical clinical
syndrome known as Whipple’s triad. The triad consists of symptomatic fasting hypoglycemia, a documented serum glucose level <50 mg/dL, and relief of symptoms with the administration of glucose. Patients can present with a profound syncopal episode or less severe symptoms that are averted by frequent eating. Common symptoms include palpitations, trembling, diaphoresis, confusion or obtundation, and seizure, and family members may report that the patient has undergone personality change.

Routine laboratory studies will uncover a low blood sugar, the cause of all of these symptoms. Serum insulin levels are elevated. C-peptide levels should also be elevated and rule out the unusual case of surreptitious administration of insulin or oral hypoglycemic agents because excess endogenous insulin production leads to excess C-peptide. The diagnosis can be clinched with a monitored fast in which blood is sampled every 4 to 6 hours for glucose and insulin levels until the patient becomes symptomatic. However, this can be dangerous and must be done with close supervision.

Insulinomas are usually localized with CT scanning and EUS. Technical advances in EUS have led to preoperative identification of >90% of insulinomas. Visceral angiography with venous sampling is rarely required to accurately localize the tumor. Insulinomas are evenly distributed throughout the head, body, and tail of the pancreas. Unlike most endocrine pancreatic tumors, the majority (90%) of insulinomas are benign and solitary, and only 10% are malignant. They are typically cured by simple enucleation. However, tumors located close to the main pancreatic duct and large (>2 cm) tumors may require a distal pancreatectomy or pancreaticoduodenectomy. Intraoperative US is useful to determine the tumor’s relation to the main pancreatic duct and guides intraoperative decision making. Approximately 90% of insulinomas are sporadic, and 10% are associated with the MEN1 syndrome. Insulinomas associated with the MEN1 syndrome are more likely to be multifocal and have a higher rate of recurrence.

**Noninsulinoma Hyperinsulinemia Hypoglycemia Syndrome**

A syndrome of noninsulinoma pancreaticogenous hypoglycemia was described by Service et al in 1999. The syndrome is associated with beta-cell hypertrophy, islet hyperplasia and increased beta-cell mass. When these findings are accompanied by ectopic islet tissue, multilobulated islets, and ductuloinsular complexes, the definition of nesideoblastosis is met. Nesideoblastosis accompanied by hyperinsulinism was previously considered a disease of neonates, where subtotal or total pancreatectomy was required to correct potentially fatal neonatal hyperinsulinism. However, dozens of cases of nesideoblastosis associated with hyperinsulinism have now been reported in patients 2 to 5 years after Roux-en-Y gastric bypass for obesity. Many of these patients have undergone partial or total pancreatectomy to prevent potentially fatal hypoglycemia. The illness in former bariatric surgery patients appears to result from an idiosyncratically-prolonged hypersecretion of the incretin hormones GIP and GLP-1 after the gastric bypass. GLP-1 is a potent stimulant of the expression of the transcription factor PDX-1, which normally regulates beta-cell development and growth. The correct treatment of this condition to prevent episodes of hypoglycemia is conversion of the gastric bypass to a form of bariatric procedure that restores normal intestinal flow of nutrients, such as the gastric sleeve, or the addition of a restriction element such as an adjustable gastric band. Pancreatic resection without conversion of the Roux-en-Y gastric bypass is not appropriate because this allows the abnormal enteroinsular relationship to continue and hyperinsulinemia persists or recurs after partial pancreatectomy.

**Gastrinoma**

Zollinger-Ellison syndrome (ZES) is caused by a gastrinoma, an endocrine tumor that secretes gastrin, leading to acid hypersecretion and peptic ulceration. Many patients with ZES present with abdominal pain, peptic ulcer disease, and severe esophagitis. However, in the era of effective antacid therapy, the presentation can be less dramatic. Although most of the ulcers are solitary, multiple ulcers in atypical locations that fail to respond to antacids should raise suspicion for ZES and prompt a work-up. At the time of diagnosis, 21% of patients with gastrinoma have diarrhea.

The diagnosis of ZES is made by measuring the serum gastrin level. It is important that patients stop taking proton pump inhibitors for this test. In most patients with gastrinomas, the level is >1000 pg/mL. Gastrin levels can be elevated under conditions other than ZES. Common causes of hypergastrinemia include pernicious anemia, treatment with proton pump inhibitors, renal failure, G-cell hyperplasia, atrophic gastritis, retained or excluded antrum, and gastric outlet obstruction. In equivocal cases, when the gastrin level is not markedly elevated, a secretin stimulation test is helpful.

**Figure 33-64.** Radioactive octreotide scan demonstrating pancreatic endocrine tumor in the body of the pancreas (arrow).
SPECIFIC CONSIDERATIONS

PART II

In 70% to 90% of patients, the primary gastrinoma is found in Passaro’s triangle, an area defined by a triangle with points located at the junction of the cystic duct and common bile duct, the second and third portion of the duodenum, and the neck and body of the pancreas (Fig. 33-65). However, because gastrinomas can be found almost anywhere, whole-body imaging is required. The test of choice is SSTR (octreotide) scintigraphy in combination with CT. The octreotide scan is more sensitive than CT, locating about 85% of gastrinomas and detecting tumors <1 cm. With the octreotide scan, the need for tedious and technically demanding selective angiography and measurement of gastrin gradients has declined. EUS is another modality that assists in the preoperative localization of gastrinomas. It is particularly helpful in localizing tumors in the pancreatic head or duodenal wall, where gastrinomas are usually <1 cm in size. A combination of octreotide scan and EUS detects >90% of gastrinomas.

It is important to rule out MEN1 syndrome by checking serum calcium levels before surgery because resection of the gastrinoma(s) in these patients rarely results in normalization of serum gastrin concentrations or a prolongation of survival. Only one-fourth of gastrinomas occur in association with the MEN1 syndrome. One-half of patients with gastrinomas will have solitary tumors while the remainder will have multiple gastrinomas. Multiple tumors are more common in patients with MEN1 syndrome. Aggressive surgical treatment is justified in patients with sporadic gastrinomas. If patients have MEN1 syndrome, the parathyroid hyperplasia is addressed with total parathyroidectomy and implantation of parathyroid tissue in the forearm.

Approximately 50% of gastrinomas metastasize to lymph nodes of the liver and are therefore considered malignant. Patients who meet criteria for operability should undergo exploration for possible removal of the tumor. Although the tumors are submucosal, a full-thickness excision of the duodenal wall is performed if a duodenal gastrinoma is found. All lymph nodes in Passaro’s triangle are excised for pathologic analysis. If the gastrinoma is found in the pancreas and does not involve the main pancreatic duct, it is enucleated. Pancreatic resection is justified for solitary gastrinomas with no metastases. A highly selective vagotomy can be performed if unresectable disease is identified or if the gastrinoma cannot be localized. This may reduce the amount of expensive proton pump inhibitors required. In cases in which hepatic metastases are identified, resection is justified if the primary gastrinoma is controlled and the metastases can be safely and completely removed. Debubling or incomplete removal of multiple hepatic metastases is probably not helpful, especially in the setting of MEN1. The application of new modalities such as radiofrequency ablation seems reasonable, but data to support this approach are limited. Postoperatively, patients are followed with fasting serum gastrin levels, secretin stimulation tests, octreotide scans, and CT scans. In patients found to have inoperable disease, chemotherapy with streptozocin, doxorubicin, and 5-fluorouracil (5-FU) is used. Other approaches such as somatostatin analogues, interferon, and chemoembolization also have been used in gastrinoma with some success.

Unfortunately, a biochemical cure is achieved in only about one-third of the patients operated on for ZES. Despite the lack of success, long-term survival rates are good, even in patients with liver metastases. The 15-year survival rate for patients without liver metastases is about 80%, while the 5-year survival rate for patients with liver metastases is 20% to 50%. Pancreatic tumors are usually larger than tumors arising in the duodenum, and more often they have lymph node metastases. In gastrinomas, liver metastases decrease survival rates, but lymph node metastases do not. The best results are seen after complete excision of small sporadic tumors originating in the duodenum. Large tumors associated with liver metastases, located outside of Passaro’s triangle, have the worst prognosis.

Vasoactive Intestinal Peptide-Secreting Tumor

In 1958, Verner and Morrison first described the syndrome associated with a pancreatic neoplasm secreting VIP. The classic clinical syndrome associated with this pancreatic endocrine neoplasm consists of severe intermittent watery diarrhea leading to dehydration, and weakness from fluid and electrolyte losses. Large amounts of potassium are lost in the stool. The vasoactive intestinal peptide-secreting tumor (VIPoma) syndrome is also called WDHA syndrome due to the presence of watery diarrhea, hypokalemia, and achlorhydria. The massive (5 L/d) and episodic nature of the diarrhea associated with the appropriate electrolyte abnormalities should raise suspicion of the diagnosis. Serum VIP levels must be measured on multiple occasions because the excess secretion of VIP is episodic and single measurements might be normal and misleading. A CT scan localizes most VIPomas, although as with all islet cell tumors, EUS is the most sensitive imaging method. Electrolyte and fluid balance is sometimes difficult to correct preoperatively and must be pursued aggressively. Somatostatin analogues are helpful in controlling the diarrhea and allowing replacement of fluid and electrolytes. VIPomas are more commonly located in the distal pancreas and most have spread outside the pancreas. Palliative debulking operations can sometimes improve symptoms for a period, along with somatostatin analogues. Hepatic artery embolization also has been reported as a potentially beneficial treatment.

Glucagonoma

Diabetes in association with dermatitis should raise the suspicion of a glucagonoma. The diabetes usually is mild. The classic necrolytic migratory erythema manifests as cyclic migrations of lesions with spreading margins and healing centers typically on
the lower abdomen, perineum, perioral area, and feet. Patients also complain of an enlarged, sensitive tongue. The diagnosis is confirmed by measuring serum glucagon levels, which are usually >500 pg/mL. Glucagon is a catabolic hormone, and most patients present with malnutrition. The rash associated with glucagonoma is thought to be caused by low levels of amino acids. Preoperative treatment usually includes control of the diabetes, parenteral nutrition, and octreotide. Like VIPomas, glucagonomas are more often in the body and tail of the pancreas and tend to be large tumors with metastases. Again, debulking operations are recommended in good operative candidates to relieve symptoms.

**Somatostatinoma**

Because somatostatin inhibits pancreatic and biliary secretions, patients with a somatostatinoma present with gallstones due to bile stasis, diabetes due to inhibition of insulin secretion, and steatorrhea due to inhibition of pancreatic exocrine secretion and bile secretion. Most somatostatinomas originate in the proximal pancreas or the pancreatoduodenal groove, with the ampulla and peripancreatic area as the most common site (60%). The most common presentations are abdominal pain (25%), jaundice (25%), and cholelithiasis (19%). This rare type of pancreatic endocrine tumor is diagnosed by confirming elevated serum somatostatin levels, which are usually >10 ng/mL. Although most reported cases of somatostatinoma involve metastatic disease, an attempt at complete excision of the tumor and cholecystectomy is warranted in fit patients.

**Nonfunctioning Pancreatic Endocrine Tumors**

Although some pancreatic endocrine neoplasms secrete one or more hormones and are associated with interesting characteristic clinical syndromes, most are not associated with elevated serum hormone levels that cause symptoms. Pancreatic endocrine tumors are considered functional if they are associated with a clinical syndrome and nonfunctioning if not associated with clinical symptoms. The majority of pancreatic endocrine tumors (PET), also called pancreatic neuroendocrine tumors (pNET), are malignant because they have the potential for uncontrolled growth and metastasis. Immunohistochemical markers such as synaptophysin, chromogranin A (CgA), and neuron-specific enolase can be helpful in the diagnosis, but the gross histology is not a reliable predictor of biologic behavior. CgA is used by some as a serum marker to monitor patients for disease recurrence or response to treatment, but the test performs poorly for this purpose. Patients often present similar to patients with pancreatic adenocarcinoma with vague pain or weight loss, but pNETs are increasingly discovered incidentally when imaging is performed for another reason. The tumor frequently enhances with arterial contrast (Fig. 33-66). Sometimes a cystic component is seen due to central necrosis. Octreoscan (somatostatin receptor scintigraphy) can be helpful to stage the disease. Surgical resection is typically recommended in fit patients in the absence of metastatic disease. For patients with tumors in the body and tail of the pancreas, this typically includes splenectomy. Enucleation and splenic preservation, although tempting in small tumors, fails to remove regional lymph nodes.

With advances in imaging, small indolent pNETs are being discovered with increasing frequency, and some surgeons are considering observation in these cases. There are several arguments in favor of observation of small (<2 cm) nonfunctional pNETs, particularly in MEN-1 patients. Pancreatic resection has significant morbidity and mortality. In the setting of MEN-1, resection is rarely curative, and most patients require reoperation. Also, the survival of these patients, even with metastatic disease, is generally excellent at least compared to pancreatic ductal adenocarcinoma. However, the 5-year survival with metastatic pNET is only 16% and radical surgery with curative intent is the standard of care, particularly in sporadic cases that are fit surgical candidates where the primary tumor can be completely excised.

Adjuvant treatment after resection is withheld in the absence of radiographically demonstrable metastatic disease even if CgA levels remain elevated. Although these tumors have a slow growth pattern compared to pancreatic ductal adenocarcinoma, many patients with pNETs will die of their disease even after an apparent complete resection, making surveillance after complete resection important. CT scan and or octreotide scan is recommended annually for 5 years after resection.

![Figure 33-66](image-url) Pancreatic neuroendocrine tumor (PNET) demonstrating enhancement during arterial phase of computed tomography scan. Pancreatic head PNET seen in (**left**) sagital, (**middle**) coronal, and (**right**) lateral views of the abdomen.
Incomplete resection (debulking) for locally advanced or metastatic pNETs of the pancreas is controversial because of the favorable survival duration of patients without surgery. However, in carefully selected fit patients with a pNET in the head of the pancreas and minimal disease in the liver, a pancreaticoduodenectomy with wedge resection of the liver metastasis might be appropriate because this avoids the morbidity of gastrointestinal hemorrhage and biliary and gastric outlet obstruction before death from the metastatic disease. The role of cytoreductive surgery in metastatic pNET is controversial, but consensus guidelines agree that aggressive resection of the primary tumor, regional lymph nodes, and liver/distant metastases should be pursued if greater than 90% of the tumor burden can be resected, which is the case in only about 10% of patients with metastatic disease.

Treatment of metastatic pancreatic neuroendocrine cancer requires a multidisciplinary approach often including a combination of cytoreductive surgery when appropriate, directed therapy for the treatment of liver metastases when possible, and systemic medical therapy. Local ablative therapies include radiofrequency ablation (RFA), cryotherapy, microwave coagulation, and ethanol injection, although RFA is the most popular and widely studied. These therapies can be performed percutaneously or during surgery via laparotomy or laparoscopy, can be repeated, and have been shown to complement resection of the primary tumor and amenable liver metastases; making palliative surgery possible for patients that would otherwise not meet criteria. This approach often improves symptoms and 5-year survival is improved to 48%.306

Transarterial chemoembolization (TACE) can be employed as palliative therapy in patients with liver metastases not amenable to surgical resection and/or ablation. It relies on the principle that metastatic tumor cells derive the majority of their oxygen supply from the hepatic artery as opposed to hepatocytes, which receive oxygen primarily from the portal vein. Performed via angiography, embolization may be performed alone (bland embolization) or in combination with chemotherapeutic agents (chemoembolization). Radioembolization, the selective distribution of radioactive yttrium-90 microspheres into the peritumoral vasculature via branches of the hepatic artery, is another alternative.

Somatostatin analogs can inhibit release of hormones from functional pNETs and reduce diarrhea in patients with VIPomas, glucagonomas, and somatostatinomas, and it can also help the rash of glucagonomas. Somatostatin analogs such as octreotide (Sandostatin LAR Depot), lanreotide (Somatuline Depot), and a new analog pasireotide (Sifnifor LAR) are also used to slow the growth of some functional and nonfunctional pNETs. Disease seen on an octreotide scan can somatostatin receptors and would be expected to respond. Patients with unresectable disease are often treated with somatostatin analogs first before targeted therapy and cytotoxic chemotherapy are utilized because the side effects are minimal.

Some targeted drugs can be helpful in treating advanced pNETs that progress despite somatostatin analogs. Sunitinib (Sutent) attacks new blood vessel growth and other targets that help cancer cells grow. Everolimus (Afinitor) works by blocking a cell protein known as mammalian target of rapamycin (mTOR) and the VEGF pathway, which normally helps cells grow and divide. These targeted agents induce stabilization much more frequently than response, but they tend to carry fewer side effects than traditional cytotoxic chemotherapy. Cytotoxic chemotherapy for pNETs is usually reserved for large tumors or quickly growing tumors that are causing symptoms, or tumors that progressed despite somatostatin analogs and targeted therapy. Platinum-based chemotherapy does, however, remain the standard of care for high-grade (poorly differentiated) pNET, yielding high response rates but typically short-lived benefit. The newer cytotoxic combination of capecitabine and temozolomide has shown activity in well-differentiated pNET.306

Neoplasms of the Exocrine Pancreas

Epidemiology and Risk Factors. It is estimated that in 2017, 53,670 Americans will be diagnosed with pancreatic cancer and 43,090 will die from the disease. Overall, pancreatic cancer has the worst prognosis of all malignancies with a 5-year survival rate of only 7.2%.307 The incidence of pancreatic cancer continues to increase, perhaps related to the increased incidence of risk factors such as obesity and diabetes, and as a result, it is predicted that pancreatic cancer will become the leading cause of cancer deaths in the United States by 2050. Pancreatic cancer recently surpassed breast cancer and is now the third leading cause of cancer death behind lung and colorectal cancer. Despite its ubiquity, this disease is extremely difficult to treat, and its exact cause is unknown. However, epidemiologic studies linking various environmental and host factors provide some clues. Recent discoveries using modern molecular biologic techniques have also improved our understanding of the causes of pancreatic cancer. The etiology of pancreatic cancer likely involves a complex interaction of genetic and environmental factors. These factors will become more fully understood as DNA sequencing is used to screen populations at risk for developing pancreatic cancer.

Pancreatic cancer is more common in older adults with most patients being 75 to 84 years old. Pancreatic cancer is more common in African Americans and slightly more common in men than women. The risk of developing pancreatic cancer is two to three times higher if a parent or sibling had the disease. Another risk factor that is consistently linked to pancreatic cancer is cigarette smoking. Smoking increases the risk of developing pancreatic cancer by at least twofold due to the carcinogens in cigarette smoke.308 Coffee and alcohol consumption have been investigated as possible risk factors, but the data are inconsistent. As in other GI cancers, diets high in fat and low in fiber, fruits, and vegetables are thought to be associated with an increased risk of pancreatic cancer.

Diabetes has been known to be associated with pancreatic cancer for many years. In fact, glucose intolerance is present in 80% of patients with pancreatic cancer, and approximately 20% have overt diabetes, a much greater incidence than would be expected to occur by chance. Preexisting type 2 diabetes increases the risk for development of pancreatic cancer by about twofold.309 The new onset of diabetes also can be an early manifestation of otherwise occult pancreatic cancer. Thus, the new onset of diabetes, or a sudden increase in insulin requirement in an older adult patient with preexisting diabetes, should provoke concern for the presence of pancreatic cancer.

Recent epidemiologic studies have confirmed the fact that patients with chronic pancreatitis, especially familial pancreatitis, have an increased risk of developing pancreatic cancer.124,196-198 Large, retrospective cohort studies of patients with pancreatitis have revealed up to a 20-fold increase in risk for pancreatic cancer. This increased risk seems to be independent of the type of pancreatitis, a finding consistent with the fact that most studies have shown little effect of alcohol ingestion per se.
on the risk of pancreatic carcinoma. The mechanisms involved in carcinogenesis in patients with preexisting pancreatitis are unknown. However, the mutated K-ras oncogene, which is present in most cases of pancreatic cancer, has been detected in the ductal epithelium of some patients with chronic pancreatitis.

**Genetics of Pancreatic Cancer.** Pancreatic carcinogenesis probably involves multiple mutations that are inherited and acquired throughout aging. The K-ras oncogene is currently thought to be the most commonly mutated gene in pancreatic cancer, with approximately 90% of tumors having a mutation. This prevalent mutation is present in precursor lesions and is therefore thought to occur early and be essential to pancreatic cancer development. K-ras mutations can be detected in DNA from serum, stool, pancreatic juice, and tissue aspirates of patients with pancreatic cancer, suggesting that the presence of this mutation or others may provide the basis for diagnostic testing in select individuals. The HER2/neu oncogene, homologous to the epidermal growth factor receptor (EGFr), is overexpressed in pancreatic cancers. This receptor is involved in signal transduction pathways that lead to cellular proliferation. Multiple tumor-suppressor genes are deleted and/or mutated in pancreatic cancer, including p53, p16, and DPC4 (Smad 4), and in a minority of cases, BRCA2. Most pancreatic cancers have three or more of the aforementioned mutations.

The genetic landscape of pancreatic adenocarcinoma has recently been investigated using exome capture technology combined with the SOLiD or Illumina next generation sequencing platforms and copy number analysis. Detailed analysis of 99 tumors reaffirmed the importance of the already known mutations such as KRAS, TP53, CDKN2A, SMAD4, MLL3, TGFB2, ARID1A, and SF3B1 in pancreatic cancer and identified eight novel significantly mutated genes involved in chromatin modification (EPC1 and ARID2), DNA damage repair (ATM) and other mechanisms (ZIM2, MAP2K4, NALCN, SLC16A4, and MAGEA6). Pathway-based analysis of recurrently altered genes also revealed the involvement of axon-guidance genes, particularly SLIT/ROBO signaling, in pancreatic carcinogenesis. Rapid and sensitive sequencing techniques will hopefully lead to better diagnostic and therapeutic approaches for pancreatic cancer.

It is estimated that up to 10% of pancreatic cancers occur as a result of an inherited genetic predisposition. A family history of pancreatic cancer in a first-degree relative increases the risk of pancreatic cancer by about twofold. Rare familial cancer syndromes that are associated with an increased risk of pancreatic cancer include BRCA2, the familial atypical multiple mole–melanoma syndrome, hereditary pancreatitis, familial adenomatous polyposis (FAP), and ataxia-telangiectasia.

**Pathology.** Pancreatic cancer probably arises through a stepwise progression of cellular changes, just as colon cancer progresses by stages from hyperplastic polyp to invasive cancer. Systematic histologic evaluation of areas surrounding pancreatic cancers has revealed the presence of precursor lesions that have been named pancreatic intraepithelial neoplasia (Fig. 33-67). Three stages of pancreatic intraepithelial neoplasia have been defined. These lesions demonstrate the same oncogene mutations and loss of tumor-suppressor genes found in invasive cancers, the frequency of these abnormalities increasing with progressive cellular atypia and architectural disarray. The ability to detect these precursor lesions in humans at a stage where the cancer can still be prevented or cured is an important goal of current pancreatic cancer research.

About two-thirds of pancreatic adenocarcinomas arise within the head or uncinate process of the pancreas; 15% are in the body, and 10% are in the tail, with the remaining tumors demonstrating diffuse involvement of the gland. Tumors in the pancreatic body and tail are generally larger at the time of diagnosis, and therefore, less commonly resectable. Tumors in the head of the pancreas are typically diagnosed earlier because they cause obstructive jaundice. Ampullary carcinomas, carcinomas of the distal bile duct, and periampullary duodenal adenocarcinomas present in a similar fashion to pancreatic head cancer but have a slightly better prognosis, probably because early obstruction of the bile duct and jaundice leads to the diagnosis.

In addition to ductal adenocarcinoma, which makes up about 75% of nonendocrine cancers of the pancreas, there are a variety of less common types of pancreatic cancer. Adenosquamous carcinoma is a variant that has both glandular and squamous differentiation. The biologic behavior of this lesion is unfortunately no better than the typical ductal adenocarcinoma. Acinar cell carcinoma is an uncommon type of pancreatic cancer that usually presents as a large tumor, often 10 cm in diameter or more, but the prognosis of patients with these tumors may be better than with ductal cancer.

**Diagnosis and Staging.** Exact pathologic staging of pancreatic cancer is important because it allows accurate quantitative assessment of results and comparisons between institutions. The tumor-node-metastasis (TNM) staging of pancreatic cancer was updated by the American Joint Committee on Cancer in 2017 (AJCC) (Table 33-21).

The important changes in the staging include more stratification for tumor size and stratification for number of lymph nodes involved. In the new system, tumors ≤2 cm remain as T1 lesions but are subcategorized as T1a (≤0.5 cm), T1b (>0.5–<1 cm), and T1c (1–2 cm). Tumors that are >2–<4 cm are categorized as T2, and tumors >4 cm are categorized as T3. Tumors that involve the celiac axis, superior mesenteric artery, and/or common hepatic artery are still categorized as T4 regardless of size. Metastasis in one to three regional lymph nodes is considered N1 disease, and involvement of ≥4 regional lymph nodes is N2. Patients without nodal involvement (N0) are stages IA, IB, and IIA based on tumor size (T1–T3, respectively). Patients with one to three regional lymph nodes involved (N2) are stage IIB regardless of tumor size. All patients with N2 disease without distant metastases are considered stage III regardless of tumor size. Patients with vascular involvement (T4) are stage III regardless of nodal involvement, and patients with distant metastases are, of course, stage IV.

Ten percent of pancreas cancer cases are diagnosed while the cancer is still confined to the primary site (localized stage); 29% are diagnosed after the cancer has spread to regional lymph nodes or directly beyond the primary site; 52% are diagnosed after the cancer has already metastasized (distant stage); and for the remaining 9%, the staging information was unknown. The corresponding 5-year relative survival rates were 31.5% for localized, 11.5% for regional, 2.7% for distant, and 5.1% for unstaged. The overall 5-year relative survival rate for patients with pancreatic cancer for 2007 to 2013 from Surveillance, Epidemiology, and End Results (SEER) was 8.2%.
The most critical deficit in the ability to treat pancreatic cancer effectively is the lack of tools for early diagnosis. The pancreas is situated deep within the abdomen, and the early symptoms of pancreatic cancer often are too vague to raise suspicion of the disease. Ultimately, the majority of patients present with pain and jaundice. On physical examination, weight loss is evident and the skin is icteric; a distended gallbladder is palpable in about one-fourth of patients. More fortunate patients have tumors situated such that biliary obstruction and jaundice occur early and prompt diagnostic tests. Unfortunately, however, the vast majority of patients are not diagnosed until weight loss has occurred—a sign of advanced disease.

Although it is often taught that carcinoma of the pancreas presents with painless jaundice (to help distinguish it from choledocholithiasis), this aphorism is not accurate. Most patients do experience pain as part of the symptom complex of pancreatic cancer, and it is often the first symptom. Therefore, awareness of the way pancreatic pain is perceived may help clinicians suspect pancreatic cancer. The pain associated with pancreatic cancer is usually perceived in the epigastrium but can occur in any
The current diagnostic and staging test of choice for pancreatic cancer is a multidetector, dynamic, contrast-enhanced CT scan, and the techniques for obtaining high-quality images are constantly improving (Fig. 33-68). The accuracy of CT scanning for predicting unresectable disease is about 90% to 95%. In contrast, CT scanning is less accurate in predicting resectable disease. CT scanning will miss small liver metastases, and predicting arterial involvement is sometimes difficult. CT findings that indicate a tumor is unresectable include involvement of ≥180° of the celiac axis, hepatic or superior mesenteric artery, enlarged lymph nodes outside the boundaries of resection, ascites, and distant metastases (e.g., liver). Invasion of the superior mesenteric vein or portal vein is not in itself a contraindication to resection as long as the veins are patent. Tumors are considered “borderline resectable” if there is abutment of ≤180 degrees of the circumference of the SMA, celiac axis, or hepatic artery or if there is a short segment of vein occlusion. Also, patients with CT findings suspicious for metastatic disease, like 1 mm liver lesions too small to characterize or biopsy, are considered “borderline resectable” as are patients with multiple comorbidities or marginal performance status. There is growing consensus that neoadjuvant treatment should be considered in all patients with any radiographic evidence of extension to adjacent vascular structures.

Currently, multidetector CT is probably the single most versatile and cost-effective tool for the diagnosis and staging of pancreatic cancer. Abdominal MRI provides essentially the same information as CT scanning. Positron emission tomography scanning in locally advanced lesions may help rule out distant metastases. EUS can be used to detect small pancreatic masses that could be missed by CT scanning and is commonly used when there is a high suspicion for pancreatic cancer but no mass is identified by the CT scan. EUS has the added advantage of providing the opportunity for transluminal biopsy of pancreatic masses, although a tissue diagnosis before pancreaticoduodenectomy is not required. However, in specific patients a histologic diagnosis may be necessary such as for those in a
neoadjuvant clinical trial or before chemotherapy in advanced tumors. EUS is a sensitive test for portal/superior mesenteric vein invasion, although it is somewhat less effective at detecting superior mesenteric artery invasion. When all of the current staging modalities are used, their accuracy in predicting resectability has improved.

As imaging continuously improves and high-quality imaging is always obtained before surgery, the chance of bringing a patient to the operating room with the intent of a curative resection and finding upon exploration that the patient has unresectable disease is becoming increasingly uncommon.

In an attempt to avoid such futile laparotomies, preliminary laparoscopy has been advocated for patients with disease felt to be resectable by CT imaging (Fig. 33-69). Diagnostic laparoscopy with the use of US is reported to improve the accuracy of predicting resectability to about 98%. The technique involves more than simple visualization with the scope and requires the placement of multiple ports and manipulation of the tissues. A general exploration of the peritoneal surfaces is carried out. The ligament of Treitz and the base of the transverse mesocolon are examined for tumor. The gastrocolic ligament is incised, and the lesser sac is examined. The ultrasound probe is used to examine the liver, porta hepatis and the portal vein, the celiac axis, and the superior mesenteric artery.

The percentage of patients in whom a positive laparoscopy helps avoid a nontherapeutic laparotomy varies from 10% to 30% in carcinoma of the head of the pancreas, but it may be as high as 50% in patients with tumors in the body and tail of the gland. Resection for pancreatic cancer is being approached laparoscopically, particularly for tumors in the body and tail of the pancreas thus eliminating the need for any separate staging laparoscopy procedure. Also, as the quality of CT scanning has improved, the value of routine diagnostic laparoscopy has decreased. The morbidity of diagnostic laparoscopy is less than that of laparotomy, and the procedure can be performed on an outpatient basis. Patients who are found to have unresectable disease recover more rapidly from a laparoscopy than a laparotomy and can receive palliative chemotherapy and radiation sooner. The potential immunosuppressive effects of a major surgical procedure also are avoided, as well as the negative psychologic impact of a major painful operation with little benefit.

Biliary obstruction can be relieved with an endoscopic approach in almost all cases. When large (10F) plastic stents are used, most patients do not require replacement for about 3 months. Metallic wall stents last about 5 months on average and usually fail only with tumor ingrowth. Keeping in mind that patients with unresectable pancreatic cancer usually live <1 year, the requirement for numerous stent changes is unlikely.

Diagnostic laparoscopy is possibly best applied to patients with pancreatic cancer on a selective basis. Diagnostic laparoscopy will have a higher yield in patients with large tumors (T3,
>4 cm), tumors located in the body or tail, patients with equivocal findings of metastasis on CT scan, and patients with other indications of advanced disease such as marked weight loss or markedly elevated CA19-9 (>1000 U/mL).

An algorithm for the diagnosis, staging, and treatment of pancreatic cancer is shown in Fig. 33-70. In practice, many of these patients are selected for neoadjuvant chemotherapy and then undergo restaging CT and staging laparoscopy prior to surgery.

**Palliative Surgery and Endoscopy.** Most patients with pancreatic cancer (85–90%) have disease that clearly precludes surgical resection. For these patients, appropriate and effective palliative treatment is critical to the quality of their remaining life. Because of the poor prognosis of the disease, it is not appropriate to use invasive, toxic, and expensive regimens in patients with extremely advanced disease and poor performance status. When patients do desire antineoplastic therapy, it is important to encourage them to enroll in clinical trials so that therapeutic advances can be made. In general, there are three clinical problems in advanced pancreatic cancer that require palliation: pain, jaundice, and duodenal obstruction. The mainstay of pain control is oral narcotics. Sustained-release preparations of morphine sulfate are frequently used. Invasion of retroperitoneal nerve trunks accounts for the severe pain experienced by patients with advanced pancreatic cancer. A celiac plexus nerve block can control pain effectively for a period of months, although the procedure sometimes needs to be repeated.

Jaundice is present in the majority of patients with pancreatic cancer, and the most troublesome aspect for the patient is the accompanying pruritus. Biliary obstruction may also lead to cholangitis, coagulopathy, digestive symptoms, and hepato cellular failure. In the past, surgeons traditionally performed a biliary and enteric bypass when unresectable disease was found at laparotomy. This is an increasingly uncommon situation for the surgeon because locally advanced unresectable disease is now detected by high-quality preoperative imaging. Metastatic disease is also more reliably predicted by preoperative imaging and, in select cases, staging laparoscopy.

In current practice, jaundice is usually palliated by an endoscopic biliary stent, often prior to surgical referral. Duodenal obstruction is usually a late event in pancreatic cancer and occurs in only about 20% of patients. Endoscopic metallic duodenal stents are an option, but the patient’s poor prognosis, the cost, and the fact that duodenal stents often do not result in ideal palliation has to be considered. The results of three RCTs examining endoscopic metallic stenting for malignant gastric outlet obstruction demonstrated that major and minor complications were comparable to gastrojejunoscopy but time to tolerating oral intake and hospital stay was shorter. Robot-assisted laparoscopic biliary-enteric bypass is now available at many centers. Although this should result in similar palliation as an open bypass and may be associated with more rapid recovery, the potential complications, the patient’s life expectancy, and cost of this procedure must be considered. As many patients today already have a bile duct stent in place by the time of referral to a surgeon, it is not clear that operative biliary bypass is required. In patients with extensive metastases, an alternative short-term palliative option to consider in patients with gastric outlet obstruction is a percutaneous endoscopic gastrostomy tube or gastrojejunal feeding tube that allows decompression of the stomach and feeding into the jejunum.

If an operative bypass is performed, choledochojejunostomy is the preferred approach. Although an easy procedure to perform, choledochoiodudenostomy is felt to be unwise because of the proximity of the duodenum to tumor. Some have discouraged the use of the gallbladder for biliary bypass; however, it...
is suitable as long as the cystic duct clearly enters the common duct well above the tumor. The jejunum is brought anterior to the colon, if possible, rather than retrocolic, where the tumor potentially would invade the bowel sooner. Some surgeons use a loop of jejunum with a jejunojejunostomy to divert the enteric stream away from the biliary-enteric anastomosis. Others use a Roux-en-Y limb with the gastrojejunalostomy located 50 cm downstream from the hepaticojejunostomy (Fig. 33-71). Potential advantages of the defunctionalized Roux-en-Y limb include the ease with which it will reach up to the hepatic hilum, probable decreased risk of cholangitis, and easier management of biliary anastomotic leaks. If a gastrojejunostomy is performed, it should be placed dependently and posterior along the greater curvature to improve gastric emptying, and a vagotomy should not be performed. Endoscopic stents are definitely not as durable as a surgical bypass. Recurrent obstruction and cholangitis is more common with stents and results in inferior palliation. However, the endoscopic approach is associated with considerably less initial morbidity and mortality than surgical bypass. Expandable metallic wall stents have superior patency and provide better palliation than plastic stents (Fig. 33-72).

If an initial diagnostic laparoscopy reveals a contraindication to the Whipple procedure, such as liver metastases, it is not appropriate to perform a laparotomy simply to create a biliary bypass. In such a patient, it is better to place an endoscopic stent. In contrast, in the uncommon scenario where a laparotomy has already been performed as part of the assessment of resectability and the Whipple procedure is not possible, a surgical bypass is usually performed. However, if the patient has a functioning endoscopic stent already in place, it may be reasonable to forego surgical bypass.

**Chemotherapy and Radiation for Locally Advanced/Metastatic Disease.** Patients with locally advanced unresectable disease are treated with chemotherapy and possibly radiation, and patients with stage IV metastatic disease are treated with systemic chemotherapy. The role of RT in unresectable, locoregionally advanced pancreas cancer remains controversial. RT may slow the progression of local disease and possibly alleviate or prevent symptoms including pain, biliary obstruction, bleeding, and bowel obstruction. However, the likelihood of micrometastatic distant disease is high, treatment is not expected to be curative, and radiation can result in toxicity. Stereotactic body radiotherapy (SBRT) has been used to limit toxicity by targeting high-dose short-course radiation to enhance local response prior to surgery. In a phase 2 multi-institutional trial evaluating gemcitabine and SBRT in patients with locally advanced unresectable pancreatic cancer, 10% of patients with locally advanced disease who would not have been candidates for initial surgery were deemed to have resectable tumors following therapy, and 8% ultimately underwent R0 and node-negative resection. A number of chemotherapy regimens are available for pancreatic cancer, but the results are not impressive. Gemcitabine (Gemzar) was approved by the U.S. Food and Drug Administration (FDA) for use in pancreatic cancer in 1996. In patients with unresectable pancreatic cancer, gemcitabine results in symptomatic


Figure 33-72. Expandable metallic biliary stent. After ERCP cannulation of the distal bile duct (left) the stent is advanced over the cannula and placed across the obstruction in the distal bile duct (right).
improvement, improved pain control and performance status, and weight gain. However, survival is improved by only 1 to 2 months. Prior to gemcitabine, 5-fluorouracil (5-FU) was used as the standard treatment for unresectable pancreatic cancer. Both of these drugs are still used today. 5-FU or capecitabine (Xeloda), a similar but orally administered drug, are frequently used as a radiosensitizer during radiation therapy. Single-agent gemcitabine is still commonly used in patients with a poor performance status.

Erlotinib (Tarceva) was approved in 2005 based on very minimal improvement in overall survival in combination with gemcitabine. The study showed that erlotinib in combination with gemcitabine results in a statistically significant improvement in overall survival in patients with advanced pancreatic cancer in the first-line setting. Although the absolute benefit in overall survival was modest with a median survival difference between the two arms of only 2 weeks.

FOLFIRINOX, a combination of three chemotherapy drugs (5-FU/leucovorin, irinotecan, and oxaliplatin) is now commonly used as first-line treatment for metastatic pancreatic adenocarcinoma in patients with a relatively good performance status. In 2010, a Phase 3 clinical trial showed positive results for patients treated with FOLFIRINOX. The objective response rate was improved from 9% to 32%, and median overall survival of patients with metastatic pancreatic cancer improved from 7 to 11 months, but the improvement was associated with increased toxicity, so patient selection is important. Patients treated with FOLFIRINOX may experience more severe side effects than those treated with gemcitabine alone, so this combination is usually reserved for patients with a good performance status.

In 2013, another combination therapy was approved as first-line treatment for metastatic pancreatic adenocarcinoma. Albumin-bound paclitaxel (Abraxane) was approved to be used in combination with gemcitabine (Gemzar). The median overall survival was improved to 8.5 months in the nabpaclitaxel–gemcitabine group as compared with 6.7 months in the gemcitabine group. Progression-free survival and the response rate were also improved. Rates of peripheral neuropathy and myelosuppression were increased, but this regimen is less toxic than FOLFIRINOX.

In 2015, Irinotecan (ONIVYDE), in combination with 5-FU (fluorouracil) and leucovorin, was approved for treatment of metastatic pancreatic adenocarcinoma that has progressed following treatment with a gemcitabine based therapy. The median overall survival in patients assigned to nanoliposomal irinotecan plus fluorouracil and folinic acid was 6.1 months vs. 4.2 months with fluorouracil and folinic acid. Common side effects for this second-line therapy were neutropenia, diarrhea, vomiting, and fatigue.

These results may warrant treatment in patients who understand the benefits and risks. However, the lack of significant survival advantage should encourage physicians to refer motivated patients for experimental protocols because it is only through continued clinical research that more meaningful treatments for pancreatic cancer will be developed.

Ablation for Locally Advanced Unresectable Disease. Persistent arterial vascular encasement after neoadjuvant therapy contraindicates resection. Irreversible electroporation utilizes delivery of high-voltage millisecond electrical pulses resulting in permanent disruption of the cellular membranes and subsequent apoptosis. This process leads to cell death, but does not injure the extracellular matrix, thus allowing cellular tumor ablation while preserving structural components of tissues. Collagen-based structures such as vessels or the pancreatic duct are not disrupted. Furthermore, because IRE is not based on thermal damage of cancer cells, the heat-sink phenomenon is not a concern, and even lesions abutting large vessels can be ablated with radical intent. Irreversible electroporation using the Nanoknife is reported to enable treatment of pancreatic tumors abutting vascular structures without compromise of the vessels or concern for the heat sink effect of nearby blood flow. Martin has created a registry and accumulated multi-institutional data on 200 patients with locally advanced pancreatic cancer showing OS of 28.3 months for patients with borderline resectable pancreatic cancer and 23.2 months in patients with unresectable pancreatic cancer. Those numbers compare favorably with the survival of patients treated with chemoradiation alone, which is 13 months in historical controls. This modality is new, but these early reports indicate it may be safe in combination with chemotherapy and are of particular interest because of the potential to downstage and offer surgery to patients initially diagnosed with locally advanced unresectable disease. It is also important to understand that a significant learning curve exists to achieve safety and optimization of the technique. Proper patient selection, technical ability with intraoperative ultrasound to allow precise IRE electrode bracketing, and standardization of the IRE energy delivery is important. Randomized prospective trials are needed before adoption of this technique can be expanded.

Surgical Resection: Pancreaticoduodenectomy. In a patient with appropriate clinical and/or imaging indications of pancreatic cancer, a tissue diagnosis before performing a pancreaticoduodenectomy is not essential. Although percutaneous CT-guided biopsy is usually safe, complications such as hemorrhage, pancreatitis, fistula, and abscess can occur. Tumor seeding along the subcutaneous tract of the needle is uncommon. Likewise, FNA under EUS guidance is safe and well tolerated. The problem with preoperative or even intraoperative biopsy is that many pancreatic cancers are not very cellular and contain a significant amount of fibrous tissue, so a biopsy may be misinterpreted as showing chronic pancreatitis if it does not contain malignant glandular cells. In the face of clinical and radiologic preoperative indications of pancreatic cancer, a negative biopsy should not preclude resection. In patients who are not candidates for resection because of metastatic disease, biopsy for a tissue diagnosis becomes important because these patients may be candidates for palliative chemotherapy trials. It is especially important to make an aggressive attempt at tissue diagnosis before surgery in patients whose clinical presentation and imaging studies are more suggestive of alternative diagnoses such as pancreatic lymphoma or pancreatic islet cell tumors. These patients might avoid surgery altogether in the case of lymphoma or warrant an aggressive approach in the case of islet cell carcinoma.

Pancreatectoduodenectomy can be performed through a midline incision from xiphoid to umbilicus or through a bilateral subcostal incision. The initial portion of the procedure is an assessment of resectability. The liver and visceral and parietal peritoneal surfaces are thoroughly assessed. The gastrohepatic omentum is opened, and the celiac axis area is examined for enlarged lymph nodes. The base of the transverse mesocolon is examined for tumor involvement.

The ascending and hepatic flexure of the colon are mobilized off the duodenum and head of the pancreas and reflected medially. A Kocher maneuver is performed by dissecting behind
the head of the pancreas. The superior mesenteric vein is identified early in the case and dissected up toward the inferior border of the neck of the pancreas. The gastroepiploic vein and artery are ligated to prevent any traction injury. Often, the middle colic vein and right gastroepiploic vein share a common trunk before entering into the superior mesenteric vein. Knowledge of this anatomy helps reduce injury to the veins and unnecessary blood loss. The relation of the tumor to the superior mesenteric vein and artery cannot be accurately assessed by palpation at this point and is not completely determined until later in the operation when the neck of the pancreas is divided and the surgeon is committed to resection. Mesenteric vascular involvement is best determined by a high quality preoperative CT scan.

It is important to assess for an aberrant right hepatic artery, which is present in 20% of patients. The aberrant artery commonly arises from the superior mesenteric artery posterior to the pancreas and ascends parallel and adjacent to the superior mesenteric and portal veins. The presence of an aberrant right hepatic artery should be apparent on the preoperative CT scan and can be identified intraoperatively by palpation on the back side of the hepatoduodenal ligament, where a prominent pulse will be felt posterior and to the right of the portal vein.

The porta hepatis is examined. Enlarged or firm lymph nodes that can be swept down toward the head of the pancreas with the specimen do not preclude resection. If the assessment phase reveals no contraindications to the Whipple procedure (Table 33-22), the resection phase commences.

If the pylorus is to be preserved, the stomach and proximal duodenum are mobilized off the pancreas, preserving the gastroepiploic vessels down to the pylorus. The proximal hepatic artery is identified usually by removing a lymph node that commonly lies just anterior to the artery. The hepatic artery is dissected and traced toward the porta hepatis. The gastroduodenal branch of the hepatic artery is identified. A test clamping is performed to ensure that a strong pulse remains in the proper hepatic artery before division of the gastroduodenal artery. In cases of celiac occlusion, flow comes from the superior mesenteric artery and retrograde through the gastroduodenal artery to the proper hepatic artery. Ligation of the gastroduodenal in this case would be equivalent to hepatic artery ligation. A bypass to the hepatic artery would be required. Once the test clamping is negative and the gastroduodenal artery is divided, the hepatic artery is retracted medially, and the common bile duct is retracted laterally to reveal the anterior surface of the portal vein behind them. Dissection is performed only on the anterior surface of the vein. If there is no tumor involvement, the neck of the pancreas will separate from the vein easily. A large, blunt-tipped clamp is a safe instrument to use for this dissection. The tunnel under the neck of the pancreas can then be completed mostly under direct vision from inferior and superior.

The gallbladder is then mobilized from the liver, the cystic duct and artery are ligated, and the gallbladder is removed. The common hepatic duct is circumferentially dissected. Either the duodenum is divided 2 cm distal to the pylorus, which defines the procedure as a pylorus-preserving pancreaticoduodenectomy, or PPPD) or the antrum is divided, as classically described by Whipple. The jejunum is divided beyond the ligament of Treitz, and the mesentery is ligated until the jejunum can be delivered posterior to the superior mesenteric vessels from left to right.

The common hepatic duct is then divided usually just above the entrance of the cystic duct, and the bile duct is dissected down to the superior margin of the duodenum. Inferior traction on the distal bile duct opens the plane to make visible the anterior portion of the portal vein. The pancreatic neck is divided anterior to the portal vein (Fig. 33-73). The use of cautery is avoided in the area of the pancreatic duct. The pancreatic head and uncinate process then are dissected off of the right lateral aspect of the superior mesenteric vein, ligating the fragile branches draining the head and uncinate process into the portal vein (Fig. 33-74). The uncinate process is then dissected off of the posterior and lateral aspect of the superior mesenteric artery. This can be the most tedious portion of the operation.

Table 33-22

<table>
<thead>
<tr>
<th>Findings contraindicating resection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liver metastases (any size)</td>
</tr>
<tr>
<td>Celiac lymph node involvement</td>
</tr>
<tr>
<td>Peritoneal implants</td>
</tr>
<tr>
<td>Hepatic hilar lymph node involvement</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Findings not contraindicating resection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invasion at duodenum or distal stomach</td>
</tr>
<tr>
<td>Involved peripancreatic lymph nodes</td>
</tr>
<tr>
<td>Involved lymph nodes along the porta hepatis that can be swept down with the specimen</td>
</tr>
</tbody>
</table>
Minimally Invasive Pancreatectomy. Laparoscopic distal pancreatectomy has been proven to be safe and is appropriate for essentially all indications for pancreatectomy. This approach is associated with decreased blood loss and quicker recovery. Centrally located lesions near the splenoportal confluence must be approached laparoscopically with caution. Robotic technology may make dissection and control of major vessels in this area easier and perhaps safer. Interest in laparoscopic pancreatectoduodenectomy is increasing throughout the United States with some early adopters reporting excellent outcomes comparable to the open procedure. Early results indicate this technique is feasible, but considerable expertise is required in both open pancreatic resection as well as advanced laparoscopic techniques to achieve these outcomes. Whether the advantages seen in other areas of minimally invasive surgery apply to the Whipple procedure is an area of current investigation.

Variations and Controversies. The preservation of the pylorus has several theoretical advantages, including prevention of reflux of pancreaticobiliary secretions into the stomach, decreased incidence of marginal ulceration, normal gastric acid secretion and hormone release, and improved gastric function. Patients with pylorus-preserving resections have appeared to regain weight better than historic controls in some studies. Return of gastric emptying in the immediate postoperative period may take longer after the pylorus-preserving operation, and it is controversial whether there is any significant improvement in long-term quality of life with pyloric preservation.

Techniques for the pancreaticojejunostomy include end-to-side or end-to-end and duct-to-mucosa sutures or invagination (Fig. 33-75). Pancreaticogastrostomy has also been investigated. Some surgeons use stents, glue to seal the anastomosis, or octreotide to decrease pancreatic secretions. No matter what combination of these techniques is used, the clinically significant pancreatic leakage rate is always about 10%. Therefore, the choice of techniques depends more on the surgeon’s personal experience.

Traditionally, most surgeons place drains around the pancreatic and biliary anastomoses because disruption of the pancreaticojejunostomy cannot be avoided in one out of 10 patients.
This complication can lead to the development of an upper abdominal abscess or can present as an external pancreatic fistula. Usually, a pure pancreatic leak is controlled by the drains and will eventually seal spontaneously. Combined pancreatic and biliary leaks are cause for concern because bile will activate the pancreatic enzymes. In its most virulent form, disruption leads to necrotizing retroperitoneal infection, which can erode major arteries and veins of the upper abdomen, including the exposed portal vein and its branches or the stump of the gastroduodenal artery. Impending catastrophe is often preceded by a small herald bleed from the drain site. Depending on the clinical situation, such an event is an indication to perform an angiogram or return the patient to the operating room to widely drain the pancreaticojejunostomy and to repair the involved blood vessel. Open packing may be necessary to control diffuse necrosis and infection. Some studies have questioned the practice of routine drain placement after pancreatectomy with reliance on postoperative percutaneous drainage when leaks occur. However, a randomized, controlled, multicenter trial showed that patients who develop a leak after pancreaticoduodenectomy are

Figure 33-75. Techniques for pancreaticojejunostomy. **A** to **D**, Duct-to-mucosa, end-to-side. **E**, Intraoperative photographs of end-to-side pancreaticojejunostomy. **F** to **J**, End-to-end invagination. **K** to **O**, End-to-side invagination.
at a substantially increased risk of mortality if a drain was not placed at the time of resection.\textsuperscript{337} In contrast, the outcome does not seem to be as dramatically affected by drain placement in the setting of distal pancreatectomy.\textsuperscript{338} In the absence of a fistula, drains should be removed early in the postoperative period, preferably by postoperative day 5.\textsuperscript{339}

Many patients with pancreatic cancer are malnourished preoperatively and suffer from gastroparesis in the immediate postoperative period. Routine placement of a feeding jejunostomy tube and gastrostomy tube has become less common, and most surgeons use these tubes selectively. Gastrostomy tubes may decrease the length of stay in patients who might

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure33-75}
\caption{(Continued)}
\end{figure}
be predicted to have severe gastroparesis. Jejunostomy tubes are certainly not benign and can result in leaks and intestinal obstruction. However, parenteral nutrition is also associated with serious complications such as line sepsis, loss of gut mucosal integrity, and hepatic dysfunction. Enteric tubes should be considered in patients at risk such as malnourished patients who have received neoadjuvant chemotherapy.

Because of the high incidence of direct retroperitoneal invasion and regional lymph node metastasis at the time of surgery, trails of more extended resections including extension of the pancreatic resection to the middle body of the pancreas, segmental resection of the portal vein, if necessary, resection of retroperitoneal tissue along the right perinephric area, and lymphadenectomy to the region of the celiac plexus were examined. In the hands of experienced surgeons, these techniques are associated with greater blood loss but no increase in mortality. However, improved survival has not been demonstrated. Total pancreatectomy has also been considered in the past. Although pancreatic leaks are eliminated, major morbidity from brittle diabetes and exocrine insufficiency outweigh any theoretical benefit.

Pancreatic cancer can recur locally after pancreaticoduodenectomy. Intraoperative radiotherapy (IORT) delivers radiation to the operative bed at the time of resection. Radiation to
surrounding normal areas is minimized, but the radiation is delivered all in one setting, rather than in fractionated doses over time. Favorable results were recently reported among a series of patients with locally advanced unresectable or borderline-resectable PDAC who received intensive neoadjuvant treatment followed by exploratory laparotomy and IORT.340

Complications of Pancreaticoduodenectomy. The operative mortality rate for pancreaticoduodenectomy has decreased to <5% in high-volume centers (where individual surgeons perform more than 15 cases per year), suggesting that patients in rural areas would benefit from referral to large urban centers.341-342 The most common causes of death are sepsis, hemorrhage, and cardiovascular events. Postoperative complications are unfortunately still very common and include delayed gastric emptying, pancreatic fistula, and hemorrhage.

Delayed gastric emptying is common after pancreaticoduodenectomy and is treated conservatively as long as complete gastric outlet obstruction is ruled out by a contrast study. In the acute phase, intravenous erythromycin may help, but the problem usually improves with time.

Considerable attention has been focused on the prevention of pancreatic leak after pancreas resection. Modifications of the anastomotic technique (end-to-side or end-to-end, duct-to-mucosa, or invaginated), the use of jejunum or the stomach for drainage, the use of pancreatic duct stents, the use of octreotide, and various sealants have all been evaluated.

Long-acting synthetic analogues of somatostatin have been evaluated as a pharmacologic therapy to reduce pancreatic secretion and the rate of pancreatic fistula after pancreatic resection. Some European studies supported benefit particularly in selected higher risk patients, while previous North American trials concluded there was no benefit.343-349 A recent single-center, randomized trial with pasireotide (a newer analog) suggested potential benefit.350

Many technical modifications to the classic pancreaticoduodenectomy have been described. However, numerous technical variations to the pancreaticoenteric anastomosis have not clearly demonstrated an objective method to consistently decrease the rate of clinically significant postoperative pancreatic fistula, which in most series is about 10%. Yeo compared the incidence of pancreatic fistula in patients who had a pancreaticoduodenectomy with reconstruction via a pancreaticogastrostomy or pancreaticojejunostomy.351 There was no significant difference between the two techniques in the incidence of pancreatic fistula. A recent meta-analysis summarized the results of 16 trials comparing pancreaticogastrostomy to pancreaticojejunostomy. All of the observational clinical studies reported superiority of pancreaticogastrostomy over pancreaticojejunostomy, most likely influenced by publication bias. In contrast, all randomized prospective trials failed to show advantage of a particular technique, suggesting both techniques provide equally good results.352

Other options to consider when performing the pancreatic anastomosis are the duct-to-mucosa vs. the invagination techniques. Some surgeons choose the technique at the time of operation, depending on the size of the pancreatic duct and the texture of pancreas favoring invagination when the duct is small and the pancreatic texture is soft.353 Other surgeons use the same technique every time. The duct-to-mucosa anastomosis results in a low pancreatic fistula rate, particularly in patients with a large pancreatic duct and a fibrotic pancreas.354

Use of a pancreatic duct stent across the anastomosis has been suggested as a means of preventing a pancreatic leak and as an aid in technical precision. Both internal stenting as well as external stenting have been practiced. A recent Cochrane analysis of eight randomized, controlled trials failed to identify any convincing evidence of benefit with internal or external pancreatic duct stents.355 Some previous studies indicated that stents might be harmful. A recent multicenter randomized trial comparing external to internal pancreatic duct stents during pancreaticoduodenectomy showed a lower rate of pancreatic fistula with internal stents, so this controversy is likely to continue.356

Reconstruction with an isolated Roux-en-Y pancreaticoenteric anastomosis has been suggested as a method to decrease postoperative pancreatic leak.357-358 The logic behind this technical modification is that the use of separate Roux-en-Y limbs for biliary and pancreatic secretions may protect the pancreatic anastomosis from activated pancreatic enzymes. However, data is limited, and this is not a common practice.

Avoiding the pancreatic anastomosis altogether by ductal ligation or occlusion has also been evaluated as a potential technique to reduce the rate of postoperative pancreatic fistula.359-360 Ductal occlusion with neoprene or prolamine, which are nonresorbable glues, has been abandoned due to pancreatic atrophy and loss of exocrine function. Duct occlusion in pancreaticojejunostomy significantly increases the risk of endocrine insufficiency without a decrease in the postoperative complication rate. To avoid long-term loss of function, absorbable glues, such as fibrin glue, have been evaluated to limit the action of pancreatic enzymes until the anastomosis is healed. Fibrin glue has been used for both duct occlusion and has also been applied to the surface of the pancreatic stump and anastomotic site without clear improvement in pancreatic fistula rate. The effect of BioGlue applied to the anastomotic surface after the Whipple procedure and pancreatic stump after distal pancreatectomy was evaluated in a retrospective cohort study. There were no statistically significant differences in the incidence or severity grades of postoperative pancreatic fistulas.361 A randomized prospective trial of application of fibrin glue to the surface of the pancreaticojejunostomy in high-risk patients did not reduce the incidence of pancreatic fistula or total complications after pancreaticoduodenectomy.362

If not combined with a biliary leak, pancreatic fistula, although serious, can usually be managed conservatively. In about 95% of cases, reoperation is not indicated, and prolonged drainage, using drains placed in the original operation or percutaneously after resection, results in spontaneous closure of the fistula.363

Hemorrhage can occur either intraoperatively or postoperatively. Intraoperative hemorrhage typically occurs during the dissection of the portal vein. A major laceration of the portal vein can occur at a point in the operation at which the portal vein is not yet exposed. Temporary control of hemorrhage is generally possible in this situation by compressing the portal vein and superior mesenteric vein against the tumor with the surgeon’s left hand behind the head of the pancreas. An experienced assistant is needed to divide the neck of the pancreas to the left of the portal vein and achieve proximal and distal control. Sometimes, the vein can be sutured closed with minimal narrowing. Other times, a segmental resection and interposition graft (internal jugular vein) may be needed.

Postoperative hemorrhage can occur from inadequate ligation of any one of numerous blood vessels during the procedure.
Hemorrhage can also occur due to digestion of retroperitoneal blood vessels due to a combined biliary-pancreatic leak. Uncommonly, a stress ulcer, or later, a marginal ulcer, can result in GI hemorrhage. Typically, a vagotomy is not performed when pancreaticoduodenectomy is performed for pancreatic cancer, but patients are placed on proton pump inhibitors.

**Outcome and Value of Pancreaticoduodenectomy for Cancer.** Survival figures indicate that perhaps few patients are cured indefinitely of pancreatic cancer with pancreaticoduodenectomy. This has led to a nihilistic view toward patients with this disease which has further contributed to poor outcomes. Using the National Cancer Data Base (1995–2004), Bilimoria reported on 9559 patients with early stage potentially resectable tumors (pretreatment clinical Stage I: T1N0M0 and T2N0M0).364 Multivariate models were employed to identify factors predicting failure to undergo surgery and assess the impact of pancreatectomy on survival. This study identified a striking underuse of pancreatectomy in the United States. Of clinical stage I patients, 71.4% (6823/9559) did not undergo surgery; 6.4% (616/9559) were excluded due to comorbidities; 4.2% (403/9559) refused surgery; 9.1% (869/9559) were excluded due to age; and 38.2% (3644/9559) with potentially resectable cancers were not offered surgery. Patients were less likely to undergo surgery if they were older than 65 years, were black, were on Medicare or Medicaid, had pancreatic head lesions, earned lower annual incomes, or had less education. Patients were less likely to receive surgery at low-volume and community centers. Patients who were not offered surgery had worse survival than patients who underwent resection. Overall survival from PDAC would significantly increase if more patients with stage I tumors were identified and offered surgery at high-volume centers (Fig. 33-76).

Although pancreaticoduodenectomy may be performed with the hope of the rare cure in mind, the operation more importantly provides better palliation than any other treatment, and it is the only modality that offers any meaningful improvement in survival. If the procedure is performed without major complications, many months of palliation are usually achieved. However, it is the surgeon’s duty to make sure patients and their families have a realistic understanding of the true goals of pancreaticoduodenectomy in the setting of pancreatic cancer.

**Adjuvant Chemotherapy and Radiation.** Small studies in the 1980s suggested that adjuvant chemotherapy with 5-FU combined with radiation improves survival by about 9 months after pancreatic resection for pancreatic adenocarcinoma.365 Subsequent, noncontrolled studies have reinforced that concept; however, the data have been criticized due to the low number of patients and low dose of radiation therapy that was given. In addition, gemcitabine has replaced 5-FU as standard therapy in pancreatic cancer but is thought to be too toxic when given with radiotherapy without dose reduction. A recent large European multicenter trial concluded that there was no value to chemoradiotherapy, although the study suggested the possibility that chemotherapy alone might have survival benefit.366 Randomized trials have failed to resolve the debate regarding the role of adjuvant radiation therapy in resectable pancreas cancer. A reasonable consideration in a disease with high rates of distant metastases is to begin with adjuvant chemotherapy, followed by radiation therapy in patients who do not progress, particularly in patients where there may be increased concern about local recurrence such as in patients with close margins.

Remarkable results in adjuvant therapy were reported by the Virginia Mason Clinic with combination 5-FU, cisplatinum, interferon-α, and external beam radiation.367 Although the toxicity was high (42% hospitalized for GI toxicity), the promising results prompted larger confirmatory studies. Unfortunately, one such study was stopped due to toxicity, and this protocol has not been widely adopted. More recent results with FOLFIRINOX in the setting of metastatic disease have encouraged clinical trials using this regimen in the adjuvant setting, which are currently underway (ClinicalTrials.gov identifier: NCT02172976). Nevertheless, pending further study, it is typical in the United States for patients with acceptable functional status to receive some form of adjuvant chemotherapy and sometimes chemoradiotherapy after surgery.

**Neoadjuvant Treatment.** There are several potential advantages to the use of chemotherapy or chemoradiation before an attempt at surgical resection. For example, it avoids the risk that adjuvant treatment is delayed by complications of surgery. Neoadjuvant treatment also may decrease the tumor burden at operation, increasing the rate of resectability and killing some tumor cells before they can be spread intraoperatively. Another potential advantage is that it allows patients with occult metastatic disease to avoid the morbidity of pancreatic resection. As many as 20% of patients treated with neoadjuvant chemoradiation develop metastatic disease detected by restaging CT and do not go on to surgery. This approach may separate patients into a subset likely to benefit from resection and a subset in whom surgery would be unlikely to provide clinical benefit. Preoperative chemoradiation has been shown not to increase the perioperative morbidity or mortality of pancreaticoduodenectomy. It may even decrease the incidence of pancreatic fistula. Prospective randomized trials investigating this concept are ongoing but are difficult to complete due to the high number of patients who fail to complete or receive a full course of either therapy. Studies have shown that neoadjuvant therapy is associated with a lower rate of lymph node positivity and improved survival is achieved among the patients who do not develop disease progression during neoadjuvant therapy and go on to resection. Neoadjuvant therapy should be considered an acceptable alternative to surgery first followed by adjuvant therapy for resectable pancreatic cancer. Patients should be encouraged to consider available
clinical trials of neoadjuvant therapy for resectable pancreatic cancer. Unfortunately, a recent trial attempting to randomize patients with resectable pancreatic cancer to neoadjuvant versus adjuvant therapy failed to demonstrate an advantage for the neoadjuvant approach. However, the study was inconclusive because it was stopped after 73 of a planned 254 patients due to slow accrual, with only 66 eligible for analysis. Other trials are ongoing. The NEOPA trial (ClinicalTrials.gov identifier: NCT01900327) is a prospectively randomized phase 3 trial of patients receiving neoadjuvant chemoradiation followed by curative surgery vs. primary surgery followed by adjuvant therapy with a primary endpoint of 3-year overall survival.

Most pancreatic surgeons agree that neoadjuvant chemotherapy, and perhaps chemoradiotherapy, should be offered to patients with locally advanced but resectable disease (vein involvement) and to patients with “borderline resectable disease” (abutment of ≤180 degrees of the circumference of the SMA, celiac axis, or hepatic artery or if there is a short segment of vein occlusion, CT findings suspicious for metastatic disease, like 1 mm liver lesions too small to characterize or biopsy, and patients with multiple comorbidities or marginal performance status). This strategy acknowledges the fact that these patients are at high risk of early distant recurrence and/or R1 resection with early local recurrence. Neoadjuvant treatment helps select the right patients for surgery and may reduce the incidence of a margin positive resection.

Postoperative Surveillance. Recurrence after successful resection usually manifests as hepatic metastases. Adjuvant chemotherapy with or without radiation is usually administered for 6 months. During this time period, patients are monitored with frequent physical examinations and laboratory tests, including CA19-9. CT scans are typically ordered every 3 months in the first 2 years after resection or when a rising CA19-9 or new symptoms suggest recurrence. Surgical therapy for recurrent disease is usually reserved only for select patients with limited disease who remain reasonable operative candidates who develop symptomatic gastric outlet or bowel obstruction.

Ampullary and Periampullary Cancer. Ampullary cancers need to be distinguished from periampullary cancers. The ampulla is the junction of the biliary and pancreatic ducts within the duodenum. Periampullary cancer includes tumors arising from the distal bile duct, duodenal mucosa, or pancreas just adjacent to the ampulla, and the ampulla can be overgrown by cancers that arise from these adjacent areas, making it impossible to determine the true site of origin. Clinically, the term periampullary cancer is, therefore, a nonspecific term used to refer to a variety of tumors arising at the intersection of these four sites. The term ampullary cancer is more specific and is reserved for tumors that arise at the ampulla. Based on their location, ampullary cancers are usually detected relatively early due to the appearance of jaundice and have a more favorable prognosis. The ampulla of Vater is lined by an epithelial layer that transitions from pancreatic and biliary ductal epithelium to duodenal mucosal epithelium. Ampullary adenocarcinomas can therefore have an intestinal and/or pancreaticobiliary histologic morphology, with the former having a better prognosis. Patients with ampullary cancer have a 10-year survival of about 35%, which is a much better prognosis than patients with pancreatic adenocarcinoma. The difference in survival is not entirely explained by an earlier presentation and lower incidence of lymph node metastases. There are biologic, particularly molecular, differences between ampullary and pancreatic adenocarcinoma of the pancreas.

Intestinal type ampullary cancers have a lower incidence of EGFr and mutant p53 overexpression, and fewer activating K-ras mutations. These tumors are more likely to have genetic changes similar to colon cancer such as microsatellite instability and adenomatous polyposis coli mutations.

Management of Periampullary Adenomas. Benign tumors such as ampullary adenomas can also originate at the ampulla. The accuracy of endoscopic biopsy in distinguishing ampullary cancer from benign adenoma is poor, with false-negative rates from 25% to 56% even if sphincterotomy precedes the biopsy. However, benign villous adenomas of the ampullary region can be excised locally. This technique is applicable only for small tumors (approximately 2 cm or less) with no evidence of malignancy upon biopsy. EUS may help to accurately determine if there is invasion into the duodenal wall. In the absence of invasion, adenomas may be amenable to an endoscopic or transduodenal excision. A longitudinal duodenotomy is made and the tumor is excised with a 2- to 3-mm margin of normal duodenal mucosa. In some centers, small periampullary adenomas can also be removed endoscopically. A preoperative diagnosis of cancer is a contraindication to transduodenal excision, and pancreaticoduodenectomy should be performed. Likewise, if final pathologic examination of a locally excised tumor reveals invasive cancer, the patient should be returned to the operating room for a pancreaticoduodenectomy. An important subset of patients are those with FAP who develop periampullary or duodenal adenomas. These lesions have a high incidence of harboring carcinoma and frequently recur unless the mucosa at risk is resected. A standard (not pylorus-sparing) Whipple is the procedure of choice in FAP patients with periampullary lesions.

Cystic Neoplasms of the Pancreas. A cystic neoplasm needs to be considered when a patient presents with a fluid-containing pancreatic lesion. Asymptomatic cystic neoplasms of the pancreas may be more frequent than previously recognized and are being identified with increasing frequency as the use of abdominal CT scanning and MRI has increased. Pancreatic cysts are now thought to be present in about 9% of the population age 80 and older. When symptoms are clearly attributable to a pancreatic cyst, resection is indicated in patients who are fit candidates for surgery. However, management of asymptomatic pancreatic cysts is nuanced and can trigger significant anxiety for patients and their surgeons. Invasive surveillance and aggressive surgical intervention can cause harm, decrease quality of life, and increase costs. While the overall risk that an incidental pancreatic cyst is malignant is very low (about 1 in 10,000), the risks of surgery are very significant with a 2% to 5% mortality and 30% to 40% morbidity. However, some of these neoplasms slowly undergo malignant transformation and thus represent an opportunity for surgical cure, which is uncommon after transformation to invasive pancreatic adenocarcinoma. The dilemma for the surgeon is to identify the minority of cysts that pose a significant risk and provide individual patients with an accurate assessment of their unique risk-benefit ratio of resection vs. surveillance.

Surveillance programs are of questionable value in patients who are not candidates for surgery due to age and or multiple comorbidities and limited life expectancy. Surgeons also need to clearly explain to the patient the risks and benefits of surveillance itself. MRI is the preferred surveillance imaging modality.
over computed tomography because MRI does not expose the patient to radiation and better demonstrates the structural relationship between the pancreatic duct and associated cyst. Also, MRI is less invasive than EUS. EUS is therefore reserved for further evaluation of higher risk cysts. Another problem to confront is when to stop surveillance in a cyst that has been stable during observation. Some clinicians stop following after 5 years but data is lacking to guide this decision. In addition, patients who have undergone resection of a pancreatic cystic neoplasm with high grade dysplasia may warrant continued surveillance after surgery particularly when there is a possible field effect such as in IPMN or concern that all of the disease has not been resected. Individualized decision-making and multidisciplinary input is ideal.372

**Pseudocysts.** The most common cystic lesion of the pancreas is the pseudocyst, which, of course, has no epithelial lining and is a nonneoplastic complication of pancreatitis or pancreatic duct injury. As discussed in “Complications of Chronic Pancreatitis,” the diagnosis is usually straightforward from the clinical history. Although not usually necessary, analysis of pseudocyst fluid would reveal a high amylase content. The danger comes in mistaking a cystic pancreatic neoplasm for a pseudocyst and incorrectly draining a cystic neoplasm into the GI tract rather than resecting the neoplasm. For this reason, biopsy of the pseudocyst wall is a requirement in the management of pancreatic pseudocysts.

**Cystadenoma.** Serous cystadenomas are essentially considered benign tumors without malignant potential. Serous cystadenocarcinoma has been reported very rarely (<1%). Therefore, malignant potential should not be used as an argument for surgical resection, and the majority of these lesions can be safely observed in the absence of symptoms due to mass effect or rapid growth. The average rate of growth is about 0.5 cm per year. About 50% of cystadenomas are asymptomatic and detected as an incidental finding. Most symptomatic patients have mild upper abdominal pain, epigastric fullness, or moderate weight loss. Occasionally, cystadenomas can grow to a size capable of producing jaundice or GI obstruction due to mass effect (Fig. 33-77). For symptomatic patients with serous cystadenoma, surgical resection is indicated. For lesions in the tail, splenectomy is not necessary, given the benign nature of the tumor. In appropriate candidates, a laparoscopic approach to distal pancreatectomy with or without splenic preservation can be considered. These cysts are frequently found in older women in which pancreatic resection for a benign neoplasm should be avoided in the absence of significant symptoms. All regions of the pancreas are affected, with half in the head/uncinate process, and half in the neck, body, or tail of the pancreas. They have a spongy appearance, and multiple small cysts (microcystic) are more common than larger cysts (macrocystic or oligocystic). These lesions contain thin serous fluid that does not stain positive for mucin and is low in CEA (<200 ng/mL). Typical imaging characteristics include a well-circumscribed cystic mass, small septations, fluid close to water density, and sometimes, a central scar with calcification. If a conservative management is adopted, it is important to be sure of the diagnosis. EUS-FNA should yield nonviscous fluid with low CEA and amylase levels, and if cells are obtained, which is rare, they are cuboidal and have a clear cytoplasm.

**Mucinous Cystadenoma and Cystadenocarcinoma.** Mucinous cystic neoplasms (MCNs) encompass a spectrum ranging from benign but potentially malignant to carcinoma with a very aggressive behavior (Table 33-23). There is often heterogeneity within the lesions with benign and malignant-appearing regions, making it impossible to exclude malignancy with biopsy. MCNs are commonly seen in perimenopausal women, and about two-thirds are located in the body or tail of the pancreas. Like cystadenomas, most MCNs are now incidental findings identified during imaging performed for other reasons. When symptoms are present, they are usually nonspecific and include upper abdominal discomfort or pain, early satiety, and weight loss. On imaging studies, the cysts have thick walls and do not communicate with the main pancreatic duct (Fig. 33-78). There may be nodules or calcifications within the wall of the cyst. The cysts are lined by tall columnar epithelium that fills the cyst with viscous mucin. The submucosal layer consists of a highly cellular stroma of spindle cells with elongated nuclei similar to the “ovarian stroma,” which is a key pathologic feature distinguishing these lesions. Elevated CEA levels in the fluid (≥200 ng/mL) are consistent with mucinous lesions and may suggest malignant transformation.373 Solid areas may contain atypical cells or invasive cancer, and extensive sampling of the specimen is necessary to accurately predict prognosis.

**Table 33-23**

<table>
<thead>
<tr>
<th>World Health Organization classification of primary tumors of the exocrine pancreas</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Benign</strong></td>
</tr>
<tr>
<td>1. Serous cystadenoma (16%)</td>
</tr>
<tr>
<td>2. Mucinous cystadenoma (45%)</td>
</tr>
<tr>
<td>3. Intraductal papillary-mucinous adenoma (32%)</td>
</tr>
<tr>
<td>4. Mature cystic teratoma</td>
</tr>
<tr>
<td><strong>B. Borderline</strong></td>
</tr>
<tr>
<td>1. Mucinous cyst tumor with moderate dysplasia</td>
</tr>
<tr>
<td>2. Intraductal papillary mucinous tumor with moderate dysplasia</td>
</tr>
<tr>
<td>3. Solid pseudopapillary tumor</td>
</tr>
<tr>
<td><strong>C. Malignant</strong></td>
</tr>
<tr>
<td>1. Ductal adenocarcinoma</td>
</tr>
<tr>
<td>2. Serous/mucinous cystadenocarcinoma (29%)</td>
</tr>
<tr>
<td>3. Intraductal mucinous papillary tumor</td>
</tr>
</tbody>
</table>

Figure 33-77. Mucinous cystic neoplasm in tail of pancreas.
Resection is the treatment of choice for most mucin-producing cystic tumors. Malignancy cannot be ruled out without removal and extensive sampling of the entire tumor. Malignancy has been reported in 6% to 36% of MCNs. Current thinking is that all of these tumors will eventually evolve into cancer if left untreated. However, for individual patients, the surgeon is often faced with a difficult decision. The risk of pancreas surgery in an older patient with multiple comorbidities and a relatively short life expectancy regardless of their pancreatic cyst is frequently weighed against the potential future risk that the cyst, particularly a small cyst, will transform into an incurable invasive cancer.

The utility of detailed DNA analysis of pancreatic cyst fluid to diagnose mucinous and malignant cysts has been evaluated in the PANDA study. The study concluded that cyst fluid K-ras mutation was helpful in the diagnosis of mucinous cysts with 96% specificity. Components of DNA analysis detecting malignant cysts included allelic loss amplitude over 82% and high DNA amount. The criteria of high amplitude K-ras mutation followed by allelic loss showed maximum specificity (96%) for malignancy. However, this test lacks sensitivity. In clinical practice, the surgeon must take all of these complementary factors into consideration when determining the malignant potential of a pancreatic cystic neoplasm.

Because most MCNs are located in the body and tail of the pancreas, distal pancreatectomy is the most common treatment. For small lesions, it may be appropriate to preserve the spleen, but splenectomy ensures removal of the lymph node basin that can potentially be involved. It is very important not to rupture the cyst during resection, and the tumor should be removed intact, not morselized. Therefore, a laparoscopic approach may not be appropriate for larger lesions. Completely resected MCNs without atypia are usually cured especially if small (<3 cm). Even patients with moderate dysplasia or carcinoma in situ are usually cured by complete resection. Noninvasive MCNs require no surveillance after resection. For MCNs with an associated invasive carcinoma, prognosis depends on the extent of the invasive component, tumor stage, and resectability. The 2-year survival rate and 5-year survival rate of patients with
Intraductal Papillary Mucinous Neoplasm. Intraductal papillary mucinous neoplasms (IPMNs) usually occur within the head of the pancreas and arise within the pancreatic ducts. The ductal epithelium forms a papillary projection into the duct, and mucin production causes intraluminal cystic dilation of the pancreatic ducts (Fig. 33-79). Imaging studies demonstrate diffuse dilation of the pancreatic duct, and the pancreatic parenchyma is often atrophic due to chronic duct obstruction. However, classic features of chronic pancreatitis, such as calcification and a beaded appearance of the duct, are not present. At ERCP, mucin can be seen extruding from the ampulla of Vater, a so-called fish-eye lesion that is virtually diagnostic of IPMN (Fig. 33-80). Initial reports suggested a male predominance, but more recent series indicate an equal distribution. Patients are usually in their seventh to eighth decade of life and present with abdominal pain or recurrent pancreatitis, thought to be caused by obstruction of the pancreatic duct by thick mucin. Some patients (5–10%) have steatorrhea, diabetes, and weight loss secondary to pancreatic insufficiency.

Some IPMNs primarily involve the main pancreatic duct, while others involve the branch ducts. The mean frequency of malignancy in main duct IPMN (MD-IPMN) is 62% (Fig. 33-81). Considering this high incidence of malignant lesions and the low 5-year survival rates (31–54%), international consensus guidelines recommend resection for all surgically fit patients with MD-IPMN.\textsuperscript{372} If the margin is positive for high-grade dysplasia, additional resection should be attempted to obtain at least moderate-grade dysplasia at the surgical margin.

The surgical management of IPMNs is complicated by the fact that the lesion itself is small and preoperative imaging studies show a dilated pancreatic duct but not necessarily the mass. Mucus can dilate the duct proximal and distal to the lesion. Furthermore, these lesions can spread microscopically along the duct, and there can be skip areas of normal duct between the diseased portions. Therefore, thorough preoperative imaging including EUS, MRCP, or ERCP, and sometimes pancreatic ductoscopy, which can also be repeated intraoperatively, is useful (see Fig. 33-80). The surgeon needs to be prepared to extend the resection, if necessary, based on intraoperative findings and frozen section of the margin. Extending the resection to the point of total pancreatectomy is controversial due to the morbidity of this operation. Like MCNs, the IPMNs require careful histologic examination of the entire specimen for an invasive cancer (Fig. 33-82).

Survival of patients with IPMN, even when malignant and invasive, can be quite good. As with MCN, patients with borderline tumors or carcinoma in situ are usually cured. For this reason, if recurrence occurs in the remaining pancreas, further resection is warranted because several series have shown that some of these cases are salvageable. Patients with IPMN are also at risk for other malignancies and should undergo colonoscopy and close surveillance.

Branch-duct type IPMNs (BD-IPMN) are often found in the uncinate process, are sometimes asymptomatic, and are less frequently associated with malignant transformation (6–46%). Asymptomatic, small suspected BD-IPMNs are frequently
Figure 33-81. Intraductal papillary mucinous neoplasm (IPMN). A. Examples of “fish-eye deformity” of IPMN. Mucin is seen extruding from the ampulla. B. Mucin coming from pancreatic duct when neck of pancreas is transected during Whipple procedure (left). Intraoperative pancreatic ductoscopy to assess the pancreatic tail (right). C. Views of pancreatic duct during ductoscopy; normal (left) and IPMN (right).
observed with serial imaging. High-risk features such as mural nodules, a dilated main duct, positive cytology or cyst fluid CEA >200 need to be ruled out. In the absence of these features, continued observation with serial imaging is appropriate, especially in patients who are not ideal operative candidates. The mean frequency of invasive cancer in resected BD-IPMN is 18%. BD-IPMN mostly occurs in elderly patients, and the annual malignancy rate is only 2% to 3%. These factors support conservative management with follow-up in patients who do not have any symptoms or risk factors predicting malignancy such as mural nodule, rapidly increasing cyst size, and high-grade atypia in cytology. There is insufficient data to support immediate resection for all BD-IPMNs <3 cm without “high-risk stigmata” and “worrisome features.” Branch-duct IPMNs ≥3 cm should be resected.372

Four histologic subtypes of IPMNs have been characterized: gastric, intestinal, pancreaticobiliary, and oncocytic. Most of BD-IPMNs are composed of gastric-type epithelium. However, intestinal type is more common in MD-IPMN. In a recent report, the four subtypes of IPMNs were associated with significant differences in survival.376 Patients with gastric-type IPMN had the best prognosis, whereas those with intestinal and pancreaticobiliary type had a bad prognosis.

**Figure 33-82.** Operative specimen of pancreas with multifocal intraductal papillary mucinous neoplasms (black arrow) and a focus of invasive adenocarcinoma (white arrow). (Reproduced with permission from Asiyanbola B, Andersen DK. IPMN. Editorial Update. accesssurgery.com McGraw-Hill Education; 2008.)

Workup of Asymptomatic Pancreatic Cystic Neoplasms

Incidentally discovered asymptomatic pancreatic cystic neoplasms are evaluated by MRI with MRCP to check for “high-risk stigmata or worrisome features.” An enhancing solid component or main pancreatic duct (MPD) dilation ≥10 mm are considered “high-risk stigmata.” Cysts ≥3 cm, thickened enhancing cyst walls, nonenhancing mural nodules, MPD size of 5 to 9 mm, abrupt change in the MPD caliber with distal pancreatic atrophy, and lymphadenopathy are considered “worrisome features.” All cysts with high risk stigmata are resected. All cysts with worrisome features and cysts >3 cm without worrisome features are further evaluated with EUS. If EUS shows a definite mural nodule, main duct features suspicious for involvement or the cytology is positive or suspicious for malignancy, then resection is recommended. In the absence of worrisome features, no further initial work-up is recommended, but surveillance is still required.

The interval between surveillance imaging is based on the size of the cyst with a lengthening of the interval once stability is established. Patients with noninvasive MCNs require no surveillance after resection, but patients with IPMNs need surveillance after resection. In the absence of residual lesions, repeat MRI at 2 and 5 years may be reasonable. If there is low or moderate-grade dysplasia at the margin, MRI every 6 months is recommended.

Resection is indicated in all surgically fit patients with MD-IPMN, and additional resection is indicated if there is high-grade dysplasia at the margin. Resection is also indicated in all surgically fit patients with MCN. BD-IPMN <3 cm without worrisome features or high-risk stigmata can undergo surveillance.

When patients are deemed unacceptable for resection due to comorbidities, ablation of the cyst can be considered. However, at this time, cyst ablation is considered experimental and should be done as part of a clinical trial. Gastroenterologists have limited experience with injection of a cytotoxic agent into the cyst in an attempt to ablate the cyst epithelium. Ethanol has been used and more recently this has been combined with paclitaxel. The combination of ethanol and paclitaxel injection resulted in elimination of the cysts, as determined by CT scanning, in 29 out of 47 (62%) of patients in a median follow-up period of 21.7 months.377

Recently, the American Gastroenterological Association (AGA) published guidelines for asymptomatic mucinous cysts (http://www.gastro.org/guidelines/pancreatic-cysts) that are different from all previously published guidelines in the following areas: 2-year interval for cyst of any size undergoing surveillance, stopping surveillance after 5 years if no change, surgery only if more than one concerning feature on MRI confirmed on EUS and only in centers with high volumes of pancreatic surgery, and no surveillance after surgery if no invasive cancer or dysplasia. Although based on extensive literature review and synthesis, these recommendations have resulted in significant controversy because, in an effort to reduce the costs of health care delivery and perhaps decrease inadvertent harm to patients, they advocate less frequent follow-up and a higher threshold before offering EUS and/or surgery.

**Solid-Pseudopapillary Tumor.** Solid-pseudopapillary tumors are rare and typically occur in young women. Previous names for this entity include, solid and cystic, solid and papillary, cystic and papillary, and papillary-cystic tumor. They are typically well circumscribed on CT (Fig. 33-83). The cysts are not true epithelial-lined cysts but rather represent a necrotic/degenerative process. Histology may be similar to neuroendocrine tumors, but they do not stain positive for neuroendocrine markers such as chromogranin. Most are cured by resection, but liver and peritoneal metastases have been reported.

**Other Cystic Neoplasms.** Rarely, typical ductal adenocarcinoma of the pancreas may undergo cystic degeneration due to central necrosis. Occasionally, this will create difficulty in the proper preoperative diagnosis and should be kept in mind when deciding to conservatively follow a cystic pancreatic neoplasm. It is more common, 5% to 10%, for neuroendocrine tumors of the pancreas to contain cysts. These cysts are filled with serosanguineous fluid rather than necrotic debris. Lymphoepithelial
cysts of the pancreas usually occur in men in their fifth to sixth decade. These benign lesions may be unilocular or multilocular and vary widely in size. The contents of the cyst are also variable and may be thin serous fluid or cheesy/caseous material if there is increased keratin formation. A substantial number of patients with von Hippel-Lindau syndrome develop pancreatic cysts that resemble serous cystadenomas. There may be multiple lesions scattered throughout the pancreas. Patients with polycystic kidney and hepatic disease may also develop benign pancreatic cysts (cystadenomas). With all of these rare cystic neoplasms, careful clinical history, high-quality pancreatic imaging, and sampling of the cyst fluid for analysis will guide proper treatment.

Pancreatic Lymphoma. Lymphoma can affect the pancreas. Primary involvement of the pancreas with no disease outside the pancreas also occurs. The clinical presentation often is similar to pancreatic adenocarcinoma, with vague abdominal pain and weight loss. Identification of a large mass often involving the head and body of the pancreas should raise suspicion. Percutaneous or EUS-guided biopsy will confirm the diagnosis in most cases. If the diagnosis cannot be confirmed preoperatively, laparoscopic exploration and biopsy are indicated. There is no role for resection in the management of pancreatic lymphoma. Endoscopic stenting to relieve jaundice followed by chemotherapy is the standard treatment, and long-term remission is often achieved.

REFERENCES

Entries highlighted in bright blue are key references.


Hofbauer B, Saluja AK, Bhatia M, et al. Effect of recombinant platelet-activating factor acetylhydrolase on two models...


HISTORICAL BACKGROUND

The spleen has been the subject of man’s musings since almost the establishment of the written word. It has been largely misunderstood, often maligned, and certainly underappreciated as a major organ for more than two millennia. The ancients, presumably through patient observation and occasional anatomic exploration, often consigned the spleen to a vestigial role.

Our current understanding of the central role played by the spleen in regulating the immune system and influencing metabolic and endocrine functions has been built upon knowledge gleaned only over the past few decades. Our early notions of the spleen as a dispensable filter of blood or seat of emotion have been dispelled as our understanding of its structure and function has evolved, informing our surgical approach to this hallowed measure of influence of surgical publication, the impact factor, which is most commonly calculated using a 2-year frame of reference.5,6

In the early 17th century, several physician scientists, Malpighi being the most prominent, began testing hypotheses on splenic function by splenectomizing dogs. He reportedly followed several dogs 5 years postoperatively, noting their healthy survival though apparent ravenous hunger and enhanced sexual appetites. The spleen, still in the era of “balanced humors,” was thus felt to play a role in balancing various appetites as well. In addition to melancholy, the spleen became associated with anger and, paradoxically, was also seen as the “seat of laughter.”2

The claim for the first human splenectomy may have predated that of canine splenectomy. Andriano Zaccavello was credited in 1549 with having performed a splenectomy on a middle-aged woman. This claim remains shrouded in controversy, and the indication for the surgery and whether in fact splenectomy was performed have been called into question. The patient apparently survived, but may in fact have undergone resection of an ovarian cyst rather than her spleen.4 Most patients who underwent splenectomy in the three centuries that followed fared badly. The vast majority of splenectomies performed were partial. Most of these patients required surgery for

Galen, in the second century AD, engaged in more serious anatomic investigation espousing the early belief that function followed structure. His investigations, though pioneering, lacked sufficient rigor, evidenced by his contention that black bile or melancholia flowed from the liver to the spleen and then through the short gastric vessels into the stomach to be excreted. The influence of Galen’s teaching endured for more than 1200 years. All the more remarkable considering the current hallowed measure of influence of surgical publication, the impact factor, which is most commonly calculated using a 2-year frame of reference.5,6

In the early 17th century, several physician scientists, Malpighi being the most prominent, began testing hypotheses on splenic function by splenectomizing dogs. He reportedly followed several dogs 5 years postoperatively, noting their healthy survival though apparent ravenous hunger and enhanced sexual appetites. The spleen, still in the era of “balanced humors,” was thus felt to play a role in balancing various appetites as well. In addition to melancholy, the spleen became associated with anger and, paradoxically, was also seen as the “seat of laughter.”2

The claim for the first human splenectomy may have predated that of canine splenectomy. Andriano Zaccavello was credited in 1549 with having performed a splenectomy on a middle-aged woman. This claim remains shrouded in controversy, and the indication for the surgery and whether in fact splenectomy was performed have been called into question. The patient apparently survived, but may in fact have undergone resection of an ovarian cyst rather than her spleen.4 Most patients who underwent splenectomy in the three centuries that followed fared badly. The vast majority of splenectomies performed were partial. Most of these patients required surgery for
left upper quadrant stab wounds sustained in battles or duels resulting in partial or complete splenic prolapse.8

It was in the early 18th century that the growing body of anatomic microstructural knowledge began to turn the tide on the long-held theory of health and disease deriving from a balance of the four humors. William Henson believed the spleen to be a ductless vascular gland similar to the thyroid and adrenals. In 1777, he wrote of the lymphatic nature of the spleen and its filtering function and even suggested its role in hematopoiesis.10

Rudolf Virchow, one of the first to discover leukemia, implicated the spleen as figuring prominently in all leukemia patients. He suspected that the spleen was responsible for generating the leukocytes in large quantities in these patients. There soon followed an enthusiastic effort by surgeons to cure leukemia by splenectomy. Dr. Thomas Bryant performed the first splenectomy in 1866 in a patient with leukemia. The patient died, as did all 14 patients who underwent splenectomy for leukemia over the next 15 years. After his second consecutive mortality in this setting, Bryant declared that “the operation is physiologically unsound & surgically unsafe for leukemia and should not be performed.”11 In 1908, Johnson reported a series of 99 splenectomies for leukemia with an 85% mortality rate. Unfortunately, it took several decades for his words to be heeded.

In 1916, a medical student from Prague named Paul Kaznelson wrote on the key role played by the spleen in the destruction of platelets leading to the first reported (and successful) splenectomy for a patient with idiopathic thrombocytopenia purpura.2

As surgeons’ experience with the procedure grew, the associated morbidity and mortality decreased. By 1920, the Mayo Clinic reported that splenectomy had a reduced mortality rate of 11%.1

O’Donnell in 1929 was the first to describe fatal post-splenectomy sepsis in a child who had undergone the surgery for hemolytic anemia.3 It took Springer’s 1973 review of almost 2800 postsplenectomy patients and the 2.5% incidence of sepsis-induced mortality (vs. 0.01% in the general population) to reorient surgeons to more conservative splenic procedures.2,3

The advent of minimally invasive surgery and laparoscopic splenectomy in the early 1990s represented a clear advance, benefitting the patient through this evolution of surgical technique. Most large series of laparoscopic splenectomy for benign and malignant indication now report a mortality rate of <1%.12,13 As even more contemporary research reveals the spleen to play a central role in immune, metabolic, and endocrine function, it follows that the surgeon’s role going forward will be to preserve this organ and its functions whenever possible.

EMBRYOLOGY AND ANATOMY

Consisting of an encapsulated mass of vascular and lymphoid tissue, the spleen is the largest reticuloendothelial organ in the body. Arising from the primitive mesoderm as an outgrowth of the left side of the dorsal mesogastrium, by the fifth week of gestation, the spleen is evident in an embryo 8 mm long.

Development begins through the formation of the splanchnic mesodermal plate, derived from the mesoderm, at embryonic day 12. The embryonic spleen is first colonized by erythroid and myeloid progenitor cells at 2 weeks of gestation. Following soon thereafter, the hematopoietic stem cells take up residence in the forming spleen.14 The spleen assumes an important hematopoietic role until the fifth month of gestation. After birth, splenic erythropoietic function may persist in some hematologic disorders.15
The organ continues its differentiation and migration to the left upper quadrant, where it comes to rest with its smooth, diaphragmatic surface facing posteroinferiorly.\textsuperscript{16}

The most common anomaly of splenic embryology is the accessory spleen. Present in up to 20\% of the population, one or more accessory spleens may also occur in up to 30\% of patients with hematologic disease. Over 80\% of accessory spleens are found in the region of the splenic hilum and vascular pedicle. Other locations for accessory spleens in descending order of frequency are the gastrocolic ligament, the pancreas tail, the greater omentum, the stomach’s greater curve, the splenocolic ligament, the small and large bowel mesentery, the left broad ligament in women, and the left spermatic cord in men (Fig. 34-1).\textsuperscript{10,16}

The abdominal surface of the diaphragm separates the spleen from the lower left lung and pleura and the ninth to eleventh ribs. The visceral surface faces the abdominal cavity and contains gastric, colic, renal, and pancreatic impressions. Spleen size and weight vary with age, with both diminishing in the elderly and in those with underlying pathologic conditions. The average adult spleen is 7 to 11 cm in length and weighs 150 g (range, 70–250 g).

The spleen’s superior border separates the diaphragmatic surface from the gastric impression of the visceral surface and often contains one or two notches, which are particularly pronounced when the spleen is greatly enlarged.

Of particular clinical relevance, the spleen is suspended in position by several ligaments and peritoneal folds to the colon (splenocolic ligament), the stomach (gastrosplenic ligament), the diaphragm (phrenosplenic ligament), and the kidney, adrenal gland, and tail of the pancreas (splenorenal ligament) (Fig. 34-2). In a related historical footnote it was widely held less than 200 years ago that a “wandering spleen” led women to experience hypochondria. Dietl in 1863 finally clarified that “it was not a patient’s temperament but rather relaxation, extension or the hypoplasia of splenic ligaments that made a spleen wander.”\textsuperscript{17} The gastrosplenic ligament contains the short gastric vessels; the remaining ligaments are avascular, with rare exceptions, such as in patients with portal hypertension. The relationship of the pancreas to the spleen also has important clinical implications. In cadaveric anatomic series, the tail of the pancreas has been demonstrated to lie within 1 cm of the splenic hilum 75\% of the time and to actually abut the spleen in 30\% of patients.\textsuperscript{2}

The spleen derives most of its blood from the splenic artery, the longest and most tortuous of the three main branches of the celiac artery. The splenic artery can be characterized by the pattern of its terminal branches. The distributed type of splenic artery is the most common (70\%) and is distinguished by a short trunk with many long branches entering over three-fourths of the spleen’s medial surface. The less common magistral type of splenic artery (30\%) has a long main trunk dividing near the hilum into short terminal branches, and these enter over 25\% to 30\% of the spleen’s medial surface. The spleen also receives some of its blood supply from the short gastric vessels that branch from the left gastroepiploic artery running within the gastrosplenic ligament. The splenic vein joins the superior mesenteric vein to form the portal vein and accommodates the major venous drainage of the spleen.

When a normal, freshly excised spleen is sectioned, the cut surface is finely granular and predominantly dark red with whitish nodules distributed liberally across its expanse. This gross observation reflects the spleen’s microstructure. The splenic parenchyma is composed of two main elements: the red pulp, constituting approximately 75\% of total splenic volume, and the white pulp (Fig. 34-3). At the interface between the red and white pulp is the narrow marginal zone.
Blood enters the red pulp through cords comprised of fibroblasts and reticular fibers, which contain many macrophages and lack an endothelial lining. The blood then passes from these “open” cords to venous sinuses, which are surrounded and separated by the same reticulum, and ultimately drains into tributaries of the splenic vein. An understanding of the microanatomy of these sinuses has elucidated the mechanical filtration function of the spleen. Unlike the cords of the red pulp, the sinuses of the red pulp are lined by endothelial cells. These cells contain unique stress fibers that connect the endothelial cells and that contain actin and myosin-like filaments capable of producing a sliding action. When activated, these filaments can create slits or gaps between the endothelial cells through which blood can then pass from the cords. Aging erythrocytes with stiffer membranes get stuck trying to pass into the sinus and are phagocytized by macrophages within the red pulp.

The red pulp thus serves as a dynamic filtration system, enabling macrophages to remove microorganisms, cellular debris, antigen-antibody complexes, and senescent erythrocytes from the circulation.

Around the terminal millimeters of splenic arterioles, a periaortic lymphatic sheath replaces the native adventitia of the vessel. The sheath is comprised of T lymphocytes and intermittent aggregations of B lymphocytes or lymphoid follicles. When antigenically stimulated, the follicles, serving as centers of lymphocyte proliferation, develop germinal centers, which regress as the stimulus or infection subsides. This white pulp consists of nodules that normally are ≤1 mm in size but can increase to several centimeters when nodules coalesce, as occurs in certain lymphoproliferative disorders. At the junction between the white and red pulp is the marginal zone, where lymphocytes are more loosely aggregated.

As well as serving as a transit area, the marginal zone is home to its own unique population of cells. Notably two specific types of macrophages reside there, marginal zone macrophages and marginal zone metallophilic macrophages. The former play an important role in the targeting and clearance of certain bacterial pathogens. The latter have been shown to be the main producers of interferons A and B in response to a viral challenge.

**PHYSIOLOGY AND PATHOPHYSIOLOGY**

The spleen is contained by a 1- to 2-mm thick capsule. In humans, the capsule is rich in collagen and contains some elastin fibers. Many mammals have splenic capsules and trabeculae with abundant smooth muscle cells, which upon autonomic stimulation contract to expel large volumes of stored blood into the general circulation. The human splenic capsule and trabeculae, by contrast, contain few or no smooth muscle cells.

Total splenic inflow of blood is approximately 250 to 300 mL/min. Blood flows through successively tapering arteries to arterioles, traverses the white pulp, crosses the marginal zone, and enters the red pulp. From that entry, the flow rate through the spleen may vary greatly. Animal studies measuring the transit times of isotopically labeled blood through the spleen have revealed three distinct velocities of flow. Humans have long been recognized to have both a fast or closed circulation—with blood passing directly from arterioles into venous sinuses—and a slower or open circulation. Most of the spleen’s filtration function occurs via the slower circulation. During open circulation, blood percolates through the reticular space and splenic cords, thus gaining access through gaps or slits in the endothelial cell lining to the sinuses as previously described. Flowing into and out of the venous sinuses through these gaps, the blood is exposed to extensive contact with splenic macrophages. These are responsible for the innate immune response of the spleen, which occurs largely within the marginal zone. The white pulp, by contrast, is involved only in adaptive immunity. In addition, because the passage of plasma through these spaces does not slow in a similar manner, a temporary and unique adhesive contact between blood cells and components of the splenic cord may occur. That there is a selective slowing of blood cell flow versus plasma flow is further evidenced by the fact that within the spleen, the erythrocyte concentration (hematocrit) is twice that of the general circulation. During this contact with splenic macrophages, it is likely that the removal of both cellular debris and senescent blood cells occurs.

The process by which the spleen removes erythrocyte inclusions, such as Heinz bodies (intracellular altered hemoglobin), without cell lysis while red blood cells travel through the spleen is not well understood. The spleen acts as the major site for clearance from the blood of damaged or aged red blood cells and, in addition, has a part in the removal of abnormal white blood cells and platelets. A minimum of 2 days of the erythrocyte’s 120-day life cycle is spent sequestered in the spleen. Daily, approximately 20 mL of aged red blood cells are removed. Evidence suggests that, as erythrocytes age, previously undetected antigens on their surfaces may attach to autoantibodies in the circulation; then macrophages may bind to the antibodies and initiate phagocytosis. It is probable that the erythrocyte is damaged over time by multiple passages through the spleen as well as delayed transit through the congested and relatively hypoxic and acidic environment of the splenic cords.

The spleen can also serve as an extra medullary site for hematopoiesis, if required. Another role played by the spleen is in recycling iron. Erythrocytes in large numbers are destroyed intravascularly throughout the body. The released hemoglobin is then bound to haptoglobin, which is ultimately scavenged from the circulation in the spleen.
The spleen plays a vital, although not indispensable, role in host defense evidenced by the healthy survival of splenectomized patients. Both innate and adaptive immune responses (historically categorized as cell-mediated and humoral immunity) occur within the spleen.

In addition to the previously noted activities of the marginal zone macrophages, marginal zone B cells serve to detect circulating pathogens and respond quickly to either differentiate into immunoglobulin M (IgM)–producing plasma cells or to function as antigen-presenting cells (APCs), which facilitate pathogen removal and destruction.

It is APC entry in the white pulp in particular that is key to the initiation of the adaptive immune response. Antigens are thus presented to immunocompetent centers within the lymphoid follicles. This gives rise to the elaboration of immunoglobulins (predominantly IgM). After an antigen challenge, such an acute IgM response results in the release of opsonic antibodies from the white pulp of the spleen. Antigen clearance is then facilitated by the splenic and hepatic reticuloendothelial systems.

The structure and immunophysiology of the white pulp is very similar to that of lymph nodes, with the notable difference being that material enters the lymph node in the lymph whereas it is delivered to the white pulp in the blood.20

The spleen also produces opsonins, tuftsin, and properdin. Circulating monocytes are converted within the red pulp into fixed macrophages that account for the spleen’s remarkable phagocytic activity.

The spleen also appears to be a major source of the protein properdin, important in the initiation of the alternate pathway of complement activation. The splenic reticuloendothelial system is better able to clear bacteria that are poorly or inadequately opsonized from the circulation than is the hepatic reticuloendothelial system. Encapsulated bacteria generally fit such a profile, hence the risk posed by pneumococcus and Haemophilus influenzae to an asplenic patient. There appears to be sufficient physiologic capacity within the complement cascade to withstand the loss of tuftsin and properdin production without an increase in patient vulnerability after splenectomy.21-23

In patients with chronic hemolytic disorders, splenic tissue may become permanently hypertrophied. The reticular spaces of the red pulp become distended with macrophages engorged with the products of erythrocyte breakdown, and splenomegaly can result. It is important to distinguish between splenomegaly and hypersplenism, two similar but distinct terms that are critical to understand when discussing splenic pathology. Splenomegaly refers simply to abnormal enlargement of the spleen. Splenomegaly is described variably within the surgical literature as moderate, massive, and hyper, which reflects a lack of consensus. Most would agree, however, that splenomegaly applies to organs weighing ≥500 g and/or averaging ≥15 cm in length.

Massive splenomegaly similarly lacks a consensus definition but has been described variably as spleens >1 kg in mass or >22 cm in length (Fig. 34-4).3 Spleens palpable below the left costal margin are thought to be at least double normal size, with an estimated weight of ≥750 g.24

There is not a single, universally accepted standard, but most would agree that an ex vivo mass of >1 kg or a pole-to-pole length of >15 cm generally qualifies as splenomegaly. Hypersplenism often is found in association with splenomegaly but is not synonymous with it. Hypersplenism is defined as the presence of one or more cytopenias in the context of a normally functioning bone marrow.

Figure 34-4. Splenomegaly. A. Computed tomography (CT) scan. B. Three-dimensional reconstruction of CT scan. C. Postoperative specimen.
Disorders causing hypersplenism can be categorized as either (a) those in which increased destruction of abnormal blood cells occurs in an intrinsically normal spleen (e.g., hemolytic anemias) or (b) primary disorders of the spleen resulting in increased sequestration and destruction of normal blood cells (e.g., infiltrative disorders).

The life cycles of cellular elements vary widely in human blood. A neutrophil in circulation has a normal half-life of approximately 6 hours. The spleen’s role in the normal clearance of neutrophils is not well established. It is clear that hypersplenism may result in neutropenia through sequestration of normal white blood cells or the removal of abnormal ones. Platelets, on the other hand, generally survive in the circulation for 10 days. Under normal circumstances, a third of the total platelet pool is sequestered in the spleen. Thrombocytopenia may result from excessive sequestration of platelets as well as accelerated platelet destruction in the spleen. Splenomegaly may result in sequestration of up to 80% of the platelet pool. The spleen may also contribute to the immunologic alteration of platelets, which leads to thrombocytopenia in the absence of splenomegaly (e.g., idiopathic thrombocytopenic purpura [ITP]).

The immunologic functions of the spleen are consistent with those of other lymphoid organs. It is a site of bloodborne antigen presentation and the initiation of T- and B-lymphocyte activities involved in humoral and cellular immune responses. Alteration of splenic immune function often gives rise to antibody production, which results in blood cell destruction.

Although the spleen contributes to the process of erythrocyte maturation, in adult humans there is little evidence of normal hematopoietic function. The spleen does have a minor role in hematopoiesis in the fourth month in the human fetus, and reactivation can occur in childhood if the bone marrow fails to meet the hematologic needs. Splenic hematopoiesis giving rise to abnormal red blood cells is seen in adults with myeloproliferative disorders. In addition, in response to some anemias, elements of the red pulp may revert to hematopoiesis.

### INDICATIONS FOR SPLENECTOMY

Generally speaking, splenectomy is performed for the purposes of cure or palliation of hematological disease including conditions of hypersplenism, to relieve the mass effect and symptoms associated with splenomegaly, to control infection or hemorrhage and finally to diagnose splenic pathology. For the purposes of this chapter, we will divide these indications into the following categories: (a) benign conditions, including red blood cell disorders, hemoglobinopathies, and platelet disorders (Table 34-1a); (b) malignant conditions, including white blood cell disorders, bone marrow disorders (myeloproliferative disorders) and tumors of the spleen (Table 34-1b); and (c) miscellaneous conditions and lesions of the spleen including infections and abscesses, cysts, vascular anomalies, and more (Table 34-1c). We will conclude this chapter with a brief discussion on splenic salvage.

Overall, the most common indication for splenectomy is trauma to the spleen, whether external trauma (blunt or penetrating) or iatrogenic injury (e.g., at the time of other operations). Inadvertent intraoperative injury to the spleen, necessitating removal, also called “incidental splenectomy” is discussed in a

<table>
<thead>
<tr>
<th>DISEASE/CONDITION</th>
<th>INDICATIONS FOR SPLENECTOMY</th>
<th>RESPONSE TO SPLENECTOMY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Essential thrombocytopenia</td>
<td>Only for advanced disease (i.e., transformation to myeloid metaplasia or AML) with severe symptomatic splenomegaly</td>
<td>Relief of abdominal pain and early satiety</td>
</tr>
<tr>
<td>Glucose-6-phosphate dehydrogenase deficiency (G6PD)</td>
<td>Excessive transfusion requirements; failure of medical therapy (controversial)</td>
<td>May be curative</td>
</tr>
<tr>
<td>Hereditary spherocytosis</td>
<td>Hemolytic anemia, recurrent transfusions, intractable leg ulcers</td>
<td>Improves or eliminates anemia</td>
</tr>
<tr>
<td>Idiopathic thrombocytopenic purpura (ITP)</td>
<td>Failure of medical therapy, recurrent disease</td>
<td>75%–85% rate of long-term response</td>
</tr>
<tr>
<td>Polycythemia vera</td>
<td>Only for advanced disease (i.e., transformation to myeloid metaplasia or AML) with severe symptomatic splenomegaly</td>
<td>Relief of abdominal pain and early satiety</td>
</tr>
<tr>
<td>Pyruvate kinase deficiency</td>
<td>Only in severe cases, recurrent transfusions</td>
<td>Decreased transfusion requirement, palliative only</td>
</tr>
<tr>
<td>Sickle cell disease</td>
<td>History of acute sequestration crisis, splenic symptoms, or infarction (consider concomitant cholecystectomy)</td>
<td>Palliative, variable response</td>
</tr>
<tr>
<td>Thalassemia</td>
<td>Excessive transfusion requirements, symptomatic splenomegaly, infarction, or hypersplenism</td>
<td>Diminished transfusion requirements, relief of symptoms</td>
</tr>
<tr>
<td>Thrombotic thrombocytopenic purpura (TTP)</td>
<td>Excessive plasma exchange requirement</td>
<td>Typically curative</td>
</tr>
<tr>
<td>Warm-antibody autoimmune hemolytic anemia</td>
<td>Failure of medical (steroid) therapy</td>
<td>60%–80% response rate, recurrences common</td>
</tr>
</tbody>
</table>
### Table 34-1b
Indications for and expected response to splenectomy in various malignant diseases, including white blood cell disorders, myeloproliferative disorders, and nonhematologic tumors of the spleen

<table>
<thead>
<tr>
<th>DISEASE/CONDITION</th>
<th>INDICATIONS FOR SPLENECTOMY</th>
<th>RESPONSE TO SPLENECTOMY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute myeloid leukemia (AML)</td>
<td>Intolerable symptomatic splenomegaly</td>
<td>Relief of abdominal pain and early satiety</td>
</tr>
<tr>
<td>Chronic lymphocytic leukemia (CLL)</td>
<td>Cytopenias and anemia</td>
<td>75% response rate</td>
</tr>
<tr>
<td>Chronic myelogenous leukemia (CML)</td>
<td>Symptomatic splenomegaly</td>
<td>Relief of abdominal pain and early satiety</td>
</tr>
<tr>
<td>Chronic myelomonocytic leukemia (CMML)</td>
<td>Symptomatic splenomegaly</td>
<td>Relief of abdominal pain and early satiety</td>
</tr>
<tr>
<td>Myelofibrosis (agnogenic myeloid metaplasia)</td>
<td>Severe symptomatic splenomegaly</td>
<td>76% clinical response at 1 y, high risk of hemorrhagic, thrombotic, and infectious complications (26%)</td>
</tr>
<tr>
<td>Hairy cell leukemia</td>
<td>Severe symptomatic splenomegaly, severe transfusion requirements; failure of medical therapy</td>
<td>Curative</td>
</tr>
<tr>
<td>Hodgkin’s lymphoma</td>
<td>Surgical staging in selected cases</td>
<td>Varied</td>
</tr>
<tr>
<td>Non-Hodgkin’s lymphoma</td>
<td>Cytopenias, symptomatic splenomegaly</td>
<td>Improved complete blood count values, relief of symptoms</td>
</tr>
<tr>
<td>Metastatic tumor of the spleen</td>
<td>If symptomatic, or as part of cancer treatment</td>
<td>Varied</td>
</tr>
<tr>
<td>Primary tumor of the spleen</td>
<td>For diagnosis and treatment of cancer</td>
<td>Varied</td>
</tr>
</tbody>
</table>

### Table 34-1c
Indications for and expected response to splenectomy in various miscellaneous conditions

<table>
<thead>
<tr>
<th>DISEASE/CONDITION</th>
<th>INDICATIONS FOR SPLENECTOMY</th>
<th>RESPONSE TO SPLENECTOMY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abscess of the spleen</td>
<td>Multiloculated, or failure of conservative measures for unilocular</td>
<td>Curative</td>
</tr>
<tr>
<td>Amyloidosis</td>
<td>Symptomatic splenomegaly</td>
<td>Improves symptoms; does not correct underlying disease</td>
</tr>
<tr>
<td>Felty’s syndrome</td>
<td>Neutropenia</td>
<td>80% durable response rate</td>
</tr>
<tr>
<td>Gaucher’s disease</td>
<td>Hypersplenism</td>
<td>Improves cytopenias; does not correct underlying disease</td>
</tr>
<tr>
<td>Niemann-Pick disease</td>
<td>Symptomatic splenomegaly</td>
<td>Improves symptoms; does not correct underlying disease</td>
</tr>
<tr>
<td>Portal/sinistral hypertension</td>
<td>Splenic vein thrombosis, symptomatic splenomegaly</td>
<td>Palliative</td>
</tr>
<tr>
<td>Sarcoidosis</td>
<td>Hypersplenism or symptomatic splenomegaly</td>
<td>Improves symptoms and cytopenias; does not correct underlying disease</td>
</tr>
<tr>
<td>Splenic artery aneurysm</td>
<td>Best for distal lesions near splenic hilum</td>
<td>Curative</td>
</tr>
<tr>
<td>Symptomatic nonparasitic cysts</td>
<td>Partial splenectomy for small cysts; unroofing for large cysts</td>
<td>Curative</td>
</tr>
<tr>
<td>Symptomatic parasitic cysts</td>
<td>Therapy of choice</td>
<td>Curative; exercise caution not to spill cyst contents</td>
</tr>
<tr>
<td>Wandering spleen</td>
<td>Abdominal pain or splenomegaly (venous congestion)</td>
<td>Curative</td>
</tr>
<tr>
<td>Traumatic rupture</td>
<td>Grades 4/5, or failure of conservative management of lower grades</td>
<td>Curative</td>
</tr>
</tbody>
</table>
later section. Management of splenic injury in the trauma patient is also beyond the scope of this chapter and discussed elsewhere. The most common indications for elective splenectomy are malignancy and hematologic autoimmune disorders, principally, idiopathic thrombocytopenic purpura (ITP) and autoimmune hemolytic anemia (AIHA).

Benign Disorders

Red Blood Cell Disorders

Congenital

Hereditary Spherocytosis. Hereditary spherocytosis (HS) is the most common type of hemolytic anemia for which splenectomy is indicated and the third most common type of congenital hemolytic anemia overall. HS results from an inherited dysfunction or deficiency in one of the erythrocyte membrane proteins (alpha or beta spectrin, ankyrin, band 3 protein, or protein 4.2). The resulting destabilization of the membrane lipid bilayer allows a pathologic release of membrane lipids. The red blood cell assumes a more spherical, less deformable shape, and the spherocytic erythrocytes are sequestered and destroyed in the spleen and hemolytic anemia ensues. HS is inherited primarily (70–80% of the time) in an autosomal dominant fashion; the estimated prevalence in Western populations is roughly 1 in 2000.

Patients with typical HS forms may have mild jaundice. Splenomegaly usually is palpable on physical examination. Laboratory examination reveals varying degrees of anemia: patients with mild forms of the disease may not have anemia; patients with moderate to severe forms may have hemoglobin levels as low as 4 to 6 g/dL. The mean corpuscular volume is typically low to normal or slightly decreased. For screening, a combined elevated mean corpuscular hemoglobin concentration and an elevated erythrocyte distribution width are an excellent predictor. Other laboratory indicators of HS include those providing evidence of rapid red blood cell destruction, including elevated reticulocyte count, elevated lactate dehydrogenase level, and increased level of unconjugated bilirubin. Spherocytes are readily apparent on peripheral blood film.

Risks and benefits should be assessed carefully before splenectomy and cholecystectomy are performed for HS. The main indications are moderate to severe symptomatic hemolytic anemia, growth retardation, skeletal changes, leg ulcers, and extramedullary hemopoietic tumors in young patients.

If gallstones coexist with spherocytosis, the gallbladder should be removed, but prophylactic cholecystectomy without gallstones is controversial. Near total splenectomy is advocated in children. Dramatic clinical improvement—despite persistent hemolysis—usually occurs after splenectomy in patients with moderate to severe disease. Because children can be affected with HS, the timing of splenectomy is important and is aimed at reducing the diminutive possibility of overwhelming postsplenectomy sepsis. Delaying such an operation until the patient is between the ages of 4 and 6—unless the anemia and hemolysis accelerate—is recommended by most experts.

Red Blood Cell Enzyme Deficiencies. Red blood cell enzyme deficiencies associated with hemolytic anemia may be classified into two groups: deficiencies of enzymes involved in glycolytic pathways, such as pyruvate kinase deficiency, and deficiencies of enzymes needed to maintain a high ratio of reduced to oxidized glutathione in the red blood cell, protecting it from oxidative damage, such as glucose-6-phosphate dehydrogenase (G6PD) deficiency.

Pyruvate Kinase Deficiency. Pyruvate kinase (PK) deficiency is the most common glycolytic defect causing congenital non-spherocytic hemolytic anemia. Since its first description in the early 1960s, vast amounts of information have been elucidated about the genetic diversity of the disease, red blood cell clearance, long-term complications and treatment options including transfusion and splenectomy. PK deficiency affects people worldwide, with a slight preponderance among those of Northern European or Chinese descent. Its estimated prevalence in the Caucasian population is 51 per million. Clinical manifestations of the disease vary widely, from transfusion-dependent severe anemia in early childhood to well-compensated mild anemia in adolescents or adults. Prenatal hydrops fetalis has also been reported. Pyruvate kinase enzyme activity is the gold standard for initial diagnostic testing or by detection of specific mutations at the complementary DNA or genomic level. Splenomegaly is common, and in severe cases, splenectomy can alleviate transfusion requirements. As with other disorders that cause hemolytic anemia in children, splenectomy should be delayed if possible to at least 4 years of age to reduce the risk of postsplenectomy infection.

Glucose-6-Phosphate Dehydrogenase Deficiency. The most common red blood cell enzyme deficiency overall is G6PD deficiency. It is far more prevalent than PK deficiency with more than 400 million people affected worldwide, although most experience only moderate health risks and no longevity reduction. Clinical manifestations—chronic hemolytic anemia, acute intermittent hemolytic episodes, or no hemolysis—depend on the variant of G6PD deficiency. The mainstay of therapy is avoidance of drugs known to precipitate hemolysis in patients with G6PD deficiency. Transfusions are given in cases of symptomatic anemia. Conventional wisdom is that splenectomy is not indicated in this disease, and certainly the overwhelming majority of patients with G6PD deficiency will neither require nor benefit from splenectomy. However, one report described a small case series of six symptomatic G6PD deficiency patients who had severe hemolytic anemia and required transfusion, all of whom were identified to share a common mutation at exon 10. All underwent splenectomy. A complete response occurred in four patients (transfusion requirement eliminated), and a partial response occurred in one patient (transfusion requirement reduced); no follow-up data were provided for the remaining patient. This study indicates that for a carefully select group of patients with severe hemolytic anemia attributable to G6PD deficiency, splenectomy may be of benefit, although more data is needed before strong recommendation can be made.

Acquired

Warm-Antibody Autoimmune Hemolytic Anemia. Autoimmune hemolytic anemias (AIHAs) are characterized by the destruction of red blood cells, whose erythrocyte life span is diminished by autoantibodies leveled against antigens. AIHA is classified as either primary or secondary, depending on whether an underlying cause, such as a disease or toxin, is identified. AIHA is also divided into “warm” and “cold” categories, based on the temperature at which the autoantibodies exert their effect. In cold-agglutinin disease, severe symptoms are uncommon and splenectomy is almost never indicated; therefore, this entity is not discussed further in this section. However, warm-antibody AIHA has clinical consequences with which the surgeon should be familiar.
Warm-antibody AIHA, although occurring primarily in midlife, can affect individuals at all ages. The disorder is more common among women, and fully half of warm-antibody AIHA cases are idiopathic. Clinical presentation may be acute or gradual. Findings include mild jaundice and symptoms and signs of anemia. One-third to one-half of patients present with splenomegaly. Sometimes in such cases the spleen is palpable on physical examination. The diagnosis relies on demonstrating hemolysis as indicated by anemia, reticulocytosis, and/or products of red blood cell destruction, including bilirubin, in the blood, urine, and stool. A positive result on direct Coombs’ test confirms the AIHA diagnosis by distinguishing autoimmune from other forms of hemolytic anemia.

Treatment of AIHA depends on the severity of the disease and whether it is primary or secondary. Severe symptomatic anemia demands prompt attention, often requiring red blood cell transfusion. The mainstay treatment for both primary and secondary forms of symptomatic, unstable AIHA remains long-term corticosteroid administration. Therapy should continue until a response is noted by a rise in hematocrit and fall in reticulocyte count, which generally occurs within 3 weeks.

Clinical response to splenectomy for AIHA varies, and the evidence from a number of small case series is conflicting. For example, one 2004 series reported a favorable response to splenectomy in patients with AIHA secondary to chronic lymphocytic leukemia, whereas more recent series found no benefit from splenectomy in patients with AIHA secondary to systemic lupus erythematosus or inflammatory bowel disease.

Favorable responses to splenectomy have been reported in patients with warm-antibody AIHA, with a recent series showing complete remission of refractory AIHA following laparoscopic splenectomy at 35-month follow-up in patients over 60 years old. Transient responses are more common, however, and many patients eventually experience hemolysis again despite splenectomy. The decision regarding splenectomy in the case of AIHA should be individualized based on careful consideration of the clinical history and frank discussion with the patient. It is considered as a third-line therapy after failure of steroids or anti-CD20 antibody administration.

**Hemoglobinopathies.** Sickle cell disease is an inherited chronic hemolytic anemia that results from the mutant sickle cell hemoglobin (HbS) within the red blood cell and is inherited in an autosomal dominant fashion. Persons who inherit an \( HbS \) gene from one parent (heterozygous) are carriers; those who inherit an \( HbS \) gene from both parents (homozygous) have sickle cell anemia.

In sickle cell disease, the underlying abnormality is the mutation of adenine to thymine in the sixth codon of the \( \beta \)-globin gene, which results in the substitution of valine for glutamic acid as the sixth amino acid of the \( \beta \)-globin chain. Mutant \( \beta \) chains included in the hemoglobin tetramer create HbS. Deoxygenated HbS is insoluble and becomes polymerized and sickled. The subsequent lack of deformability of the red blood cell, in addition to other processes, results in microvascular congestion, which may lead to thrombosis, ischemia, and tissue necrosis. The disorder is characterized by painful intermittent episodes.

Sequestration occurs in the spleen, with splenomegaly resulting early in the disease course. In most patients, subsequent infarction of the spleen and autosplenectomy occur at some later time. The most frequent indications for splenectomy in sickle cell disease are recurrent acute sequestration crises, hypersplenism, and splenic abscess. The occurrence of one major acute sequestration crisis, characterized by rapid painful enlargement of the spleen and circulatory collapse, generally is considered sufficient grounds for splenectomy. Both partial and total splenectomy have been shown to control clinical symptoms in children but may not change hematologic parameters. Preoperative preparation should include special attention to adequate hydration and avoidance of hypothermia.

Transfusions are often indicated for anemia, for moderately severe episodes of acute chest syndrome (i.e., a new infiltrate on chest radiograph associated with new symptoms, such as fever, cough, sputum production, or hypoxia), and preoperatively before splenectomy. Patients experiencing stroke or a severe crisis may require hydration and an exchange transfusion, which may be performed manually or with automated apheresis equipment. Hydroxyurea is an oral chemotherapeutic agent that upregulates fetal hemoglobin, which interferes with polymerization of HbS and thus reduces the sickling process.

**Thalassemia.** Thalassemia is the term for a group of inherited disorders of hemoglobin synthesis prevalent among people of Mediterranean extraction and classified according to the globin chain (\( \alpha \), \( \beta \), or \( \gamma \)) affected. As a group, the thalassemias are the most common genetic diseases known to arise from a single gene defect. Most forms of this disorder are inherited in Mendelian recessive fashion from asymptomatic carrier parents. In the so-called thalassemia belt that extends throughout the shores of the Mediterranean as well as through the Arabian Peninsula, Turkey, Iran, India, and southeastern Asia, the incidence of thalassemia is between 2.5% and 15%. However, thalassemias have been found in people of all ethnic origins.

In all forms of thalassemia, the primary defect is absent or reduced production of hemoglobin chains. From this abnormality, two significant consequences arise: (a) reduced functioning of hemoglobin tetramers, yielding hypochromia and microcytosis; and (b) unbalanced biosynthesis of individual \( \alpha \) and \( \beta \) subunits, which results in insoluble red blood cells that cannot release oxygen normally and may precipitate with cell aging. Both underproduction of hemoglobin and excess production of unpaired hemoglobin subunits contribute to thalassemia-associated morbidity and mortality.

A diagnosis of thalassemia major (homozygous form) is made by demonstrating hypochromic microcytic anemia associated with randomly distorted red blood cells and nucleated erythrocytes (target cells) on peripheral blood smear. Elevated reticulocyte count and white blood cell count are among the associated findings. Because \( \alpha \) chains are needed to form both fetal hemoglobin and adult hemoglobin, \( \alpha \)-thalassemia becomes symptomatic in utero or at birth. By contrast, \( \beta \)-thalassemia becomes symptomatic at 4 to 6 months because \( \beta \) chains are involved only in adult hemoglobin synthesis.

The clinical spectrum of the thalassemias is wide. Heterozygous carriers of the disease are usually asymptomatic. Homozygous individuals, on the other hand, typically present before 2 years of age with pallor, growth retardation, jaundice, and abdominal swelling due to liver and spleen enlargement. Among other characteristics of thalassemia major are intractable leg ulcers, head enlargement, frequent infections, and the need for periodic blood transfusions. Untreated individuals usually die in late infancy or early childhood from severe anemia.

Treatment for thalassemia involves red blood cell transfusions to maintain a hemoglobin level of >9 mg/dL, along with...
intensive parenteral chelation therapy with deferoxamine. Splenectomy is indicated for patients with excessive transfusion requirements (>200 mL/kg per year), discomfort due to splenomegaly, or painful splenic infarction. Careful assessment of the risk-benefit ratio is essential. Thalassemia patients are at high risk for pulmonary hypertension after splenectomy; however, the precise pathophysiology of this sequela remains unclear.\textsuperscript{45,46} The increase in infectious complications is likely to be due to a coexisting immune deficiency, in large part brought about by iron overload, which may be associated both with the thalassemia itself and with transfusions. The disproportionately high rate of overwhelming postsplenectomy infection in thalassemia patients has led some investigators to consider partial splenectomy in children; some success in reducing mortality has been reported.\textsuperscript{47} However, splenectomy should be delayed until after the age of 4 years unless it is absolutely necessary.

**Platelet Disorders**

**Idiopathic Thrombocytopenic Purpura**  Idiopathic thrombocytopenic purpura (ITP), also called immune thrombocytopenic purpura, is an autoimmune disorder characterized by a low platelet count and mucocutaneous and petechial bleeding. The low platelet count stems from premature removal of platelets opsonized by antiplatelet immunoglobulin G autoantibodies produced in the spleen. This clearance occurs through the interaction of platelet autoantibodies with Fc receptors expressed on tissue macrophages, predominantly in the spleen and liver. The estimated incidence of ITP is 100 persons per million annually, about one-half of whom are children.\textsuperscript{48} Adult-onset and childhood-onset ITP are strikingly different in their clinical course and management.

Patients with ITP typically present with petechiae or ecchymoses, although some experience major bleeding from the outset. Bleeding may occur from mucosal surfaces in the form of gingival bleeding, epistaxis, menorrhagia, hematuria, or even melena. The severity of bleeding frequently corresponds to the deficiency in platelets: Patients with counts greater than 50,000/mm\(^3\) usually present with incidental findings; those with counts between 30,000 and 50,000/mm\(^3\) often have easy bruising; those with platelet counts between 10,000 and 30,000/mm\(^3\) may develop spontaneous petechiae or ecchymoses; and those with counts less than 10,000/mm\(^3\) are at risk for internal bleeding.\textsuperscript{49} The incidence of major intracranial hemorrhage is approximately 1%, and it usually occurs early in the disease course. The duration of the bleeding helps to distinguish acute from chronic forms of ITP. Children often present at a young age (peak age of approximately 5 years) with sudden onset of petechiae or purpura several days to weeks after an infectious illness. In contrast, adults experience a more chronic form of disease with an insidious onset. Splenomegaly is uncommon with ITP in both adults and children, and its occurrence should prompt a search for a separate cause of thrombocytopenia. Up to 10% of children, however, have a palpable spleen tip.

Diagnosis of ITP is based on exclusion of other possibilities in the presence of a low platelet count and mucocutaneous bleeding. Other diseases resulting in secondary forms of ITP, such as systemic lupus erythematosus, antiphospholipid syndrome, lymphoproliferative disorders, human immunodeficiency virus (HIV) infection, and hepatitis C, should be identified and treated when present. In addition, any history of use of a drug known to cause thrombocytopenia, such as certain antimicrobials, anti-inflammatories, antihypertensives, and antidepressants, should be sought. In addition to low platelet count, another laboratory finding characteristic of ITP is the presence of large, immature platelets (megakaryocytes) on peripheral blood smear.

The usual first line of therapy is oral prednisone at a dose of 1.0 to 1.5 mg/kg per day.\textsuperscript{49,50} No consensus exists as to the optimal duration of steroid therapy, but most responses occur within the first 3 weeks. Response rates range from 50% to 75%, but relapses are common. IV immunoglobulin, given at 1.0 g/kg per day for 2 to 3 days, is indicated for internal bleeding when platelet counts remain less than 5000/mm\(^3\), when extensive purpura exists, or to preoperatively boost platelets. IV immunoglobulin is thought to impair clearance of immunoglobulin G–coated platelets by competing for binding to tissue macrophage receptors. An immediate response is common, but a sustained remission is not. The medical approach to ITP has been modified with the advent of rituximab and thrombopoietin receptor agonists.\textsuperscript{51} Therapeutic recommendations are summarized in Table 34-2. Splenectomy is selectively indicated for failure of medical therapy, for prolonged use of steroids with undesirable effects, and in selected cases after first relapse.\textsuperscript{1} Prolonged use of steroids can be defined in various ways, but a persistent need for more than 10 to 20 mg/d for 3 to 6 months to maintain a platelet count of greater than 30,000/mm\(^3\) generally prompts referral for splenectomy. Splenectomy is an effective option for refractory ITP and provides a permanent response without subsequent need for steroids in 75% to 85% of patients. Two recent reviews and meta-analyses have assessed the global results of splenectomy for ITP, specifically after the use of laparoscopic techniques.

<table>
<thead>
<tr>
<th>Table 34-2</th>
<th>Second-line treatment options for immune thrombocytopenia</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Splenectomy</strong></td>
<td><strong>Advantages</strong></td>
</tr>
<tr>
<td></td>
<td>• Effective, time-honored treatment</td>
</tr>
<tr>
<td></td>
<td>• Most effective response (overall 88%, partial 22%, complete 66%)</td>
</tr>
<tr>
<td></td>
<td><strong>Disadvantages</strong></td>
</tr>
<tr>
<td></td>
<td>• Risk of overwhelming postsplenectomy infection</td>
</tr>
<tr>
<td></td>
<td>• Increased risk of thrombotic events</td>
</tr>
<tr>
<td></td>
<td>• Morbidity and mortality associated with an operation</td>
</tr>
<tr>
<td><strong>Rituximab</strong></td>
<td><strong>Advantages</strong></td>
</tr>
<tr>
<td></td>
<td>• Nonsurgical</td>
</tr>
<tr>
<td></td>
<td>• Good experience since 1999</td>
</tr>
<tr>
<td></td>
<td>• Initial immediate response 63%</td>
</tr>
<tr>
<td></td>
<td>• 31% response at 2 years</td>
</tr>
<tr>
<td></td>
<td><strong>Disadvantages</strong></td>
</tr>
<tr>
<td></td>
<td>• Severe toxicity in 2%–6%</td>
</tr>
<tr>
<td><strong>Thrombopoietin Receptor Agonist</strong></td>
<td><strong>Advantages</strong></td>
</tr>
<tr>
<td></td>
<td>• Nonsurgical</td>
</tr>
<tr>
<td></td>
<td>• Oral agent self-administered weekly</td>
</tr>
<tr>
<td></td>
<td>• 80% response rate</td>
</tr>
<tr>
<td></td>
<td><strong>Disadvantages</strong></td>
</tr>
<tr>
<td></td>
<td>• Long-term treatment required</td>
</tr>
<tr>
<td></td>
<td>• Potential toxicity</td>
</tr>
<tr>
<td></td>
<td>• Long-term data lacking at this time</td>
</tr>
</tbody>
</table>
In 2004, Kojouri and colleagues reviewed 135 case series from 1966 to 2004, pooling 4955 patients.\textsuperscript{51} The long-term platelet count response was assessed, as was the ability to predict response and the incidence of complications. Complete response was achieved in 66\% of cases with a follow-up ranging from 1 to 153 months, and complete and partial responses occurred in as many as 88\% of patients, regardless of the duration of follow-up. They also analyzed 12 preoperative demographic, clinical, and laboratory parameters and found no predictive capability of platelet response in any of them. Mortality was very low (1\%), and morbidity was 10\%. Limitations of this review included old case series and a low percentage of laparoscopic splenectomies. In a 2009 systemic review involving 1223 patients, Mikael and colleagues evaluated the short- and long-term outcomes after laparoscopic splenectomy.\textsuperscript{52} The conversion rate to open surgery was 5.6\%, and the immediate nonresponder rate was 8.2\%; however, eventually a clinical response was achieved in 72\% of the patients on long-term follow-up (5 years). The initial concerns regarding the potential for missing accessory spleens, longer operative times, and increased cost related to laparoscopic versus open splenectomy have been resolved. Laparoscopy is now the approach of choice for elective splenectomy for ITP, with recent studies showing improved short-term results and comparable long-term results to conventional open splenectomy in this condition.\textsuperscript{53-56}

In children with ITP, the course is typically self-limited, with durable and complete remission in greater than 70\% of patients regardless of therapy. Because of the good prognosis without treatment, the decision to intervene surgically is controversial and is largely to obviate intracranial hemorrhage as discussed earlier. Thus, children with typical ITP—and certainly those without hemorrhage—are managed principally by observation, with short-term therapy in select cases.\textsuperscript{57} Urgent splenectomy, in conjunction with aggressive medical therapy, may play a role in the rare circumstance of severe, life-threatening bleeding in both children and adults.

**Thrombotic Thrombocytopenic Purpura** Thrombotic thrombocytopenic purpura (TTP) is a serious disorder characterized by thrombocytopenia, microangiopathic hemolytic anemia, and neurologic complications. Abnormal platelet clumping occurs in arterioles and capillaries, reducing the lumen of these vessels and predisposing the patient to microvascular thrombotic episodes. The reduced lumen size also causes shearing stresses on erythrocytes, which leads to deformed red blood cells subject to hemolysis. Hemolysis may also be due in part to sequestration and destruction of erythrocytes in the spleen. Research has demonstrated that the underlying abnormality is likely related to the persistence of unusually large multimers of von Willebrand factor associated with platelet clumping in the patient’s blood.\textsuperscript{58,59}

TTP occurs in approximately 3.7 individuals per million, but this rare disorder’s dramatic clinical sequelae and favorable response to early therapy demand an understanding of its clinical presentation to ensure an early diagnosis. Clinical features of the disorder include petechiae, fever, neurologic symptoms, renal failure, and, infrequently, cardiac symptoms such as heart failure or arrhythmias. Petechial hemorrhages in the lower extremities are the most common presenting sign. Along with fever, patients may experience flu-like symptoms, malaise, or fatigue. Neurologic changes range from generalized headaches to altered mental status, seizures, and even coma. Generally, however, the mere presence of petechiae and thrombocytopenia is sufficient to lead to the diagnosis of TTP and consideration of treatment.

The diagnosis is confirmed by the peripheral blood smear, which shows schistocytes, nucleated red blood cells, and basophilic stippling. Although other conditions such as tight aortic stenosis or prosthetic valves may lead to the presence of schistocytes, these conditions generally are not accompanied by thrombocytopenia. TTP may be distinguished from autoimmune causes of thrombocytopenia, such as Evans’ syndrome (ITP and autoimmune hemolytic anemia) or systemic lupus erythematosus, by a negative result on Coombs’ test.

Plasma exchange is the first-line therapy for TTP. This treatment consists of the daily removal of a single volume of the patient’s plasma and its replacement with fresh-frozen plasma until the thrombocytopenia, anemia, and associated symptoms are corrected. Therapy is then tapered over 1 to 2 weeks.\textsuperscript{40} Splenectomy should be considered in patients who experience relapse or who require multiple plasma exchanges to control severe symptoms, and generally is well tolerated without significant morbidity.\textsuperscript{59,60}

**Malignant Conditions**

**White Blood Cell Disorders.** The role of splenectomy in patients with white blood cell disorders varies. As for the myelogenous diseases mentioned previously, splenectomy for white blood cell disorders can be effective therapy for symptomatic splenomegaly and hypersplenism, improving some clinical parameters but generally not altering the course of the underlying disease or long-term prognosis. Historically, splenectomy has played a role during surgical staging for Hodgkin’s disease, although this practice has become largely obsolete with the advent of imaging technologies (Computed tomography [CT] scan and 18F-fluorodeoxyglucose positron emission tomography [18F-FDG PET]).\textsuperscript{61-63} Careful consideration of the intended benefits of splenectomy must be weighed against the significant perioperative and postsplenectomy risks in this often complex patient population.\textsuperscript{64}

**Hairy Cell Leukemia** Hairy cell leukemia (HCL) is an uncommon blood disorder, representing only 2\% of all adult leukemias. HCL is characterized by splenomegaly, pancytopenia, and large numbers of abnormal lymphocytes in the bone marrow. These lymphocytes contain irregular hair-like cytoplasmic projections identifiable on the peripheral smear. Many HCL patients have few symptoms and require no specific therapy. Splenectomy for HCL was historically performed as a palliative procedure, alleviating the symptoms associated with splenomegaly and normalizing peripheral blood counts in the majority of patients, but not leading to morphological bone marrow remissions. However, with the advent of diverse new drugs (e.g., rituximab, pentostatin, cladribine), splenectomy has become rarely performed.\textsuperscript{65} Splenectomy should be considered after exhaustive systemic therapy or for those with nontraumatic rupture, and it has also been used to treat pregnant women with HCL to delay onset of chemotherapy.\textsuperscript{66}

**Hodgkin’s Lymphoma** Hodgkin’s Lymphoma (HL) is a disorder of the lymphoid system characterized by the presence of Reed-Sternberg cells (which actually form the minority of the Hodgkin’s tumor). More than 90\% of patients with HL present with lymphadenopathy above the diaphragm. Lymph nodes can become particularly bulky in the mediastinum, which may result in shortness of breath, cough, or obstructive pneumonia.
Lymphadenopathy below the diaphragm is rare on presentation but can arise with disease progression. The spleen is often an occult site of spread, but massive splenomegaly is not common. In addition, large spleens do not necessarily signify involvement.62

Four major histologic types exist: lymphocyte predominance type, nodular scleros type, mixed cellularity type, and lymphocyte depletion type. The histologic type, along with location of disease and symptomatology, influence survival for patients with HD. Stage I disease is limited to one anatomic region; stage II disease is defined by the presence of two or more contiguous or noncontiguous regions on the same side of the diaphragm; stage III disease involves disease on both sides of the diaphragm, but limited to lymph nodes, spleen, and Waldeyer’s ring (the ring of lymphoid tissue formed by the lingual, palatine, and nasopharyngeal tonsils); and stage IV disease includes involvement of the bone marrow, lung, liver, skin, gastrointestinal tract, or any organ or tissue other than the lymph nodes or Waldeyer’s ring.62

Staging laparotomy for HL is less commonly performed in the current era of minimally invasive surgery and advanced imaging techniques. More liberal use of diagnostic tools (CT scan and 18F-FDG PET) and chemotherapy for patients with HL has drastically reduced the indications for surgical staging. Current indications for surgical staging include clinical suspicion of lymphoma without evidence of peripheral disease or patients requiring restaging for suspicion of failure after chemotherapy.62,67

**Non-Hodgkin’s Lymphoma** Non-Hodgkin’s lymphoma (NHL) encompasses all malignancies derived from the lymphoid system except classic HL.63 A proliferation of any one of the three predominant lymph cell types—natural killer cells, T cells, or B cells—may be included in the category of NHL. Because of the wide net cast by NHL, the clinical presentations of the disorders under its umbrella vary. The subentities of NHL may be clinically classified into nodal or extranodal, as well as indolent, aggressive, and very aggressive groups. Patients with indolent lymphomas may present with mild or no symptoms and seek medical attention for a swollen lymph node, whereas the aggressive and very aggressive lymphomas create easily noticeable symptoms, such as pain, swelling due to obstruction of vessels, fever, and night sweats. Surgical staging is no longer indicated for NHL because the combination of history and physical examination, chest radiograph and abdominal/pelvic CT scan, biopsy of involved lymph nodes (including laparoscopically directed nodal and liver biopsies), and bone marrow biopsy is sufficient.63 Splenomegaly exists in some, but not all, forms of NHL. Splenectomy does not alter the natural history of the disease, but it is indicated in cases where a diagnosis cannot be established by obtaining peripheral tissue and clinical suspicion remains or for management of symptoms related to an enlarged spleen as well as for improvement of cytopenias.68-72

**Chronic Lymphocytic Leukemia** Chronic lymphocytic leukemia (CLL) is currently considered a subtype of NHL. The main characteristic of CLL is a progressive accumulation of old and nonfunctional lymphocytes.63,73 Symptoms of CLL are nonspecific and include weakness, fatigue, fever without illness, night sweats, and frequent bacterial and viral infections. The most frequent finding is lymphadenopathy. When the spleen is enlarged, it may be massive or barely palpable below the costal margin. Splenectomy is indicated to improve cytopenias and was shown to be 75% effective in a combined group of patients who had either CLL or nonmalignant HD.29 Splenectomy may facilitate chemotherapy in patients whose cell counts were prohibitively low before spleen removal. Palliative splenectomy also is indicated for symptomatic splenomegaly.

**Bone Marrow Disorders (Myeloproliferative Disorders).** The myeloproliferative disorders are characterized by an abnormal growth of cell lines in the bone marrow. They include chronic myeloid leukemia, acute myeloid leukemia, chronic myelomonocytic leukemia, essential thrombocytopenia, polycythemia vera, and myelofibrosis, also known as agnogenic myeloid metaplasia (see “Myelofibrosis [Agnogenic Myeloid Metaplasia]” later in this chapter). The common underlying problem leading to splenectomy in these disorders is symptomatic splenomegaly. Symptoms due to splenomegaly are due to mass effect and consist of early satiety, poor gastric emptying, heaviness or pain in the left upper quadrant, and even diarrhea. Hypersplenism, when it occurs in these conditions, usually is associated with splenomegaly. Splenectomy performed in the setting of the myeloproliferative disorders is generally for treatment of the pain, early satiety, and other symptoms of splenomegaly. Radiation has been used since 1903 to treat symptomatic splenomegaly in myeloproliferative disorders, but today it is principally used in situations in which splenectomy is not an option.

**Chronic Myelogenous Leukemia** Chronic myelogenous leukemia (CML) is a disorder of the primitive pluripotent stem cells in the bone marrow that results in a significant increase in erythroid, megakaryotic, and pluripotent progenitors in the peripheral blood smear. The genetic hallmark is a transposition between the bcr gene on chromosome 9 and the abl gene on chromosome 22. CML accounts for 7% to 15% of all leukemias, with an incidence of 1.5 in 100,000 in the United States.44 It is often asymptomatic, but CML can cause fatigue, anorexia, sweating, and left upper quadrant pain and early satiety secondary to splenomegaly. Enlargement of the spleen is found in roughly one-half of patients with CML. Current therapy includes imatinib or allogeneic stem cell transplantation. Splenectomy is indicated to relieve symptoms of massive splenomegaly, but it does not prevent blast crisis or alter the disease process.45,70

**Acute Myeloid Leukemia** Like CML, acute myeloid leukemia (AML) involves the abnormal growth of stem cells in the bone marrow. Unlike CML, AML has a presentation that is more rapid and dramatic. The proliferation and accumulation of hematopoietic stem cells in the bone marrow and blood inhibit the growth and maturation of normal red blood cells, white blood cells, and platelets. Death usually results within weeks to months if AML goes untreated. The incidence of AML is approximately 9200 new cases each year in the United States, and it accounts for 1.2% of all cancer deaths.74 Patients with other myeloproliferative disorders, such as polycythemia vera, primary thrombocytosis, or myeloid metaplasia, are at increased risk for leukemic transformation to AML. Presenting signs and symptoms of AML include a viral-like illness with fever, malaise, and frequently bone pain due to the expansion of the medullary space. Standard treatment is combined induction therapy with daunorubicin, cytarabine, and stem cell transplantation. Splenectomy is indicated in AML only in the uncommon circumstance that left upper quadrant pain and early satiety become unbearable. The benefit must be weighed against the heightened risk of postsplenectomy infection in AML patients immunocompromised due to neutropenia and chemotherapy.74
**Chronic Myelomonocytic Leukemia** Like CML and AML, chronic myelomonocytic leukemia (CMML) is characterized by a proliferation of hematopoietic elements in the bone marrow and blood. CMML differs from CML in that it is associated with monocytosis in the peripheral smear (>1 × 10^3 monocytes/mm^3) and in the bone marrow. Splenomegaly occurs in one-half of these patients, and splenectomy can result in symptomatic relief.75

**Essential Thrombocytopenia** Essential thrombocytopenia (ET) represents abnormal growth of the megakaryocyte cell line, resulting in increased levels of platelets in the bloodstream. The diagnosis is made after the exclusion of other chronic myeloid disorders such as CML, polycythemia vera, and myelofibrosis that may also present with thrombocytosis.76 Clinical manifestations of ET include vasomotor symptoms, thrombohemorrhagic events, recurrent fetal loss, and the transformation to myelofibrosis with myeloid metaplasia or AML. Hydroxyurea is used to reduce thrombotic events in ET but does not alter transformation to myelofibrosis or leukemia. Splenomegaly occurs in one-third to one-half of patients with ET. Splenectomy is not felt to be helpful in the early stages of ET and is best reserved for the later stages of disease, when myeloid metaplasia has developed.69 Even in these circumstances, candidates should be chosen selectively because significant bleeding has been reported to complicate splenectomy in these patients.

**Polycythemia Vera** Polycythemia vera (PV) is a clonal, chronic, progressive myeloproliferative disorder characterized by an increase in red blood cell mass, frequently accompanied by leukocytosis, thrombocytosis, and splenomegaly. Patients with PV typically enjoy longer survival than those affected by hematologic malignancies but remain at risk for transformation to myelofibrosis or AML. The disease is rare, with an annual incidence of 5 to 17 cases per million population.76,77 Physical findings include ruddy cyanosis, conjunctival plethora, hepatomegaly, splenomegaly, and hypertension. Treatment should be tailored to the risk status of the patient and ranges from phlebotomy and aspirin administration to the use of chemotherapeutic agents. As in ET, splenectomy is not helpful in the early stages of disease and is best reserved for patients with late-stage disease in whom myeloid metaplasia has developed and splenomegaly-related symptoms are severe.76,77

**Myelofibrosis (Agnogenic Myeloid Metaplasia)** The term myelofibrosis may be used to describe either the generic condition of fibrosis of the bone marrow (which may be associated with a number of benign and malignant disorders) or a specific, chronic, malignant hematologic disease associated with splenomegaly, the presence of red blood cell and white blood cell progenitors in the bloodstream, marrow fibrosis, and extramedullary hematopoiesis, otherwise known as agnogenic myeloid metaplasia (AMM). AMM also can be referred to as myelosclerosis, idiopathic myeloid metaplasia, and osteosclerosis. In this chapter, the term myelofibrosis is synonymous with AMM.

In AMM, fibrosis of the bone marrow is believed to be a response to a clonal proliferation of hematopoietic stem cells. Marrow failure is common. The true incidence of AMM is unknown due to the scarcity of epidemiologic data, but one study estimated its U.S. incidence at 1.46 per 100,000 population.78-80 The diagnosis is made by a careful examination of the peripheral blood smear and bone marrow. Nucleated red blood cells and immature myeloid elements in the blood are present in 96% of cases and strongly suggest the diagnosis. Teardrop poikilocytosis is another frequent finding. Care must be taken, however, to exclude a history of a primary neoplasm (such as lymphoma or adenocarcinoma of the stomach, lung, prostate, or breast) or tuberculosis because patients with these conditions may develop secondary myelofibrosis.

Treatment depends on symptoms: Asymptomatic patients are closely followed, whereas symptomatic patients undergo therapeutic intervention targeted to their symptoms. The only curative therapy is allogeneic bone marrow transplantation in younger, high-risk patients. Supportive therapy for clinically symptomatic anemia includes steroids, danazol, erythropoietin, or blood transfusion.78,80 Splenomegaly-related symptoms are best treated with splenectomy. Although some chemotherapeutic agents (busulfan, hydroxyurea, interferon-α) and low-dose radiation can reduce splenic size, their discontinuation usually results in rapid splenic regrowth.

A thorough preoperative workup must precede splenectomy in patients with AMM. The candidate must possess acceptable cardiac, pulmonary, hepatic, and renal reserve for the operation. The coagulation system should be examined; testing should include measurement of coagulation factors V and VIII and fibrin split products, platelet count, and bleeding time. Low platelet counts may require administration of steroids and/or platelet transfusion at the time of surgery. Splenectomy provides durable, effective palliation for nearly all patients with AMM, although postoperative complications are more common in patients with AMM than in those with other hematologic indications. The Mayo Clinic recently published its 30-year experience with 314 myelofibrosis patients who underwent splenectomy. Nearly half of the operations (49%) were performed to alleviate the mechanical symptoms of splenomegaly; the remainder were undertaken to manage anemia, thrombocytopenia, or portal hypertension. Response to splenectomy was 76% overall at 1 year; overall complication rate was 28%, including 21 perioperative deaths.80 Thrombosis, hemorrhage, and infection complications were common, with perioperative thrombocytopenia an independent predictor of mortality risk. These data underscore the severity of this malignancy and emphasize the need for careful patient selection when considering splenectomy in AMM.70

**Tumors and Metastasis.** Primary tumors of the spleen are typically benign or malignant variants of vascular neoplasms. The most common primary tumors of the spleen are sarcomas, many of which have been linked to various environmental and occupational exposures such as vinyl chloride or thorium dioxide. Isolated splenic metastases are extremely unusual but may occur in the setting of a concomitant carcinoma.81 Lung carcinoma is the tumor that most commonly spreads to the spleen, although, colorectal, ovarian, and melanoma may also metastasize to the spleen.82,83 If after a thorough examination, an isolated splenic metastasis is confirmed, a laparoscopic splenectomy with intact spleen retrieval may be considered.81-83

**Miscellaneous Disorders and Lesions**

**Infections and Abscesses.** Primary infections of the spleen although uncommon in immunocompetent adults, are particularly reported in the immunocompromised population or those with a history of intravenous recreational drug abuse.84,85 The potential effects of certain systemic infections on the spleen merit close attention, mostly because of the dreaded risk of spontaneous splenic rupture. Infectious mononucleosis due to
either Epstein-Barr virus or cytomegalovirus infection imparts a small but often-discussed risk of spontaneous splenic rupture in both adults and children. The true incidence may be underreported, however. Recent case reports abound in the literature regarding spontaneous splenic rupture due to a variety of infectious causes (malaria, *Listeria* infection, fungal infections, dengue, and Q fever, to name a few) as well as a variety of neoplastic and other noninfectious causes (lymphoma, angiosarcoma, amyloidosis, pregnancy). The presumed pathophysiologic mechanism is inhibition of the splenic parenchyma with inflammatory cells, which distorts the architecture and fibrous support system of the spleen and thins the splenic capsule. In this setting, splenic rupture can occur spontaneously or after a seemingly minor external trauma or even a Valsalva maneuver.

Abscesses of the spleen are uncommon, with an incidence reported to be 0.14% to 0.7%. They occur more frequently in tropical locations, where they are associated with thrombosed splenic vessels and infarction in patients with sickle cell anemia. Five distinct mechanisms of splenic abscess formation have been described: (a) hematogenous infection; (b) contiguous infection; (c) hemoglobinopathy; (d) immunosuppression, including HIV infection and chemotherapy; and (e) trauma. The presentation of splenic abscess frequently is delayed, with most patients enduring symptoms for 16 to 22 days before diagnosis. Clinical manifestations include fever, left upper quadrant pain, leukocytosis, and splenomegaly in about one-third of patients. The diagnosis is confirmed by ultrasound or CT scan, which has a 95% sensitivity and specificity. Common organisms are aerobic microbes (streptococci and *Escherichia coli*), but other microorganisms have also been isolated (*Mycobacterium tuberculosis* and *Salmonella typhi*). Upon discovery of a splenic abscess, broad-spectrum antibiotics should be initiated, with adjustment to more specific therapy based on culture results and continuation of treatment for a minimum of 14 days. Splenectomy is the operation of choice, but percutaneous and open drainage are options for patients who either cannot tolerate splenectomy or where the clinical scenario warrants splenic preservation, particularly in children. Percutaneous drainage is often successful for patients with unilocular disease and may result in avoidance of splenectomy. Patients with multilocular disease will often require multiple drains and therefore often benefit from total splenectomy.

**Cysts.** Splenic cysts (Fig. 34-5) can be categorized according to a number of criteria; one clinically relevant scheme is to characterize true splenic cysts as either parasitic or nonparasitic. Parasitic Infections. Parasitic infection is the most common cause of splenic cysts worldwide, and the majority are due to *Echinococcus* species. These cysts are more commonly found in areas where the pathogen is endemic. Symptoms, when present, generally are related to the presence of a mass lesion in the left upper quadrant or a lesion that impinges on the stomach. Ultrasound can establish the presence of a cystic lesion and occasionally incidentally detect asymptomatic lesions as well. Serologic testing for echinococcal antibodies can confirm or exclude the cystic lesion as parasitic, an important piece of information when planning operative therapy. Symptomatic parasitic cysts are best treated with splenectomy. Avoidance of spillage of parasitic cyst contents into the peritoneal cavity to avoid the possibility of anaphylactic shock is an important principle in surgical management despite its rare occurrence.

Cysts resulting from trauma are termed *pseudo* cysts due to their lack of an epithelial cell lining. Less common examples of nonparasitic cysts are dermoid, epidermoid, and epithelial cysts. The treatment of nonparasitic cysts depends on whether or not they produce symptoms. Asymptomatic nonparasitic cysts may be observed with close follow-up by ultrasound to exclude significant growth or expansion. Patients should be advised of the risk of cyst rupture with even minor abdominal trauma if they elect nonoperative management for large cysts. Small symptomatic nonparasitic cysts may be excised with splenic preservation, and large symptomatic nonparasitic cysts may be unroofed. Both of these operations may be performed laparoscopically.

**Storage Diseases and Infiltrative Disorders**

**Gaucher’s Disease** Gaucher’s disease is an inherited lipid storage disorder characterized by the deposition of glucocerebrosides in cells of the macrophage-monocyte system. The underlying abnormality is a deficiency in the activity of a lysosomal hydrolase. Abnormal glycolipid storage results in organomegaly, particularly hepatomegaly and splenomegaly. Patients with Gaucher’s disease frequently experience symptoms related to splenomegaly, including early satiety and abdominal discomfort, and to hypersplenism, including thrombocytopenia, normocytic anemia, and mild leukopenia. These latter findings occur as a result of excessive sequestration of formed blood elements in the spleen. Other symptoms in patients with Gaucher’s disease include bone pain, pathologic fractures, and jaundice. Splenectomy may alleviate hematologic abnormalities in these patients with hypersplenism, but it does not correct the underlying disease process.

**Niemann-Pick Disease** Niemann-Pick disease is an inherited disorder of abnormal lysosomal storage of sphingomyelin and cholesterol in cells of the macrophage-monocyte system. Four types of the disease (A, B, C, and D) exist, with unique clinical presentations. Types B and D result from a deficiency in lysosomal hydrolase and are the forms most likely to demonstrate splenomegaly with its concomitant symptoms. Symptoms of splenomegaly are relieved by splenectomy.

**Amyloidosis** Amyloidosis is a disorder of abnormal extracellular protein deposition and protein conformation disorder associated with a clonal plasma cell dyscrasia. There are multiple forms of amyloidosis, each with its own individual clinical presentation, and the severity of disease may range from asymptomatic to multiorgan failure. Patients with primary amyloidosis, associated with plasma cell dyscrasia, have splenic involvement in approximately 5% of cases. Secondary amyloidosis, associated with chronic inflammatory conditions, also may present with an enlarged spleen and even spontaneous rupture. Symptoms of splenomegaly are relieved by splenectomy.

**Sarcoidosis** Sarcoidosis is an inflammatory disease of young adults characterized by noncaseating granulomas in affected tissues. Signs and symptoms of the disease range in severity and typically are nonspecific, such as fatigue and malaise. Any organ system may be involved. The most commonly involved organ is the lung, followed by the spleen. Splenomegaly occurs in approximately 25% of patients. Massive splenomegaly (>1 kg) is rare with a reported incidence of 3% to 6%. Other affected tissues include the lymph nodes, eyes, joints, liver, spleen, and heart. When splenomegaly occurs and causes symptoms related to size or hypersplenism, splenectomy effectively relieves symptoms and corrects hematologic abnormalities such as anemia.
as anemia and thrombocytopenia. Spontaneous splenic rupture has been reported in sarcoidosis.94

**Splenic Artery Aneurysm** Although splenic artery aneurysm is rare, it is the most common visceral artery aneurysm. Women are four times more likely to be affected than men. The aneurysm usually arises in the middle to distal portion of the splenic artery.99,100 Indications for treatment include presence of symptoms, pregnancy, intention to become pregnant, and presence of pseudoaneurysms associated with inflammatory processes. Aneurysm resection or ligation alone is acceptable for amenable lesions in the mid-splenic artery, but distal lesions in close proximity to the splenic hilum should be treated with concomitant splenectomy. An excellent prognosis follows elective treatment.

Splenic artery embolization has been used to treat splenic artery aneurysm; however, this may result in painful splenic infarction or abscess.

**Portal Hypertension** Portal hypertension can result from numerous causes but is usually due to liver cirrhosis. Splenomegaly and splenic congestion often accompany portal hypertension, which leads to sequestration and destruction of circulating cells in the spleen. Splenectomy is not indicated for hypersplenism per se in patients with portal hypertension as there is no correlation between the degree of pancytopenia and long-term survival in these patients.7 In rare circumstances in which splenectomy is required to reduce bleeding from esophageal varices exacerbated by thrombocytopenia, a concomitant splenorenal shunt procedure may be performed to decompress the portal system.

Sinistral hypertension secondary to splenic vein thrombosis, on the other hand, is potentially curable with splenectomy. Patients that are bleeding from isolated gastric varices who have normal liver function test results, especially those with a previous history of pancreatitis, should be examined for splenic vein thrombosis and treated with splenectomy if findings are positive.

Idiopathic portal hypertension (IPH) is extremely rare, and it is characterized by the absence of cirrhosis or other clear etiologies such as hepatic or portal vein thrombosis, cardiac failure, or hematologic disorders.101-104 Hypersplenism often develops in these patients due to hyperactivity of the spleen; defined as a triad of splenomegaly, pancytopenia,
and normocellularity of bone marrow warranting intervention by way of splenic artery embolization or splenectomy. Partial splenic embolization (PSE) has recently been utilized with good success in prohibitively high operative risk patients with severe hypersplenism. 105

**Felty’s Syndrome** The triad of rheumatoid arthritis, splenomegaly, and neutropenia is called *Felty’s syndrome*. It exists in approximately 3% of all patients with rheumatoid arthritis, two-thirds of whom are women. Immune complexes coat the surface of white blood cells, which leads to their sequestration and clearance in the spleen with subsequent neutropenia. This neutropenia (<2000 neutrophils/mm³) increases the risk for recurrent infections and often drives the decision for splenectomy. The size of the spleen is variable, from nonpalpable in 5% to 10% of patients to massively enlarged upwards of 4 times heavier than normal in others. Corticosteroids, hematopoietic growth factors, methotrexate, and splenectomy have all been used to treat the neutropenia of Felty’s syndrome. More recently, rituximab has been tried as a second line agent in refractory Felty’s syndrome. 106,107

Overall response to splenectomy is excellent, with >80% of patients showing a durable increase in white blood cell count. More than one-half of patients who had infections before surgery may clear their infections after surgery. 108 Besides symptomatic neutropenia, other indications for splenectomy include transfusion-dependent anemia and profound thrombocytopenia.

**Wandering Spleen** A very uncommon anatomic abnormality is the “wandering spleen.” In this condition, the spleen “floats” inside the abdominal cavity due to an anomaly during embryogenesis and may present itself in a variety of ways including acute abdomen. 109 The wandering spleen is not normally attached to adjacent viscera in the splenic fossa. This may lead to splenic torsion and infarction. Splenectomy or splenectomy may be required. 110

**Partial Splenectomy and Splenic Salvage.** The increased awareness of asplenia-related life-threatening complications such as overwhelming postsplenectomy infection (OPSI) has led to the development of parenchyma sparing splenic resections for select patients and disorders. The first successful partial splenectomy was reported in 1979 by Shapiro and was followed by Uranus who would perform the first laparoscopic partial splenectomy in 1995. 111,112 Previously, many surgeons had been reluctant to perform partial splenectomy because of the technical difficulties and bleeding risk. However, with a better understanding of the segmental vascular anatomy of the spleen and the development of improved laparoscopic skills and technologies, laparoscopic partial splenectomy has been used successfully in patients with hematologic diseases such as hereditary spherocytosis in children, who may benefit the most from splenic preservation, as well as for benign splenic cysts. 113-119 The technique of partial splenectomy will be discussed later in this chapter.

**IMAGING FOR EVALUATION OF SIZE AND PATHOLOGY**

Thorough assessment of anatomic detail and functional status of the spleen are essential for proper surgical planning. Special preoperative consideration needs to be given to patients with splenomegaly because minimally invasive methods of resection may be fraught with additional difficulty in patients with very large spleens, even in skilled hands. Other indications for splenic imaging include trauma, investigations of left upper quadrant pain, characterization of splenic lesions such as tumors, cysts, and abscesses, and guidance for percutaneous procedures. 140,146

Preoperative imaging of the spleen is primarily performed to obtain an accurate assessment of the splenic volume in order to confirm and document splenomegaly as well as to exclude any large splenic lesion that could affect the surgical resection plane. Identification of the presence of accessory spleens in the preoperative setting is also important, although lack of accessory tissue on the imaging should not preclude a thorough intraoperative search.

The guidelines of the European Association for Endoscopic Surgery suggest that for all patients in whom splenectomy is indicated, preoperative imaging should be obtained. 24 Two of the most commonly used imaging modalities include ultrasound and CT, both enabling measurement of splenic size and volume. When desired, the splenic volume may be calculated using a formula for the volume of a prolate ellipsoid:

\[
\text{Volume (cc)} = \text{length (cm) \times width (cm) \times height (cm) \times 0.5264}.
\]

Other imaging modalities, although not as commonly used, include nuclear medicine studies and magnetic resonance imaging (MRI).

**PREOPERATIVE CONSIDERATIONS**

As part of preoperative discussion prior to splenectomy, patients should be consulted on potential complications associated with this procedure, including overwhelming postsplenectomy sepsis, splenic vein thrombosis, bleeding, arterial thrombosis (myocardial infarction, stroke), deep vein thrombosis, and pulmonary hypertension.

**Vaccination and Patient Education**

Considering that infection is the most common complication, patient education and vaccinations against encapsulated pathogens are the mainstay of preventive therapy. 52,120 Although rare, the most feared and extreme infectious complication is overwhelming postsplenectomy sepsis (OPSI). (See later section, “Overwhelming Postsplenectomy Infection,” for detailed discussion.) Patients undergoing splenectomy for hematologic or malignant indications have the greatest risk, whereas patients who undergo splenectomy for trauma or iatrogenic injury have the lowest risk. OPSI is more common in the pediatric population, with 4.4% of children less than 16 years of age versus 0.9% of adults developing this life-threatening condition. The risk has been observed to be the greatest in the first 2 years after splenectomy; however, asplenic patients remain at lifelong risk. 121-123 Considering that the spleen is the site for special adaptation of macrophages that target encapsulated organisms, asplenic patients are at higher risk of infection caused by *Streptococcus pneumoniae* (responsible for >50% of OPSI), *H influenzae* type b, *Neisseria meningitidis*, and *Capnocytophaga canimorsus* (transmitted by dog bites). 124

In the setting of elective splenectomy, patients should be vaccinated two weeks prior to surgery to optimize antigen recognition and processing. If splenectomy is performed emergently, vaccinations can be administered postoperatively and consideration should be given to delaying administration...
for 2 weeks to avoid the transient immunosuppression associated with surgery. International guidelines also recommend annual influenza vaccine for asplenic patients. The influenza vaccination provides protection from influenza syndrome and secondary bacterial infection. This immunization is associated with a 54% reduced risk of death compared with unimmunized asplenic persons.125

Preoperative and postoperative patient education regarding OPSI is paramount because patients with OPSI may rapidly progress from a febrile illness to circulatory collapse and death within a matter of hours. In one study, 28% of asplenic patients were unaware of the potential infection risks, and the main reasons for this lack of awareness were that correct advice was not given or that that advice was forgotten.121,126 The use of currently available vaccines against pneumococcus and other encapsulated organisms has led to a drop in the overall incidence of OPSI to <1%. The mechanism by which vaccination protects asplenic patients is not entirely understood. Serum antibody titers do not necessarily correspond to clinical immunity. Moreover, antibody levels after pneumococcus vaccination decline steadily within 5 to 10 years. Revaccination is reasonably recommended for these patients, although the efficacy of this measure is not proven.121-124,126

**Deep vein Thrombosis Prophylaxis**

Deep vein thrombosis (DVT) after splenectomy is not infrequent, especially in cases involving splenomegaly and myeloproliferative disorders.76 Risk of portal vein thrombosis (PVT) may reach 50% for patients presenting with both splenomegaly and myeloproliferative disorders.127,128

Postsplenectomy PVT typically presents with anorexia, abdominal pain, leukocytosis, and thrombocytosis. Effective PVT treatment is possible by maintaining a high index of suspicion, achieving early diagnosis with contrast enhanced CT, and starting anticoagulation immediately. DVT prophylaxis, including use of sequential compression devices and subcutaneous administration of heparin (5000 U), should be initiated for patients undergoing splenectomy.16,77 Each patient’s risk factors for DVT should be evaluated, and when elevated risk exists (obesity, history of prior venous thromboembolism, known hypercoagulable state, older age), a postoperative antithrombotic regimen of up to two weeks of low molecular weight heparin should be maintained.

**Splenectomy Techniques**

**Patient Preparation**

Assessment of the potential need for transfusion of blood products and optimization of preoperative coagulation status are necessary. It is the authors’ practice to order blood typing and antibody screening tests for normosplenic patients undergoing elective splenectomy. Anemic patients should be transfused before surgery to a hemoglobin level of 10 g/dL. In more complex cases, including patients with splenomegaly, at least 2 to 4 units of cross-matched blood should be available at the time of surgery. Thrombocytopenia may be transiently corrected with platelet transfusions. Thrombocytopenic patients preferably should not undergo transfusion before the day of surgery and ideally not before the intraoperative ligation of the splenic artery. Several authors recommended a platelet count of $30 \times 10^9$ /L before the surgery; this may require treatment with IV immunoglobulin or oral corticosteroids if the platelets are below this number. Pooled normal human immunoglobulin is effective in elevating the platelet count in approximately 75% of patients.129-131

Patients who have been maintained on corticosteroid therapy preoperatively should receive parenteral corticosteroid therapy perioperatively. Bowel preparation is not routinely performed for patients undergoing elective splenectomy. All splenectomy patients do receive DVT prophylaxis, as discussed previously. After endotracheal intubation, a nasogastric (NG) tube is inserted for stomach decompression.

**Open Splenectomy**

Although laparoscopic surgery increasingly has achieved acceptance as the standard approach for normosplenic patients requiring splenectomy, open splenectomy (OS) is still widely practiced. The largest published series is a report of the Nationwide Inpatient Sample (NIS), where of 37,006 nontraumatic splenectomies identified during a 6-year study period (2005–2010), 81.4% of those cases were approached by open surgery.132

Traumatic rupture of the spleen continues as the most common indication for OS. Several other clinical scenarios favor an OS approach, including massive splenomegaly, ascites, portal hypertension, multiple prior operations, extensive splenic irradiation, and possible splenic abscess.

During OS, the patient is placed in the supine position with the surgeon situated at the patient’s right. A left subcostal incision paralleling the left costal margin and lying two finger-breadths below it is preferred for most elective splenectomies. A midline incision is optimal for exposure when the spleen is ruptured or massively enlarged. The spleen is mobilized by dividing ligamentous attachments, usually beginning with the splenocolic ligament (Fig. 34-6). In patients with significant splenomegaly, once lesser sac access has been achieved through either the gastrosplenic or gastrohepatic attachments, ligating the splenic artery in continuity along the superior border of the pancreas may be preferable. This maneuver may serve several purposes: allowing safer manipulation of the...
spleen and dissection of the splenic hilum, facilitating some shrinkage of the spleen, and providing an autotransfusion of erythrocytes and platelets. Further medial mobilization of the spleen is achieved by incising its lateral peritoneal attachments, most notably the splenophrenic ligament. Then follows individual ligation and sequential division of the short gastric vessels, steps that if carefully executed reduce the risk of these vessels’ retracting and bleeding. Splenic hilar dissection then takes place. Whenever possible, care should be taken to dissect and individually ligate the splenic artery and vein (in that order) before dividing them. As noted in the discussion of splenic anatomy, the tail of the pancreas lies within 1 cm of the splenic hilum in 75% of patients; therefore, during hilar dissection, great care must be taken to avoid injuring the pancreas. Once the spleen is excised, hemostasis is secured by irrigating, suctioning, and scrupulously inspecting the bed of dissection. The splenic bed is not routinely drained. A thorough search for accessory spleens must be undertaken when a hematologic disorder has occasioned splenectomy. At the completion of surgery, the nasogastric tube is removed.

**Laparoscopic Splenectomy**

Laparoscopic splenectomy (LS) has become the procedure of choice over the last two decades, since it has been first described Delaitre and Maignien in 1991. In fact, LS is now the gold standard for elective splenectomy in patients with normal-sized spleens. In experienced hands, LS is associated with decreased intraoperative blood loss, shorter hospital length of stay, and lower morbidity rates when compared to OS. Since the introduction of the lateral approach, most LS procedures are now performed with the patient in the right lateral decubitus position (Fig. 34-7). A midway “double-access” technique in which the patient is in a 45° right lateral decubitus position has also been advocated. This positioning permits concomitant surgery, such as laparoscopic cholecystectomy, more easily than does the lateral approach. The double-access technique requires the placement of five or six trocars. The lateral approach routinely involves the use of three or four trocars positioned as shown in Fig. 34-7. Use of an angled (30° or 45°) laparoscope (2, 5, or 10 mm) greatly facilitates the procedure. Exposure of the vital anatomy in a manner that allows for a more intuitive sequence of dissection, paralleling that of OS, may be considered an additional advantage of the lateral approach.

Placement of trocars in the left upper quadrant should be performed under laparoscopic visualization, particularly if any degree of splenomegaly exists, because the latter can significantly reduce the available operating space. As with OS, the splenocolic ligament and the lateral peritoneal attachments are divided with resultant medial mobilization of the spleen. The short gastric vessels may be divided usually with hemostatic energy sources such as ultrasonic dissection, diathermy, or radiofrequency ablation. With the lower pole of the spleen gently retracted, the splenic hilum is accessible to further applications of clips or an endovascular stapling device. The splenic artery and vein are divided separately when possible. Good long-term outcomes, however, are increasingly being achieved with mass hilar stapling (Fig. 34-8). Using the lateral approach with the spleen thus elevated, the surgeon can easily visualize the tail of the pancreas and avoid injury when placing the endovascular stapler within the sack and allows piecemeal extraction; a blunt instrument should be used to disrupt and remove the spleen to avoid the risk of sack rupture, spillage of contents, and subsequent splenosis (Fig. 34-9 and Fig. 34-10).

**Hand-Assisted Splenectomy**

When LS is performed in patients with splenomegaly, there have been reports of high rates of both complications and conversion to open splenectomy. Hand-assisted laparoscopic surgery (HALS) has been described as an alternative to the LS approach. Spleens greater than 22 cm in craniocaudal length or 19 cm in width may benefit from HALS over a purely laparoscopic approach. It has been reported that the

---

**Figure 34-7.** Patient positioning and trocar placement for laparoscopic splenectomy.
use of this technique has resulted in a marked reduction in average operative time for patients with massive splenomegaly (146 vs. 295 minutes). Although HALS does require a small incision (7–8 cm) for hand insertion and specimen extraction, no differences in length of stay were observed when comparing these patients to those managed purely laparoscopically.

When performing HALS splenectomy, patient positioning is similar to that of LS (Fig. 34-11). For patients with massive spleens, lateral positioning is altered slightly, such that the patient is placed supine with the left side elevated at 45°. Depending on the hand dominance of the surgeon, the hand-assist device can be placed in either a midline position for right-handed dominant or a subcostal position for left-hand dominant surgeons. A 7- to 8-cm incision should be made 2 to 4 cm caudal to the inferior pole of the enlarged spleen. The hand-assisted technique allows for a tactile feedback and atraumatic manipulation of the enlarged spleen.

The nondominant hand provides medial retraction and rotation of the spleen through a hand-assist port, while the dominant hand carries out the dissection using laparoscopic instruments. The anterior and posterior attachments of the spleen are taken down as in the laparoscopic approach, and the hilar pedicle is ligated using an endoscopic vascular stapler.

**Single-Incision Laparoscopic Surgery Splenectomy**

Single-incision laparoscopic surgery (SILS) splenectomy emphasizes the concept of surgery through one small transabdominal incision rather than the standard multiple trocar sites, with theoretical benefits of less pain and better cosmetics. The incision can be hidden periumbilically and is used as the specimen extraction site at the end of the case. This approach for solid organs poses several technical challenges. Since all instruments are closely aligned together, “fencing” of instruments and the laparoscope and limited degrees of movement are commonly encountered. The spleen, being a solid organ, cannot be grasped, and thus retraction may be more challenging in these cases. Furthermore, it has been reported that periumbilical port position may result in technical challenges when dealing with high body mass index or tall patients, precluding

**Figure 34-8.** Splenic hilum can be divided laparoscopically en masse once the spleen has been rotated medially having been mobilized from its lateral attachments.

**Figure 34-9.** Spleen extraction. A. Spleen is placed into a ripstop nylon bag before morcellation. B. Splenic morcellation.
the surgeon from adequately reaching the spleen. Other alternatives to single port placements have been reported, although to date, no proven benefits of SILS splenectomy have been demonstrated.\textsuperscript{140-142}

**Robotic Splenectomy**

The da Vinci surgical robot (Intuitive Surgical, Sunnyvale, CA) was cleared by the U.S. Food and Drug Administration (FDA) for use in humans in the year 2000 and has been applied to clinical practice in a variety of abdominal procedures including splenectomy since 2002.\textsuperscript{143,144} The term robotic surgery, referencing the da Vinci device should more accurately be described as computer assisted surgery (CAS), as it requires a surgeon sitting at a console controlling distant end effectors. A truly robotic, automated system has not yet been devised nor deployed to perform splenectomy. The reported advantages of CAS have inspired many surgeons to investigate its potential and to broaden its application in the minimally invasive surgery armamentarium. Some of these advantages include increased degrees of freedom as compared to standard “straight-stick” laparoscopy, improved optics including three-dimensional imaging of the operative field, improved instrument stabilization and reduction in hand tremor, and finally purported enhanced ergonomic and comfort factors for the operating surgeon.\textsuperscript{145-147}

However, detractors of this emerging technology cite high capital expenses as well as on-going disposable costs and the loss of haptic feedback as major downsides of CAS such that the benefits of the robotic platform are still not fully clear in the extant literature. Recent studies suggest that ergonomic benefits of robotic surgery may not be as pronounced as previously thought, as “robotic surgeons” still manifested chronic pain related to poor ergonomics.\textsuperscript{148}

There have been few published works comparing conventional laparoscopic splenectomy to CAS splenectomy. In one recent retrospective case-matched analysis, the authors compared hospital length of stay, operating room times and cost between these two groups. Although they were able to conclude that the application of CAS was feasible and safe for splenectomy, they cited a nearly 30-minute increase in operative time and over $4000 increase in cost of the procedures.\textsuperscript{149} This is consistent with a prior study supporting the notion that robotic-assisted splenectomy takes longer and is more costly than conventional laparoscopic surgery.\textsuperscript{150}

Splenectomy has also been reported using a single-incision CAS approach.\textsuperscript{151} However, this procedure is likely associated with an even steeper learning curve than traditional multiport robotic-assisted splenectomy, which calls into question the utility of this technique.

Robotic splenectomy has been performed in the pediatric population, also proving feasibility and safety, but much like the adult population, larger series are needed to determine the true benefits of this procedure as compared to conventional laparoscopic splenectomy in children.\textsuperscript{152}

The role of CAS in the patient requiring more complex surgery, for example, partial splenectomy or resection of a massive spleen, is yet to be determined. At present there appears to be no argument supporting the application of CAS over standard laparoscopy for routine splenectomy in terms of clinical or cost advantage. The entry into the market of competing devices is likely to favorably alter these considerations.

**Partial Splenectomy**

The past few decades have witnessed ever-widening endorsement for and practice of partial splenectomy. This technique, initially reported in the early 18th century, is particularly indicated to minimize the risk of postsplenectomy sepsis in children. Certain lipid storage disorders leading to splenomegaly (e.g., Gaucher’s disease), some forms of traumatic splenic injury (blunt and penetrating), spherocytosis in children, and focal benign splenic lesions are amenable to treatment with partial splenectomy.\textsuperscript{153} Both the laparoscopic and open approaches for partial splenectomy have been well described.\textsuperscript{153} The spleen must be adequately mobilized, and the splenic hilar vessels attached to the targeted segment, ligated, and divided. The devascularized segment of spleen is transected along an obvious line of demarcation.

A useful technical tip is to transect the parenchyma 1 cm inside the ischemic demarcation line to minimize blood loss.\textsuperscript{154} Bleeding from the cut surface of the spleen usually is limited and can be controlled by various methods, including cautery, argon coagulation, or application of direct hemostatic agents such as cellulose gauze and fibrin glue.

**Inadvertent Intraoperative Splenic Injury**

Inadvertent intraoperative injury to the spleen is a noted occurrence in the surgical literature, familiar to and dreaded by the abdominal surgeon. The true incidence is unknown. The gravity of such injury is not to be underestimated. Significant short-term morbidity is associated with injury to the spleen, including increased blood loss, need for transfusion, and prolonged hospital stay.\textsuperscript{155}

Intraoperative injury to the spleen has been linked with numerous operations, such as gastric fundoplication, colectomy,
Figure 34-11. A. Patient table placement for hand-assisted laparoscopic splenectomy (HALS) in case of splenomegaly. B and C. Intraoperative images of HALS.
paraesophageal hernia repair, nephrectomy, and abdominal and pelvic vascular surgery. There are also reports of splenic injuries after endoscopic procedures, such as colonoscopy.

Improper traction on the spleen against its peritoneal attachments is the most common mechanism of intraoperative injury. Capsular tears are the most common type of injury, but parenchymal lacerations and subcapsular hematomas also occur. The lower pole is more commonly injured, owing to its orientation and the greater concentration of peritoneal attachments found here.

As with all hemorrhage, prompt temporary control of bleeding should be obtained by direct compression of the spleen itself, packing of the left upper quadrant, compression of the vessels at the splenic hilum, or pressure on the splenic artery at the superior pancreatic margin. The spleen should then be mobilized from its peritoneal attachments and the nature of the injury assessed. Overall, the patient’s condition is the primary determinant of whether splenic salvage can be attempted, although hilar injury is best managed by splenectomy. When dealing with capsular tears (most common injury), strong consideration should be given to splenorrhaphy techniques: application of topical hemostatics, suture plication of disrupted parenchyma with or without omental buttress, and the use of bioabsorbable mesh sheets.

The time-honored surgical tenets of liberal exposure and visualization are particularly germane to the avoidance of splenic injury. Incisions and approaches must be tailored to both patient circumstances and surgeon experience. There is some evidence to support the assertion that use of the laparoscopic approach may reduce the incidence of splenic injury for certain operations. As with all hemorrhage, prompt temporary control of bleeding is required. Direct compression of the spleen itself, packing of the left upper quadrant, compression of the vessels at the splenic hilum, or pressure on the splenic artery at the superior pancreatic margin can slow or stop hemorrhage and allow more deliberate consideration of management options.

The type of injury plays a role as well; it has been suggested that hilar injury is best managed by splenectomy.84 Barring these unfavorable circumstances, however, and recalling that the majority of intraoperative splenic injuries are capsular tears, it is reasonable to expect that splenic preservation can be achieved in many appropriately selected situations. Presented with one of these situations, the surgeon has at his or her disposal a number of useful and well-described splenorrhaphy techniques: application of topical hemostatics, suture plication of disrupted parenchyma with or without omental buttress, and the use of bioabsorbable mesh sheets.

Currently, incidental splenectomies during laparoscopic procedures such as colorectal resections are rare events, but they are associated with worse short-term outcomes.156

Preoperative Grading Score to Predict Technical Difficulty in Laparoscopic Splenectomy

A splenectomy grading system based on preoperative parameters was developed to predict the surgical difficulty and morbidity for elective laparoscopic splenectomies.157,158 Preoperative data concerning demographic, clinical, pathological, anatomical, laboratory and radiological factors were compared with three surgical outcomes: operative time, intraoperative bleeding, and surgical conversion. Four preoperative parameters (male gender, age, type of pathology, and spleen weight) were found to be associated with a difficult splenectomy (Table 34-3). This grading score is simple to calculate from the physical examination, laboratory tests, and US or CT images and could be highly practical in a daily clinical setting. It could facilitate training and development of skills while simultaneously fostering dissemination of laparoscopic procedures.

### Table 34-3

<table>
<thead>
<tr>
<th>Difficulty Score</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>≤40 years</td>
</tr>
<tr>
<td></td>
<td>40–60 years</td>
</tr>
<tr>
<td></td>
<td>≥60 years</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Male</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pathology group</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITP</td>
</tr>
<tr>
<td>Other benign</td>
</tr>
<tr>
<td>Malignant</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Spleen weight*</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;400 gr</td>
</tr>
<tr>
<td>400–1000 gr</td>
</tr>
<tr>
<td>&gt;1000 gr</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Difficulty grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
</tr>
<tr>
<td>Medium</td>
</tr>
<tr>
<td>High</td>
</tr>
</tbody>
</table>

Note: Minimum possible score: 2 points; maximum possible score: 10 points.

*Spleen weight formula: width (cm) × length (cm) × height (cm) × 0.6 = splenic weight in grams.

Splenectomy Outcomes

Changes in blood composition resulting from splenectomy include the appearance of Howell-Jolly bodies and siderocytes. After splenectomy, leukocytosis and increased platelet counts are common as well. Although platelet counts most often rise within 2 days, they may not peak for several weeks in patients with preoperative thrombocytopenia (see “Hematologic Outcomes” later). Similarly, within 1 day after splenectomy, the white blood cell count typically rises, and such elevation may continue for several months.

Overwhelming Postsplenectomy Infection

The prevalence of asplenia in the United States is estimated to be 1 million; which is comparable to the number of patients carrying the human immunodeficiency virus (HIV).185,186 As with other forms of immunodeficiency, asplenic patients bear an increased susceptibility to specific types of infections for the remainder of their lives. Asplenic patients are at highest risk for infection with encapsulated organisms, most commonly *Streptococcus pneumoniae*, but also *Haemophilus influenzae*.
Complications
Complications of splenectomy may be classified as pulmonary, hemorrhagic, infectious, pancreatic, and thromboembolic. Left lower lobe atelectasis is the most common complication after OS; pleural effusion and pneumonia also can occur. Hemorrhage can occur intraoperatively or postoperatively, presenting as subphrenic hematoma. Transfusions have become less common since the advent of LS, although the indication for operation influences the likelihood of transfusion as well. Subphrenic abscess and wound infection are among the perioperative infectious complications. The placement of a drain in the left upper quadrant may be associated with postoperative subphrenic abscess and is not routinely recommended. Pancreatitis, pseudocyst, and pancreatic fistula are among the pancreatic complications that may result from intraoperative trauma to the pancreas during dissection of the splenic hilum.

Hematologic Outcomes
The results of splenectomy may be appraised according to the level of hematologic response (e.g., rise in platelet and hemoglobin levels) in those disorders in which the spleen contributes to the hematologic problem. Hematologic responses may be divided into initial and long-term responses. For thrombocytopenia, an initial response typically is defined as a rise in platelet count within several days of splenectomy. Reported series demonstrate the effectiveness of LS in providing a long-term platelet response in approximately 80% of individuals with ITP (Table 34-4). These results are consistent with the long-term success rate associated with OS.

For chronic hemolytic anemias, a rise in hemoglobin levels to >10 g/dL without the need for transfusion signifies a successful response to splenectomy. By this criterion, splenectomy has been reported to be successful for the vast majority of patients with chronic hemolytic anemia. For hemolytic anemia due to spherocytosis, the success rate is usually higher, ranging from 90% to 100%.

Splenectomy results also may be examined in terms of surgical and postsurgical characteristics, including operative time, recovery time, and morbidity and mortality rates, all of which tend to vary according to hematologic indication (see Tables 34-4 and 34-5).

Results of few prospective, randomized trials comparing LS and OS have been published. However, several recent meta-analyses of published comparative series including 38 papers with more than 2914 patients indicate that the laparoscopic approach typically results in longer operative times, shorter hospital stays, lower morbidity rates, similar blood loss, and similar mortality rates compared with OS. Questions of the cost effectiveness of LS persist, although analysis of this issue is hindered by a lack of universally accepted metrics as well as a paucity of recent objective data. Proponents of LS argue that the generally higher operating room charges are offset by the reduced hospital stay and presumably shorter time of lost productivity. For those institutions with experienced personnel and technical capability, the laparoscopic approach has emerged as the standard for elective, nontraumatic splenectomy.

Cancer
A Taiwanese population-based study found that individuals who had splenectomy have higher risks of developing certain types of cancer (adjusted hazard ratios were 2.64 and 1.29 for nontraumatic and traumatic reasons, respectively). Splenectomy patients were found to have significantly higher risks in esophagus, stomach, liver, other head and neck, non-Hodgkin’s lymphoma, and leukemia cancers. Although the exact mechanism for the possible association between splenectomy and cancer remains unclear, a plausible explanation is that the spleen is thought to be involved in immunological defenses and provides active response through humoral and cell-mediated pathways and that splenectomy may impair immune surveillance in the host.

Table 34-4

<table>
<thead>
<tr>
<th>Outcome after splenectomy</th>
<th>Studies (N)</th>
<th>Patients (N)</th>
<th>Pooled Results and CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
<td>38</td>
<td>2914</td>
<td>−0.01 (−0.02, 0)</td>
<td>1</td>
</tr>
<tr>
<td>Complications</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pooled</td>
<td>38</td>
<td>2914</td>
<td>−0.11 (−0.16, −0.05)</td>
<td>&lt;.00001</td>
</tr>
<tr>
<td>Minor</td>
<td>36</td>
<td>2914</td>
<td>−0.03 (−0.05, −0.01)</td>
<td>.13</td>
</tr>
<tr>
<td>Severe</td>
<td>36</td>
<td>2745</td>
<td>−0.07 (−0.11, −0.03)</td>
<td>&lt;.00001</td>
</tr>
<tr>
<td>Thrombosis</td>
<td>35</td>
<td>2695</td>
<td>−0.01 (−0.02, 0.01)</td>
<td>1</td>
</tr>
<tr>
<td>Organ injury</td>
<td>34</td>
<td>2639</td>
<td>0.01 (−0.02, 0)</td>
<td>1</td>
</tr>
<tr>
<td>Acces. spleen</td>
<td>29</td>
<td>2135</td>
<td>0.02 (−0.01, 0.05)</td>
<td>.87</td>
</tr>
<tr>
<td>Operative time</td>
<td>18</td>
<td>1370</td>
<td>57.4 min (43.3, 71.4)</td>
<td>&lt;.00001</td>
</tr>
<tr>
<td>Blood loss</td>
<td>10</td>
<td>759</td>
<td>−41 mL (−87, 4.71)</td>
<td>&lt;.00001</td>
</tr>
<tr>
<td>Length of stay</td>
<td>20</td>
<td>1566</td>
<td>−2.48 d (−2.89, −2.07)</td>
<td>&lt;.00001</td>
</tr>
</tbody>
</table>

Ultrasound

Ultrasound is the least invasive mode of splenic imaging. It is rapid, relatively easy to perform, and does not expose the patient to ionizing radiation. It is often the first imaging modality used to evaluate the spleen in a trauma patient, although questions of sensitivity and specificity remain. In the elective setting, such as for routine diagnostic purposes or for preoperative planning, it is the least costly modality available, and the sensitivity of ultrasound for detecting textural lesions of the spleen can be quite good in experienced hands. When examining a normal spleen, differentiation between red and white pulp, the spleen appears heterogeneously enhancing due to the possibility of segmental blood flow and its effect on area of resection.

Percutaneous ultrasound-guided procedures for splenic disease (e.g., cyst aspiration, biopsy), historically avoided due to the risk of hemorrhage and other complications, are becoming more common as the safety of these procedures has become increasingly demonstrated. Computed Tomography

CT affords a high degree of resolution and detail of the splenic parenchyma, vasculature, and its relationship to neighboring structures, making it the preferred imaging modality for many surgeons. CT has become an invaluable tool in the evaluation and management of the blunt trauma patient, and standardized scoring systems for splenic trauma based on CT images now aid in management decisions. In the nontrauma setting, CT is extremely useful for assessment of splenomegaly, identification of solid and cystic lesions, and guidance of percutaneous procedures. The use of iodinated contrast material adds diagnostic clarity to CT imaging of the spleen, although at the cost of, the small but real risks, possible renal impairment or allergic reactions. Three-dimensional reconstruction after CT scan may help to predict the difficulty of the procedure and to choose the best surgical approach.

The appearance of normal splenic tissue on a noncontrast CT is uniform parenchymal attenuation with values ranging between 40 and 60 Hounsfield units (HU). On a contrast-enhanced CT, the appearance of the spleen depends largely on the timing of the intravenous bolus administration of contrast material. Due to the different rates of flow through the red and white pulp, the spleen appears heterogeneously enhancing during the first minute after initiation of intravenous administration of contrast material during the arterial and early portal venous phases. The frequency of these artifacts increases with advancing patient age. When evaluating for splenic abscess, a contrast-enhanced CT should be utilized.

Table 34-5

<table>
<thead>
<tr>
<th>Laparoscopic splenectomy results by hematoologic indication</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR time (min)</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>EBL (mL)</td>
</tr>
<tr>
<td>LOS (days)</td>
</tr>
<tr>
<td>Conversions from LS to OS</td>
</tr>
<tr>
<td>Complications</td>
</tr>
</tbody>
</table>

EBL = estimated blood loss; ITP = idiopathic thrombocytopenic purpura; LOS = length of hospital stay; LS = laparoscopic splenectomy; OR = operating room; OS = open splenectomy; TTP = thrombotic thrombocytopenic purpura.


Plain Radiography

Rarely is plain radiography used for primary splenic imaging. Plain films can indirectly provide an outline of the spleen in the left upper quadrant or suggest splenomegaly by revealing displacement of adjacent air-filled structures (e.g., the stomach or splenic flexure of the colon). Plain films may also demonstrate splenic calcifications. Splenic calcifications are often found in association with splenomegaly but are otherwise a nonspecific finding. Splenic calcifications can indicate a number of benign, neoplastic, or infectious processes, including phlebolith, splenic artery aneurysm, sickle cell changes, tumors (e.g., hemangioma, hemangiosarcoma, lymphoma), echinococcosis, or tuberculosis.

Magnetic Resonance Imaging

Although MRI offers excellent detail and versatility in abdominal imaging, it is more expensive than CT scan or ultrasound and offers no obvious advantage for primary imaging of the spleen. MRI can be a valuable adjunct to the more commonly used imaging techniques when splenic disease is suspected but not definitively diagnosed.

Magnetic resonance (MR) signal characteristics of the spleen are related to the relative ratio of red and white pulp and the relationship of the timing of the intravenous (IV) contrast bolus and the time of image acquisition. The spleen will generally have a homogeneous MR signal on noncontrast images.
On contrast MRI, the spleen appears to have heterogeneous enhancement during the arterial phase of contrast enhancement.

**Angiography**

Angiography of the spleen most commonly refers to invasive arterial imaging, and when it is combined with therapeutic splenic arterial embolization (SAE), there are multiple applications for this procedure: localization and treatment of hemorrhage in select trauma patients; delivery of a variety of therapies in patients with cirrhosis or portal and sinistral hypertension and in transplant patients; and adjunct (or, more controversially, as an alternative) to splenectomy for treatment of hematologic disorders such as ITP or hypersplenism. Preoperative or intraoperative SAE for elective splenectomy is also a common, although not universal, practice. Few prospective data have been published in the last 5 years on preoperative SAE. Preoperative SAE is purported not only to facilitate less intraoperative blood loss but also possibly to allow a laparoscopic approach in patients whose spleens had previously been considered too large for, or otherwise not amenable to, safe laparoscopic resection. Limited success in using partial SAE as an alternative to therapeutic splenectomy in chronic ITP has been previously reported. Its detractors argue that the need for increased analgesics and occasional extended hospital stay preoperatively, the possibility of pancreatitis, and the well-described risks of invasive arteriography associated with the passage of wires and catheters through the vasculature, may negate any presumed benefits of preoperative SAE.

**Nuclear Imaging**

Radioscintigraphy with technetium-99m sulfur colloid demonstrates splenic location and size. It may be especially helpful in locating accessory spleens after unsuccessful splenectomy for ITP and has recently proven useful in diagnosing splenosis. Unfortunately, no conclusive outcome benefit has been shown for preoperative technetium scanning before splenectomy. When dealing with diseases of platelet sequestration, indium-labeled autologous platelet scanning (ILAPS) demonstrates whether platelet sequestration is predominantly in the spleen, liver, or both. This becomes important in deciding whether or not a patient will benefit from a splenectomy. ILAPS is a nuclear imaging modality in which autologous platelets are infused into the patient after ex vivo labeling. Subsequent scintigraphy demonstrates the site(s) of platelet sequestration and clearance. It has been proposed that patients with purely or predominantly splenic sequestration determined by ILAPS may be more likely to respond to splenectomy than those exhibiting hepatic, mixed, or diffuse patterns.

An emerging and novel application for spleen scintigraphy may be as a noninvasive method to diagnose nonalcoholic steatohepatitis (NASH). Conventional imaging methods are reliable for the detection of moderate to severe fatty changes in the liver, though they are not reliable for detecting NASH or hepatic fibrosis. NASH, which may lead to cirrhosis, can result from nonalcoholic fatty liver disease (NAFLD), the most common cause of steatosis. With the alarming rise of obesity worldwide, NAFLD is also increasingly common with prevalence ranging from 6.3% to 33%. The diagnosis of progression from NAFLD to NASH has been dependent on histologic assessment of tissue obtained from liver biopsy. Characteristics unique to NASH have been reported, among them the association of splenic enlargement, not seen to a similar degree in NAFLD.

In addition, the ratio of liver-to-spleen uptake determined by scintigraphy has been found to be predictably altered in NASH patients. The liver-to-spleen uptake ratio is significantly decreased in NASH patients, but not NAFLD patients, leading some to conclude that technetium-99m-phytate scintigraphy is a reliable tool to differentiate NASH from NAFLD. Although additional studies are needed to identify the role of these nuclear medicine studies in prognostication and monitoring of those patients at high risk for the development of NASH, recent studies show promise in this regard.

**Microbiology and Pathogenesis.** Life-threatening infection in the asplenic patient is attributable to four main factors: loss of splenic macrophages, diminished tuftsin production, loss of the spleen’s reticuloendothelial screening function, and dysregulated coagulation. In the normal host, these factors work in concert to eliminate opsonized bacteria from the bloodstream. This system is particularly suited for the removal of encapsulated bacteria, whose polysaccharide coating is a natural defense against opsonization (such as S. pneumoniae, H influenzae, and N meningitidis are the classic examples). Infections with protozoa that invade the red blood cell, such as Babesia microti (transmitted by tick bites), Ehrlichia, and Plasmodium, occur more frequently in splenectomized individuals than in normal hosts. Other potential infectious bacterial sources include group A streptococci, C canimorsus (transmitted by dog bites), group B streptococci, Enterococcus species, Bacteroides species, Salmonella species, and Bartonella species. In the absence of the spleen, elimination of these pathogens from the bloodstream falls solely to the liver, a process that has been demonstrated to be less effective. Further, the pathophysiology of infection in asplenic patients has also been implicated in their increased risk of thrombosis and pulmonary hypertension.

More recently, the bacterial patterns of splenectomy sepsis have been changing. After the introduction of vaccinations and new oral antibiotics, postsplenectomy patients can suffer from diverse strains of bacterial infection, which are not strictly correlated with the splenic function. In recent cohort series, gram negative bacteria are prevalent, representing 45% to 50% of infections in asplenic patients. In vaccinated patients, the rate of sepsis by pneumococcus is very low. In fact, encapsulated bacteria, such as S. pneumoniae, N meningitidis, and H influenzae, were rarely encountered in those series in whom vaccination was routinely adopted.

Sepsis by uncommon bacteria as well by protozoa infections such as malaria and babesiosis are also known to affect asplenic patients.

**Clinical Features.** OPSI is uniformly fatal without treatment, and thus sepsis in a splenectomized patient is a medical emergency. Therefore, any clinical suggestion of infection, including seemingly isolated fevers, must be viewed with a high index of suspicion and treated empirically as thorough investigation proceeds. OPSI may begin with a relatively mild-appearing prodrome of symptoms. In addition to fever, nonspecific symptoms such as malaise, myalgias, headache, vomiting, diarrhea, abdominal pain, and others should be viewed with alarm in the asplenic patient. This process can progress rapidly to fulminant bacteremic septic shock, with hypotension, anuria, and disseminated intravascular coagulation.

The true incidence of OPSI is not precisely known because defining criteria vary among published series. Overall lifetime risk remains low, ranging from <1% to 5%. Among
those who develop OPSI, some characteristics can be identified that impart greater risk. Reason for splenectomy is the single most influential determinant of OPSI risk. Case series demonstrate that those who undergo splenectomy for hematologic disease (malignancy, myelodysplasia, or hemoglobinopathy) are far more susceptible to OPSI than patients who undergo splenectomy for trauma or iatrogenic reasons. Age is also an important consideration, with children 5 years of age or less and adults 50 years or older being at elevated risk. Finally, time interval from spleen removal must be considered. A large number of OPSI cases occur many years to decades after splenectomy.\textsuperscript{186,189} This observation underscores both the threat of this lethal disease and the need for lifelong vigilance.

**Antibiotics and the Asplenic Patient.** Antibiotic therapy for the asplenic patient can be considered in three contexts: therapy for established or presumed infections, prophylaxis in anticipation of invasive procedures (e.g., dental procedures), and general prophylaxis. The most critical action in the treatment of established or presumed OPSI is the immediate use of broad-spectrum intravenous antibiotics, ideally after the collection of blood cultures. Vancomycin provides broad-spectrum Gram-positive coverage, including coverage against penicillin-resistant *S pneumoniae*.\textsuperscript{190} Ceftriaxone should be added to include Gram-negative coverage for *N meningitidis* and *H influenzae*.\textsuperscript{190} Early implementation of antibiotics and goal-directed therapy for sepsis can significantly reduce mortality rates.\textsuperscript{21,191} For the latter two indications, unfortunately, evidence supporting efficacy is scant, and guidelines for antibiotic prophylaxis are not uniform. Optimal duration of chemoprophylaxis in children also remains unclear; however, a daily dose of antibiotics until 5 years of age or at least 5 years after splenectomy are commonly recommended, although some advocate continuation into at least early adulthood.\textsuperscript{121,122,192} Concerns regarding compliance and bacterial resistance have been raised, which have led some authors to suggest that lifelong daily antibiotic prophylaxis be recommended only for those patients whose antibody titers fail to respond appropriately to vaccination or, alternately, that asplenic patients be advised to carry at all times a reserve supply of antibiotic to be self-administered at the earliest sign of infection.\textsuperscript{122} Considering the grave consequences of OPSI and its relatively low incidence, controlled trials resulting in meaningful data on this issue seem unlikely to be performed.

**Education.** Several risk management strategies are commonly recommended to patients following splenectomy, including wearing a medical bracelet, carrying a laminated medical alert card, possessing a medical letter with specific empiric therapy instructions (including drug names and dosages), and keeping a 5-day supply of standby antibiotics, particularly when travel is anticipated.\textsuperscript{187,189} The need for a high index of suspicion, prompt action, and aggressive education of the patient, family, and medical providers cannot be overstated in asplenic patients.

**REFERENCES**

Entries highlighted in bright blue are key references.


27. Schilling RF. Risks and benefits of splenectomy versus no 
splenectomy for hereditary spherocytosis—a personal view. 


29. Casale M, Perrotta S. Splenectomy for hereditary spherocy-
tosis: complete, partial or not at all? Expert Rev Hematol. 
2011;4(6):627-635.

30. Prchal JT, Gregg XT. Red cell enzymopathies. In: Hoffman R, 

2015;90(9):825-830.

32. Beutler E, Gelbart T. Estimating the prevalence of pyruvate 
kinase deficiency from the gene frequency in the general white 

vate kinase deficiency in northern European population in the 
north of England. Northern Region Haematologists Group. 

34. Hamilton JW, Jones FG, McMullin MF. Glucose-6-
phosphate dehydrogenase Guadalajara—a case of chronic 
non-spherocytic haemolytic anaemia responding to splenec-
tomy and the role of splenectomy in this disorder. *Hematology*. 

35. Michel M. Classification and therapeutic approaches in auto-

C. Efficacy and safety of splenectomy in adult autoimmune 

37. Hill J, Walsh RM, McHam S, Brody F, Kalaycio M. La- 
paroscopic splenectomy for autoimmune hemolytic anemia in 
patients with chronic lymphocytic leukemia: a case series and 

38. Plikat K, Rogler G, Scholmerich J. Coombs-positive autoim-
une hemolytic anemia in Crohn’s disease. *Eur J Gastroen-

39. Rigal D, Meyer F. Autoimmune haemolytic anemia: diagnosis 
strategy and new treatments (in French). *Transfus Clin Biol*. 

40. Rees DC, Williams TN, Gladwin MT. Sickle-cell disease. 

41. Rice HE, Enghum BR, Rothman J, et al. Clinical outcomes of 
splenectomy in children: report of the splenectomy in congenital 

42. Al-Salem AH. Indications and complications of splenec-

43. Rachmilewitz EA, Giardina PJ. How I treat thalassemia. 

44. Phrommintikul A, Sukonthasarn A, Kanjanavanit R, 
Nawarawong W. Splenectomy: a strong risk factor for pul-
monary hypertension in patients with thalassemia. *Heart*. 
2006;92(10):1467-1472.

45. Sheikha AK, Safih ZT, Kasnazzan KH, et al. Prevention of 
overwhelming post-splenectomy infection in thalassemia 
patients by partial rather than total splenectomy. *Can J Surg*. 

46. Dedegolu S, Bornaun H. Pulmonary hypertension in chil-
dren with beta thalassemia major, are splenectomy and high-
ferritin levels related or not? *J Pediatr Hematol Oncol*. 

47. Cines DB, Blanchette VS. Immune thrombocytopenic purpura. 

48. George JN. Sequence of treatments for adults with primary 
S12-S15.

49. Rodighiero F, Besalduch J, Michel M, Provan D, Grotzinger K, 
Thompson G. Treatment practices in adults with chronic 

50. Auger S, Duny Y, Rossi JF, Quittet P. Rituximab before 
splenectomy in adults with primary idiopathic thrombo-
158(3):386-398.

51. Kojouri K, Vesely SK, Terrell DR, George JN. Splenectomy 
for adult patients with idiopathic thrombocytopenic purpura: a 
systematic review to assess long-term platelet count responses, 
prediction of response, and surgical complications. *Blood*. 
2004;104(9):2623-2634.

52. Mikael J, Northridge K, Lindquist K, Kessler C, Deuson R, Danese M. Short-term and long-term failure of 
laparoscopic splenectomy in adult immune thrombocyto-
2009;84(11):743-748.

Chronic idiopathic thrombocytopenic purpura (ITP) is a surgi-

54. Keidar A, Feldman M, Szold A. Analysis of outcome of 
laparoscopic splenectomy for idiopathic thrombocyto-
80(2):95-100.

55. Tada K, Ohta M, Saga K, et al. Long-term outcomes of la-
paroscopic versus open splenectomy for immune thrombo-


outcome following splenectomy for chronic and persistent 
immune thrombocytopenia (ITP) in adults and children: sple-

58. Dubois L, Gray DK. Case series: splenectomy: does it still 
play a role in the management of thrombotic thrombocyto-

59. Targaron EM. Spleen, hematological disorders. In: Matteotti R, 
Becker JM, Ashley SM, eds. *Minimally Invasive Surgical Oncology: 

60. Joly BS, Coppo P, Veyradier A. Thrombotic thrombocytopenic 


62. Shankland KR, Armitage JO, Hancock BW. Non-Hodgkin 

63. Habermann TM, Rai K. Historical treatments of in hairy cell 
leukemia, splenectomy and interferon: past and current uses. 

64. Chow KU, Luxembourg B, Seifried E, Bonig H. Spleen Size 
is Significantly Influenced by Body Height and Sex: Estab-
lishment of Normal Values for Spleen Size at US with a 
306-313.

node biopsy in intra-abdominal lymphoma: high diagnostic 
accuracy achieved with a minimally invasive pro-
175-178.


184. Tovo CV, de Mattos AZ, Coral GP, Branco FS, Suwa E, de Mattos AA. Noninvasive imaging assessment of non-alcoholic fatty liver disease: focus on liver scintigraphy. World J Gastroenterol. 2015;21(15):4432-4439.


ABDOMINAL WALL

General Considerations

Incision and closure of the abdominal wall is among the most common surgical procedures performed. Knowledge of its layered anatomy is critical in the management of surgical patients. Cranially defined by the costal margin and xiphoid process and caudally ending over the pubic bones of the pelvis, the abdominal wall provides support and protection to the peritoneal and retroperitoneal organs.

Surgical Anatomy

Embryologically, the abdominal wall is derived from the mesoderm and envelops the future abdominal contents as bilateral migrating layers originating from the paravertebral area. The leading bars of these sheets fuse in the midline as the linea alba at 7 weeks after gestation and reach the umbilicus at 8 weeks.

The abdominal wall consists of nine distinct layers: skin, subcutaneous tissue, superficial fascia, external oblique muscle, internal oblique muscles, transversus abdominis muscle, transversalis fascia, preperitoneal adipose tissue, and peritoneum. The subcutaneous tissue is composed of superficial adipose tissue that is contiguous with Camper’s fascia anteriorly. Deep to this, Scarpa’s fascia consists of a fibrous matrix of tissue that fuses with the anterior layer of fascia of the flank and back.

The muscles of the abdominal wall consist of the rectus abdominis medially and the external oblique, internal oblique, and transversus abdominis laterally. The rectus abdominis is a paired longitudinal muscle spanning the length of the abdomen. Divided by the linea alba, both rectus muscles originate at the pubic symphysis and crest and insert on the xiphoid process, the fifth and sixth ribs, and the seventh costal cartilage. Three tendinous insertions cross the rectus muscle along its length. The muscle is contained within an aponeurotic sheath formed from the fusion of differing components of the lateral fascial layers.

Laterally, the three muscular layers (external oblique, internal oblique, and transversus abdominis) have obliquely oriented fibers relative to one another. The external oblique arises from the eighth rib and inserts medially into the linea alba and anterior iliac crests. The fibers of this muscular layer travel medial and caudal from its insertion. The internal oblique originates from the thoracolumbar fascia. Its fibers travels cranially and anteriorly, inserting on the lower costal margin and the xiphoid process. The deep muscular layer, the transversus abdominis muscle, begins at the costal margin and lumbar fascia, runs horizontally and anteriorly, and inserts on the linea alba, xiphoid process, and pubis symphysis.

Altogether, these muscular layers provide fascial aponeurotic sheaths anteriorly. These fascial layers provide variable contributions to the separate layers of the rectus sheath and ultimately fuse in the midline as the linea alba. The rectus sheath that encloses the rectus muscles has differing compositions above and below the arcuate line. The anterior rectus sheath is composed of the external oblique aponeurosis throughout its length. The internal oblique aponeurosis is bilaminar and contributes to the anterior and posterior rectus sheaths above the arcuate line. Below this line, the internal aponeurosis contributes only to the anterior rectus sheath. The transversus abdominis aponeurosis contributes the posterior rectus sheath above the arcuate line and the anterior rectus sheath below the arcuate line. Therefore, below the arcuate line, all the aponeurotic layers of the lateral musculature form the anterior sheath, leaving the transversalis fascia as the only posterior fascial covering. This layer is a weak fibrous layer separated from the peritoneum by preperitoneal fat.

Along the posterior abdominal wall are folds corresponding to underlying vasculature and embryologic remnants. The median umbilical fold is formed by the obliterated urachus traveling from the dome of the bladder to the umbilicus in the midline. The bilateral medial folds are formed by remnants of the umbilical arteries. Lastly, the lateral folds are associated with the inferior epigastric vessels.

The deep blood supply of the abdominal wall is supplied mostly from the inferior and superior epigastric arteries. The superior epigastric artery is the final branch of the internal thoracic artery. It enters the rectus sheath below the costal margin.
Key Points

1. There are differences in the anatomic structure in the rectus sheath above and below the arcuate line. Below the arcuate line, all the lateral fascial layers combine and travel anteriorly forming the anterior rectus sheath and leaving the posterior aspect of the lower portion of the rectus muscles without an aponeurotic covering. Above the arcuate line, the posterior rectus sheath is formed by a portion of the internal oblique aponeurosis and the transversus abdomenus sheath, and the anterior rectus sheath is formed by the remaining fibers of the internal oblique and the external oblique aponeuroses.

2. Two randomized trials have found that closure of midline incisions with small fascial stitches of five to eight mm length and five mm width is beneficial in preventing incisional hernias of the abdominal wall.

3. Primary repair of ventral incisional hernias results in high recurrence rates, and repair utilizing other methods, including prosthetic mesh and component separation, are preferred techniques of repair.

4. Laparoscopic incisional hernia repair results in similar recurrence rates and wound infections, compared to open repair. Laparoscopic repair may result in a faster recovery and shorter hospitalization; however, there may be an increase in bowel injury when compared to an open repair.

5. Desmoid tumors should be monitored for asymptomatic patients as there is a possibility of spontaneous regression. For patients with rapidly growing or symptomatic tumors resection is recommended. If complete pathologic resection is not achievable without significant morbidity, more modest resection is recommended along with treatment with adjuvant therapies.

6. The omentum provides an immunogenic and fibrotic response to foreign stimuli, allowing the abdomen to wall off infections preventing diffuse peritonitis.

7. The role of surgery in the treatment of sclerosing mesenteritis is minimal and is most often undertaken to obtain tissue for diagnosis. Most cases should be treated medically with surgical interventions reserved for cases of bowel obstruction and ischemia.

8. Surgical interventions for retroperitoneal fibrosis include obtaining tissue for pathologic diagnosis, relieving ureteral obstructions via ureterolysis or ureteral stenting, and relieving vascular obstructions via endovascular stenting of affected vessels. Most cases are treated successfully with the use of steroids.

and travels along the posterior surface of the anterior sheath and forms an anastomosis with the inferior epigastric artery at the umbilicus. The inferior epigastric artery arises from the external iliac artery. These arteries provide a collateral circulation between the vasculature of the upper and lower extremities (Fig. 35-3). The abdominal wall is also supplied by branches of the subcostal and lumbar arteries. Superficially, the abdominal wall subcutaneous and skin tissue is supplied by branches of the superficial epigastric arteries, femoral arteries, superficial external pudendal, and superficial circumflex arteries. Venous drainage of the abdominal wall is variable but typically follows the aforementioned arteries. Above the umbilicus, the lymphatics of the abdominal wall drain into the superficial axillary nodes. Below the umbilicus, these drain into the inguinal nodes. Lymphatics near the umbilicus can drain along the falciform ligament toward the hepatic nodes.

Innervation of the abdominal wall is segmental, leading to a dermatomal sensory pattern. Afferent branches of the T4 to L1 nerve roots provide sensation of the abdominal wall (see Fig. 35-3). The muscles of the abdominal wall are innervated by the efferent branches of spinal nerves T6 to T12.

Physiology

Aside from providing protection of the intraabdominal and retroperitoneal organs, the abdominal wall muscles assist with flexion, extension, and rotation of the torso along with the muscles of the back and trunk. Working as a unit, the rectus muscles and external/internal obliques activate to flex the trunk anteriorly and laterally. Rotation of the torso is performed by simultaneous stimulation of the contralateral external oblique and ipsilateral internal oblique. To rotate the torso to the right requires simultaneous contraction of the right internal oblique muscle and left external oblique muscle, and vice versa for leftward rotation. Altogether, the muscles of the abdominal wall can act to raise intraabdominal pressure providing assistance with respiration, coughs, defecation, urination, and parturition.

Abdominal Surgical Incisions and Closure

Knowledge of the abdominal wall anatomy is an important aspect of safe entry into the abdomen. The goal of an efficacious incision is to provide adequate exposure to perform the procedure with minimal perturbation of the abdominal wall function.

Incisions for open abdominal surgery are generally located in proximity to operative targets. There are two general types of incisions: longitudinal or transverse/oblique (Fig. 35-4). There does not appear to be differences in early or late postoperative complications or recovery time between these two types of incisions. However, transverse incisions may be associated with lower incisional hernia rates but higher rates of wound infection. Overall, without clear evidence of superiority the choice of incision remains a surgeon-dependent decision. Several retractor systems can be used to provide exposure using these open incisions. Examples include the Bookwalter, Omni-Tract, and Thompson retractors (Fig. 35-5).

The most common longitudinal incision is the midline incision. Providing access to most intraabdominal organs and some retroperitoneal structures, the midline incision is carried down to the linea alba to allow access to the abdomen with minimal injury to skeletal muscles, nerves, and vessels. Paramedian longitudinal incisions are made lateral to the midline through the rectus sheath or in the pararectus location. These incisions restrict access to the contralateral abdomen and pelvis and risk damage to the musculature, vessels, and nerves.

Closure of the midline incision requires reapproximation without undue tension or strangulation of the tissue to prevent...
perioperative dehiscence or incisional hernia formation. Classically, the midline fascia of the incision is closed using sutures placed 1 cm from the edge and a width of 1 cm apart. Two European randomized controlled trials demonstrated reduced rates of incisional hernia with shorter stitch width (5–8 mm) compared to the standard 1 cm stitches. Studies are ongoing in the United States at this time. Several studies have evaluated the use of prophylactic mesh implantation in the closure of the midline incision. While these studies demonstrate reduced incisional hernia formation in the short-term, more long-term data is needed to determine the incidence of mesh-related complications. Furthermore, the optimal location of mesh implantation and type of mesh used are still being evaluated.

When performing a transverse or oblique incision, a surgeon can either divide or separate the muscle fibers of the abdominal wall. The classic McBurney’s incision for appendectomy, an oblique incision performed one-third of the way from the iliac spine to the umbilicus, is an example of a muscle-splitting approach. A subcostal incision can be used to access the upper abdomen, liver, gallbladder, spleen, pancreas, or adrenals. On closure of these incisions, two layers must be approximated. The deep layer includes the internal oblique, transversus abdominus muscle, and the transversalis fascia. The superficial layer includes the anterior aponeurotic tissue of the rectus sheath medially and the external muscle/aponeurosis laterally. A bilateral subcostal incision, or a chevron incision, can be used to access organs of the upper abdomen and diaphragm. A Mercedes-Benz modification includes a midline incision superiorly, providing further access to the upper abdomen or the lower mediastinum.

Pelvic procedures are commonly performed through a Pfannenstiel incision. This incision is performed via a transverse skin incision carried down to the anterior rectus sheath. The sheath is also transversely incised and dissected off the underlying rectus muscle. The rectus muscles are separated, and access through the transversalis fascia is performed longitudinally. Closure of this incision requires approximation of the peritoneum and rectus muscles and closure of the anterior rectus sheath.

Laparoscopic port site incision placement must be carefully planned based on approach angles and working distances both to the operative site and between ports. Placement of a...
nasogastric tube and Foley catheter may help to decompress the stomach and bladder reducing injury to these structures on entry. Initial entry into the abdomen may be completed using the open Hasson or closed Veress needle techniques. The Hasson technique involves direct visualization and systematic opening of each fascial layer encountered on entry. The closed technique utilizes the Veress needle to access the abdomen in a controlled fashion. Access to the peritoneum is confirmed using one or a combination of several techniques (saline drop method or measurement of intraabdominal pressure), and gas insufflation is begun. The abdomen can then be accessed using either a visual entry port or a nonvisualized entry port at the site of the needle track or an alternative site. Choice of entry method is still controversial and based on surgeon preference. Retrospective reviews suggest that complications may be lower using the Hasson technique, but randomized trials (although small) have not found significant differences in techniques.

Abdominal incisions can lead to significant patient morbidity. Complications include hematomas/seromas, surgical site infections, fascial dehiscence, incisional hernias, and nerve injuries among others. In general, minimizing operative incision lengths to only that which is necessary for a safe operation is prudent to reduce these complications and morbidity.
Congenital Abnormalities

The folds of mesodermal cells that form the abdominal wall layers begin to form in the early weeks after gestation. The folds develop in the cephalic, caudal, and right/left lateral directions and converge at the umbilicus around 8 weeks. At around 6 weeks of development the contents of the abdominal cavity outgrow the space allowed by the surrounding layers of the...
abdominal wall, resulting in temporary herniation of the abdominal contents through a central defect. At this time, the vitelline duct (omphalomesenteric duct) and allantois also pass through the central defect. The vitelline duct serves as a conduit to the embryologic midgut and the yolk sac. As the midgut develops outside of the abdomen, it undergoes a 270° counterclockwise rotation and reenters the abdomen at around 12 weeks. Failure of the midgut to reenter the abdomen leads to the congenital abdominal wall defect known as an omphalocele, in which the contents of the defect protrude through an open umbilicus and are covered by an amniotic/peritoneal membrane. Gastrochisis, on the other hand, results from either malformation or disruption of the abdominal wall, either from genetic defects or vascular compromise. Gastrochisis presents as protruding viscera through a defect lateral to the umbilicus (usually along the right side) without the amniotic sac covering.

The vitelline duct usually involutes at around the eighth to ninth week after gestation. Failure of vitelline duct regression can lead to several abnormalities depending on the spectrum of involution. This spectrum includes total persistence of the vitelline duct leading to omphalomesenteric fistula causing drainage of intestinal contents at the umbilicus, to partial closure leading to omphalomesenteric cyst. If the vitelline remnant persists at the ileal border, a Meckel’s diverticulum forms. The vitelline duct (omphalomesenteric duct) and allantois also pass through the central defect. At this time, the vitelline duct may also persist as a fibrous attachment of the intestine to the abdominal wall predisposing the patient to bowel obstructions. Fistulas, cysts, and fibrous attachments should be resected when diagnosed.

The urachus is the proximal portion of the allantois forming as the bladder descends into its pelvic position. The urachus closes and forms the median umbilical ligament of the abdominal wall as previously described. Failure of the urachus to close results in urinary fistula or cyst. These are treated by urachal excision and closure of the bladder defect.

Acquired Abnormalities

Abdominal Wall Hernias. A protrusion or bulge of abdominal contents through the abdominal wall muscle/fascia represents an abdominal wall hernia. This may be present at birth or acquired from weakening or disruption of the overlying fascia, or from failed healing of a surgical incision. Hernias may present as asymptomatic bulges that increase with Valsalva maneuvers, or with significant discomfort. On physical exam, the patient’s abdominal wall should be evaluated with the patient both standing and in the recumbent position. Hernias may reduce spontaneously or without manual pressure. If a hernia is incarcerated, it cannot be reduced and generally requires surgical correction. If intestine is incarcerated in the hernia defect, bowel obstruction may ensue, which represents a surgical emergency. Incarcerated hernias present with significant pain, nausea, and vomiting. A hernia is considered strangulated if blood supply to its contents is compromised. Localized ischemia may lead to infarction and eventual perforation if left untreated.

Nonincisional hernias are named based on their location on the abdominal wall. Epigastric hernias are defects in the abdominal wall located between the umbilicus and the xiphoid process. These hernias are usually small but may be associated with multiple defects. They result from multiple factors, including muscle weakness, congenitally weakened epigastric fascia, or increases in intra-abdominal pressure. Epigastric hernias rarely contain bowel and usually contain portions of the omentum or falciform ligament. Given the rarity of incarceration, repair of an epigastric hernia is indicated for symptomatic patients only. Laparoscopic repair can be attempted, but this type of hernia usually can be managed with a small incision where the defect is closed with interrupted sutures.

Umbilical hernias may be congenital or acquired. Umbilical hernias are common in newborns, especially in premature infants. Closure of an umbilical defect occurs after birth as the muscles of the rectus abdominis grow toward one another. Most umbilical hernias close spontaneously by 5 years of age and can be monitored as they will spontaneously resolve. Indications for repair include incarceration, symptomatic hernia, failure to decrease in size or if the defect fails to close by the age of 5 years.

In adults, umbilical hernias form because of increased abdominal pressure due to pregnancy, obesity, or ascites. Females are at higher risk for this type of hernia than men. Small, asymptomatic hernias may be followed clinically. However, if an umbilical hernia enlarges in size, causes symptoms, or incarcerates surgical treatment should be offered. Hernias can be repaired laparoscopically or with an open procedure. Mesh should be employed for large defects where the fascial edges cannot be approximated without tension. In this case, mesh should be placed as a sublay technique (below the fascia) and sutured in place to prevent migration.

Patients with cirrhosis and associated ascites with an umbilical hernia pose a significant clinical dilemma. Umbilical defects enlarge in these patients because of high intra-abdominal pressure associated with uncontrolled ascites. With severe liver disease, these patients are at high risk of operative complications. Most hernias contain ascites; however, omentum and bowel may also enter the defect. Given the high pressure, skin breakdown may ensue leading to hernia rupture or weeping as well as risk of spontaneous bacterial peritonitis. All attempts should be made to control the patient’s ascites prior to repair. Therefore, asymptomatic patients should be managed conservatively with aggressive management of ascites. Liver transplant candidates should undergo repair at the time of transplantation as pretransplant repair has high morbidity and mortality. Patients with incarcerated hernias or with thinning or ruptured skin overlying the hernia should be treated emergently.

Hernias that occur along the arcuate line are known as Spigelian hernias. While rare, these hernias form due to the anatomic weakness of lack of a posterior rectus sheath below the arcuate line. As the hernia develops, peritoneum that passes through the arcuate line will pass laterally toward the external oblique muscle given the overlying aponeurosis (Fig. 35-6). Most patients present with pain and swelling in the mid to lower abdomen. Incarceration is common as up to 20% of patients present with a nonreducible hernia. Given the high rate of incarceration, surgical repair is usually recommended. Either open or laparoscopic repair can be performed. The defect is closed by approximating the medial and lateral edges of the transversalis fascia to the rectus sheath.

Incisional Hernias. Hernias that develop at sites of previous abdominal incisions are known as incisional hernias. Hernias can develop at the site of any previous abdominal incision. Up to 20% of midline incisions will develop hernias eventually. Vertical incisions may have a higher risk of hernia formation than transverse or oblique incisions. Upper
abdominal incisions are also at higher risk than lower incisions. Laparoscopic port sites may also develop hernias. The etiology of incisional hernias is complex. Several patient-derived factors increase the risk of hernia, including diabetes, immunosuppressant use, obesity, smoking, malnutrition, and connective tissue disorders. Local operative factors may also be implicated, including technique, wound infection, or high tension at the time of closure. Hernias can develop up to 10 years after surgery but normally occur in the early postoperative period. Incisional hernias can present as asymptomatic bulges or with severe discomfort. Multiple hernias can be present along the length of the incision. Elective surgery should be recommended in patients who are symptomatic. Small defects pose a higher risk of incarceration and should be repaired. To improve operative outcome, patient associated factors, including smoking and obesity, should be remedied prior to surgical repair.

Surgical management of incisional hernias include either primary tissue or mesh repairs. Hernias can also be repaired via a laparoscopic or open approach. Simple suture repair is associated with recurrence rates as high as 54%. A Cochrane review of several randomized controlled trials found that open mesh repair improved hernia recurrence rates when compared to simple closure (33% with simple repair vs. 16% with mesh repair). Mesh repairs are, however, associated with a higher rate of infections.

To reduce tension at the suture line associated with primary suture repair Ramirez described the components separation technique in 1990. This procedure entails dividing portions of the bilateral external oblique aponeuroses forming musculo-fascial advancement flaps. The posterior rectus sheath can also be incised, allowing up to 10 cm of medial mobilization and tensionless approximation of the midline fascia. This technique can cause large skin flaps and initially had high rates of infection. Overtime, techniques have been developed to decrease flap formation and decreased rates of surgical site infection. Endoscopic component separation can also be used to mobilize flaps with minimal ischemia to overlying subcutaneous and skin tissue, theoretically decreasing infection rates. Mobilization of midline fascia is reduced with endoscopic methods. Mesh can also be utilized to reinforce the repair. Overall, component separation without mesh compares to standard mesh repair with respect to hernia recurrence, but precludes the risk of mesh implantation. When mesh is added to component separation, recurrence rates may be as low as 4% to 10% depending on follow up period.

Mesh repair has become the standard for elective management of most incisional hernias. Position of mesh placement is controversial. Mesh can be placed above the midline fascia (overlay), bridged across fascial defects (interlay), underneath fascia (sublay), or within the abdominal cavity (underlay). A systematic review found that sublay placement of mesh may reduce hernia recurrence and prevent wound related complications. The sublay technique is performed by developing the plane between the rectus muscle and the posterior sheath, and affixing mesh in this space. The anterior sheath can be approximated if there is no tension.

Material used for mesh manufacturing can be classified into two classes: synthetic and biologic. Synthetic meshes can be either permanent or degradable, while all biologic meshes are degradable. Permanent mesh is currently made of either polypropylene, polyethylene terephthalate polyester, or expanded polytetrafluoroethylene. Permanent mesh is durable and of relatively low cost. Degradeable synthetic mesh, including Vicryl mesh, is eventually eliminated and loses structural support, but it does offer the advantage of lower mesh infection rates. Degradeble mesh is associated with high recurrence rates, but it can be used for temporary abdominal wall closure in contaminated or infected fields. Newer synthetic biomaterial meshes, including Gore BioA or Phasix, degrade over a longer period of time and may reduce recurrence rates, but long-term effectiveness is unknown. Biologic meshes are decellularized, collagen-rich porcine, bovine, or human tissue. These meshes are designed to allow host cellular ingrowth, promoting incorporation and eventual replacement of the mesh with host tissue. Biologic meshes are a high cost alternative to synthetic degradable mesh and can be used in infected fields. However, their efficacy in preventing recurrence is unclear. Composite products contain two components and are used during intraperitoneal repair. One side of composite mesh, which is placed on the abdominal wall side of implantation, is made of typical nondegradable synthetic material and promotes integration of host tissue. The other side is covered in a synthetic or biologic material, allowing contact with viscera and preventing adhesion formation. These materials include polyglactin, collagen, cellulose, titanium, omega-3, and hyaluronate. This allays concerns of direct mesh contact
with viscosa that may cause adhesions, erosion, and eventual fistula formation. Pore size and mesh weight are also important aspects of mesh design. Recently, large-pore, lightweight mesh has been developed. This was initially thought to delay incorporation, but that has not been seen in practice. They do allow the theoretical advantage of increased incorporation of host tissue and potentially better elasticity and improved postoperative pain when compared to microporous heavier meshes. There is also initial data which suggests lower rates of mesh infection with the use of large pore mesh.

Initially described by LeBlanc and Booth in 1993, laparoscopic repair is now an accepted modality for treatment of incisional hernias. Several studies have found that laparoscopic repair has a lower incidence of surgical site and mesh infections compared to open repair. It also seems that laparoscopic repair allows faster recovery with less postoperative pain. A meta-analysis of 11 studies comparing laparoscopic and open ventral hernia repair found no difference in hernia recurrence and lower rates of wound infection and wound drainage. There was however, a higher risk of bowel injury in the laparoscopic group. Another meta-analysis of six randomized controlled trials had similar findings. Follow-up in these studies is relatively short, and more long-term data is needed at this time to compare these two modalities of repair.

Laparoscopic hernia repair is performed by initially placing lateral ports for midline defects and contralaterally placed ports for lateral defects. Adhesions between the abdominal wall and intestine are carefully taken down, and the hernia contents are completely reduced. The sac is normally left in situ. Once the fascial defect is defined, a mesh is properly shaped and fashioned over the hernia. Transfascial sutures are placed circumferentially to position the mesh with sufficient overlap (4–5 cm) with healthy abdominal wall. Spiral tacks may be placed according to surgeon preference. Even more recently, robotic surgery has been established as another surgical modality in the treatment of ventral hernias. The theoretical advantage of improved visualization and articulating instruments may improve outcomes, but the cost-effectiveness of robotic repair is unclear. Overall, more studies are needed to evaluate the role of robotics in ventral hernia surgery.

Although still rare, given the increase in laparoscopic procedures, the incidence of laparoscopic port site hernias is becoming more common. Given the size of the hernias, there is a substantial risk of bowel strangulation and ischemia. These hernias commonly present as Richter’s hernia, or a hernia containing only a portion of bowel wall. A recent meta-analysis found that the incidence of port site hernia after laparoscopic procedure was less than 1%. Patients can present either early or several years after surgery. Risk factors are similar to other incisional hernias. The most common site of herniation is at an umbilical incision, but it may be found elsewhere. In adults, hernias usually occur in ports that are greater than 5 mm in size, but they can occur in any size ports in children. Depending on the timing of presentation, these are usually repaired via an open approach by increasing the size of the skin incision, reducing the hernia, and approximating all layers of fascia.

Rectus Abdominis Diastasis. Rectus abdominis diastasis (diastasis recti) is an anatomic term referring to an abnormal separation of rectus muscles and a laxity at the linea alba. Although there is controversy regarding normal separation distances between rectus muscles, a distance of two centimeters is usually considered abnormal in the midline abdomen above the umbilicus. This can either be a congenital or acquired abnormality. This is not a true hernia as the midline fascia is intact, and as such incarceration and strangulation do not occur. Risk factors for acquired rectus abdominis diastasis include conditions that elevate intraabdominal pressure, including obesity and pregnancy, as well as conditions which weaken the abdominal wall, including connective tissue disorders or prior abdominal surgery. Risk factors of developing a diastasis recti after pregnancy include older age at the time of pregnancy, multiple pregnancies, and recurrent cesarean sections. Postpartum exercise reduces the risk of developing diaasthesis recti. Most patients with diastasis recti can be diagnosed based on physical exam where a fusiform bulge is usually apparent. This bulge worsens with contraction of the rectus muscles or Valsalva maneuver. If imaging is needed, ultrasonography can be used to confirm diagnosis and rule out hernia. CT scan can also be used to confirm diagnosis and measure distance between muscle pillars. Rectus diastasis does not require surgical repair and may be improved via weight loss and exercise. Indications for repair include disability of abdominal wall muscular function or cosmesis. Surgical repair includes both open and laparoscopic plication of the rectus sheath. Mesh can also be used to bridge the muscle; however, complication rates increase with mesh usage. These procedures do, unfortunately, have a high risk of recurrence long term and introduce a new risk of incisional hernia.

Rectus Sheath Hematoma. Disruption of one of the branches of the inferior epigastric artery as well as an inability to tamponade the hemorrhage results in a rectus sheath hematoma. This occurs commonly around the arcuate line where the artery and its branches are relatively fixed causing vulnerability to shearing forces. Several risk factors are associated with rectus hematoma formation via either proclivity to disruption of blood vessels or by inability to cease bleeding. Trauma to the abdominal wall, including iatrogenic trauma with laparoscopic trocar placement, can lead to disruption of blood vessels. Vigorous contraction of the rectus muscle, either with coughing, sneezing, or exercise, can also induce hemorrhage formation. Chronic pulmonary disease can lead to hemorrhage because of coughing fits. Patients on anticoagulation also present with higher risk of hematoma formation. This condition presents with acute abdominal pain and a palpable abdominal wall mass. Rectus sheath hematoma may be mistaken with intraperitoneal pathology, including appendicitis if on the right side. However, in patients with rectus sheath hematomas, pain usually increases with contraction of the rectus muscles as opposed to intraperitoneal conditions. In addition, palpation of a mass that does not change during contraction of the rectus muscle, known as Fothergill’s sign, is also associated with rectus sheath pathology. The diagnosis should be confirmed via ultrasound or CT scan with intravenous contrast. Obtaining a type and screen, hemoglobin/hematocrit, and coagulation factors are critical in the management of these patients.

Treatment of patients with rectus sheath hematoma depends on the hemodynamic stability of the patient as well as the size of the hematoma. Hemodynamically stable patients with small hematomas, stable serial hemoglobin/hematocrits, and normal coagulation factors may be observed without hospitalization. Hemodynamically stable patients with larger or bilateral hematomas and decreases in hemoglobin should be monitored in the hospital setting, with serial hemoglobin levels, compression of the hematoma, and bedrest. If anticoagulated, reversal is necessary and transfusions of packed red blood cells may
be required in some situations. Patients in hypovolemic shock should be aggressively resuscitated with the use of blood products and treated via angiographic embolization. Angiographic intervention may also be required if the hematoma increases in size or if clinical deterioration occurs. Surgical therapy can be performed if angiographic intervention has failed. Surgical treatment includes operative evacuation of hematoma and ligation of bleeding vessels.

**Desmoid Tumors.** Also known as aggressive fibromatosis, desmoid tumors are fibroblastic neoplasms with aggressive infiltrative behavior but no metastatic potential. These tumors can occur anywhere in the body but commonly occur in the abdomen or abdominal wall. Desmoid tumors are rare and usually occur sporadically. They are, however, also associated with familial adenomatous polyposis (FAP), with an even greater risk in patients with Gardner’s syndrome. Of patients with FAP, 10% to 15% develop desmoid tumors. After prophylactic colectomy, desmoid tumors becomes the primary cause of death in patients with FAP. Risk factors for sporadic development of desmoid tumors include previous surgical incision, pregnancy, hormonal exposure, and trauma. Females have a higher predilection for formation of desmoid tumors. Diagnosis can be performed via core-needle or incisional biopsy. Larger tumor size, young patient age, and extra-abdominal tumor location all predict poor recurrence free survival.

The gold standard of treatment of abdominal wall desmoid tumors, historically, is margin-negative resection with immediate mesh reconstruction. However, more recently there has been controversy as to whether complete microscopic resection is necessary. It is now commonly agreed that while complete microscopic resection is ideal, with the emergence of adjuvant therapy a positive margin may not require additional surgery, especially if re-resection would cause high morbidity. There is also some evidence to suggest a period of close watchful waiting, as some tumors appear to remain stable or even regress over time. In one cohort of 106 abdominal wall desmoids managed initially without surgery, 16% of patients went on to require surgery in a follow-up period of three years. Interestingly, 29 patients had spontaneous regression of their tumors over that time. The National Comprehensive Cancer Network (NCCN) now suggests initial close observation for patients with asymptomatic, non–life-threatening tumors. Surgery is indicated in patients with symptomatic disease, risk of invasion of surrounding structures, or enlarging tumors. There may be a role for adjuvant or neoadjuvant radiation therapy, although the data are unclear at this time. Primary radiation therapy may be an option for patients who are not surgical candidates. There may also be a role for systemic therapy, especially if tumors are unresectable. Options include hormonal therapy, nonsteroidal anti-inflammatory agents, cytotoxic chemotherapies (doxorubicin or carboplatin), or imatinib.

**Other Abdominal Wall Tumors.** Various tumors may also be found within the abdominal wall including lipomas and neurofibromas (Fig. 35-7). Surgical resection is recommended for symptomatic or enlarging lesions. Abdominal wall malignancies are exceedingly rare and include several histologic subtypes of sarcomas, dermatofibrosarcoma protuberans, schwannomas, and melanomas. Workup of abdominal wall tumors should include core needle biopsy or excisional biopsy if the tumor is small enough. Magnetic resonance imaging (MRI) or CT scan with IV contrast should be utilized to define local extent of disease.

MRI is preferred as this modality provides more detail on extent of disease. Chest CT should also be obtained to rule out pulmonary metastasis in high-grade tumors. Chest X-ray may be sufficient to stage tumors that are low grade, given the lower risk of distant disease. Surgical resection is the mainstay of treatment for nonmetastatic disease. For most soft tissue sarcomas, 1 cm margins are usually sufficient, but 2 cm margins may be needed for dermatofibrosarcoma protuberans. Tumors of the superficial abdominal wall should be resected with the underlying fascia, which may require use of mesh to prevent abdominal wall laxity or hernia.

Given the rarity of these tumors, the efficacy of adjuvant or neoadjuvant chemoradiation is unclear. Adjuvant or neoadjuvant therapy may be recommended in patients with large (>5 cm) or high-grade tumors. Some centers administer neoadjuvant radiation therapy to patients with high-risk tumors to monitor response. Chemotherapy may also be used in certain situations. If tumors involve underlying viscera, en bloc resection may be required. Primary closure may be feasible, but prosthetic mesh use (even in the setting of bowel resection) may be required. Options for abdominal wall closure after resection include absorbable or biologic mesh reinforcement, and myocutaneous flap reconstruction.

**OMENTUM**

**Surgical Anatomy**

The omentum is a fibrous adipose apron providing support and protection of the intraabdominal viscera. Embryologically, the omentum originated from the dorsal mesogastrium. Anatomically, the omentum is divided into the greater and lesser omentums. The greater omentum begins to form during the fourth week of gestation. Initially, the omentum forms as a double-layered structure, with the spleen developing between the two layers. As development proceeds, the layers of mesentery fuse, the spleen assumes its position in the peritoneum, and the gastroplenic ligament forms. The greater omentum, therefore, is a double-layered sheet of visceral fibroadipose tissue descending from the greater curvature of the stomach, covering the small intestines. The omentum folds back on itself and attaches onto the anterior peritoneum of the transverse colon (Fig. 35-8).
In an adult, the greater omentum lies between the abdominal wall and the hollow viscera, usually extending into the pelvis. The greater omental tissue connecting the stomach to transverse colon, as well as the stomach to the spleen, are known as the gastrocolic and gastrosplenic ligaments, respectively. The blood supply to the greater omentum is derived from the right and left gastroepiploic arteries. The venous system parallels the arterial supply and ultimately drains into the portal system. Lymphatic drainage of the greater omentum occurs via the subpyloric or splenic nodes depending on laterality ultimately culminating in the celiac nodes and subsequently the thoracic duct.

Forming the anterior boundary of the lesser sac, the lesser omentum extends between the liver and lesser curvature of the stomach. Also known as the hepatoduodenal and hepato gastric ligaments, the portal triad (including the common bile duct, portal vein, and hepatic artery) is located within the inferolateral margin of the lesser omentum. This free edge of lesser omentum forms the foramen of Winslow, which is used to encircle the portal triad during a pringle maneuver (see Fig. 35-8).

**Figure 35-8.** Greater and lesser omentum. The greater omentum begins along the greater curvature of the stomach, drapes over the transverse colon into the pelvis, and folds back on itself inserting along the posterior wall of the transverse colon. The greater omentum includes the gastrophrenic and gastro splenic ligament. The lesser omentum includes the gastrohepatic and hepatoduodenal ligaments, covering the lesser sac of the abdomen. The free lateral edge of the lesser omentum includes the portal triad, forming the foramen of Winslow below. (Reproduced with permission from Moore KL, Agur AM: Essential Clinical Anatomy, 5th edition. Philadelphia, PA: Lippincott Williams & Wilkins; 2014.)

**Physiology**

Initially described by the British surgeon Rutherford Morison as the “policeman of the abdomen,” the greater omentum is understood as a peritoneal defense organ. The omentum forms fibrin adhesions at sites of inflammation, effectively attempting to wall off peritoneal infections. The greater omentum is known to respond to a foreign stimulus by expanding stromal tissue that expresses chemotactic, inflammatory, and hemostatic factors, which promote tissue inflammation and subsequent repair. Forming a fibrin bridge between omental tissue and the injured site allows passage of inflammatory factors and immune cells. The Graham Patch repair of a perforated ulcer utilizes the complex inflammatory interactions of the omentum to improve leak rates after repair.

The visceral fat of the greater omentum is also involved in metabolic functions of the body. Increased visceral fat, including the greater omentum, is an independent risk factor for insulin resistance and high triglyceride levels. This may be, in part, mediated by the increase in visceral inflammation seen in intra-abdominal obesity. Several adipokines, or cell-signaling molecules secreted by adipose tissue, are secreted by the greater omentum and act on peripheral tissues, affecting metabolic health throughout the body.

**Acquired Disorders**

**Omental Infarction.** Omental infarction is a rare cause of acute abdominal pain. Omental infarction may be primary or
secondary depending on the etiology. Primary infarction may be caused by torsion of the omentum from sudden or forceful movements, thrombosis or vasculitis of omental vessels, or omental venous outflow obstruction. Secondary causes of torsion are most often due to hernias, tumors, or adhesions. Only about 250 cases of primary omental infarction have been reported in the literature. This occurs most often in male and obese patients. Omental infarction may mimic other intra-abdominal pathologies such as appendicitis, cholecystitis, or diverticulitis. Abdominal exam usually demonstrates peritoneal tenderness, with a possible palpable mass. Ultrasonography (US) and abdominal CT scan are used to assist diagnosis. US may demonstrate a hyperechoic, noncompressible intra-abdominal mass attached to the abdominal wall. CT scan demonstrates a streaking whirling pattern of fatty tissue in the anterior abdomen (Fig. 35-9). Regardless, only a small percentage of patients are preoperatively diagnosed. Omental infarction can be treated conservatively, but this leads to several possible complications, including abscess or adhesion formation. Laparoscopic exploration and resection of infarcted tissue is the usual treatment of choice and leads to rapid resolution of symptoms.

Omental Cysts. Cystic lesions of the omentum are rare disorders, related in pathogenesis to mesenteric cysts. Most are thought to form through degeneration or inclusion of lymphatic structures. Case reports described these lesions presenting as vague abdominal pain or nausea or vomiting. Cysts may also present as a lead point for omental torsion and infarction. Physical exam may or may not reveal evidence of an intra-abdominal mass. Cysts may also be diagnosed incidentally on imaging exams performed for other reasons. CT scan and US are largely diagnostic and reveal a well-circumscribed cystic appearing lesion arising from the greater omentum. A very small percentage may transform to malignancy. Most are resected, especially if symptomatic, via open or laparoscopic approaches. Cysts may relapse if treated conservatively via laparoscopic unroofing or percutaneous drainage.

Omental Neoplasms. Most omental neoplasms are metastatic disease. Ovarian cancer is the most common cancer to have omental involvement. Other metastatic cancers include gastrointestinal tract tumors, melanoma, endometrial cancer, and kidney cancer.

Primary tumors of the omentum are exceedingly rare. Benign masses may include lipomas, myxomas, and desmoid tumors. Extra gastrointestinal stromal tumors are a rare malignant tumor of the omentum that have been described in several case series. These tumors share many genetic and immunophenotypic similarities to classical GISTs, including c-kit and PDGFRA mutations. In one review of case reports, the median age of diagnosis was noted to be 65 years and had equal predilection for male and female patients.

**MESENTERY**

**Surgical Anatomy**

The mesentery is a contiguous structure suspending and fixing bowel to the abdominal wall providing housing for arterial, venous, nervous, and lymphatic structures connecting hollow viscera with the body. The mesentery is derived from the mesodermal germ layer becoming the dorsal mesentery. Previous theories of development included sliding and regression models of mesenteric development. However, with a new contiguous model of development of the mesentery, several key simple steps are proposed to take place. Initially, the mesentery provides a point of suspension from vascular connections. As the intestine and mesentery elongate, they undergo a 270° counterclockwise rotation leaving the primordial abdominal cavity. Eventually, the duodenum and ascending/descending colon flatten against the posterior abdominal wall returning the bowel back into the abdomen and taking its normal shape. The resultant development of the white line of Toldt along the lateral border of the ascending and descending large bowel provides an avascular fascial plan between the colon, its mesentery, and the underlying retroperitoneal space. The small intestine mesentery, transverse colon mesentery, and sigmoid colon mesentery remain mobile.

Defects in the proper rotation and fixation of the bowel results in the spectrum of congenital disorders known as intestinal malrotation. In this scenario, the intestine and mesentery are simply suspended from vascular connection, making intestinal volvulus common. Defects that form in anatomical positions may act as routes for intestinal herniation. Common sites of herniation include the paraduodenal or mesocolic areas, leading to either acute or chronic intestinal obstruction in pediatric or adult populations. Areas where mesenteric attachment is incomplete may predispose patients to volvulus. Attachments of the cecum may degrade overtime leading to an area of laxity and possible twisting. The sigmoid mesentery can also increase in size over time also leading to possible sigmoid volvulus. Rarely, other areas of the colon can develop volvulus if differential mesenteric lengths form.

**Sclerosing Mesenteritis**

Sclerosing mesenteritis is a rare disorder characterized by idiopathic fibrosis of the mesentery, affecting hollow viscera as well as mesenteric vessels. The disease is part of a spectrum of inflammation and fibrosis, which when localized is known as mesenteric lipodystrophy and when diffuse is known as mesenteric panniculitis. The cause of this disease is unknown, but it may be instigated by antecedent abdominal surgery, an autoimmune disease, a paraneoplastic syndrome, a previous infection (including typhoid, tuberculosis, influenza, and rheumatic fever), or vascular insult. This disease most commonly effects
white patients between 50 and 70 years of age, although it has been rarely described in children. Most studies report a predilection for male patients. Most patients with this disease present with abdominal pain. Other symptoms include nausea and vomiting, weight loss, anorexia, and altered bowel habits. This may appear as a chronic or acute disorder. On physical exam, patients may be found to have tenderness and distension. Up to 50% of patients are found to have an abdominal mass that often transmits aortic pulsations.

Abdominal CT with IV contrast is used to assist in diagnosis (Fig. 35-10). The most common finding is that of a soft tissue mass with a higher density than normal mesenteric tissue. Although it is sometimes difficult to distinguish mesenteric fibrosis from a mesenteric tumor, two CT findings may add specificity. The “tumor pseudocapsule” refers to a hypodense zone around the associated fibrotic mass, and the “fat ring sign” refers to an area of preserved fat near mesenteric vessels coursing through areas of fibrosis. These lesions may also be calcified on CT scan.

Pathologic confirmation is required to confirm the diagnosis. This usually requires laparoscopic or open biopsy to provide adequate tissue for confirmation. Treatment of sclerosing mesenteritis is complex, with surgery having a minimal role. Patients who present with bowel ischemia may require bowel and mesenteric resection; however, the extent and location of mesenteric involvement may preclude complete resection. If obstructive symptoms are dominant, intestinal bypass may be indicated. Aggressive surgical treatment is not indicated because in many cases symptoms may improve with medical treatment or even without intervention. Steroids, hormonal therapy, colchicine, thalidomide, and cyclophosphamide have all been reported to be beneficial.

Mesenteric Cysts
Mesenteric cysts are a rare benign disorder with an incidence ranging from 1 in 27,000 to 1 in 250,000 admissions. Cysts are thought to be caused by disruption of the lymphatics in the mesentery either by traumatic disruption, mechanical obstruction, or congenital lymphatic malformations. Most cysts are unilocular, but they also may have multiple loculations. They are usually lined with a single layer of columnar epithelial cells. Presentation of mesenteric cyst is varied, with some being found incidentally on imaging exams and others causing acute abdominal pain because of cyst rupture or bowel torsion. Chronic symptoms are usually nonspecific, including abdominal pain or discomfort, anorexia, distension, nausea, vomiting, or changes in bowel habits. Symptoms are due to local compression of abdominal structures. Up to 45% of cysts are found incidentally.

Physical exam reveals an abdominal mass in up to 60% of patients. The classic Tillaux’s sign is an abdominal mass lesion that is only mobile laterally, contrasting omental cysts which are usually freely mobile in all directions. CT scan and ultrasound can be used to make an accurate diagnosis. Cystic lesions usually appear as a fluid filled mass without solid components (Fig. 35-11). It can sometimes be difficult to differentiate cystic masses from solid tumors based on imaging. Mesenteric cystic lymphangioma may present as numerous cysts on imaging. Up to 3% of mesenteric cysts contain malignancy, mostly as a sarcomatous lesion. In one recently published series, 19% of patients harbored malignancy. Solid components within the cystic structure are associated with higher rates of malignancy.

Most mesenteric cysts are treated surgically. Marsupialization and simple aspiration have high rates of recurrence and...
are generally discouraged. Benign lesions should be enucleated, and malignant cysts should be resected with clear margins. Intestines or surrounding viscera may require resection if the associated vasculature is excised. Resection can be performed either via a laparoscopic or open procedure.

**Mesenteric Tumors**

Primary mesenteric tumors are rare but represent several histologic patterns. Benign tumors include desmoid tumors, lipomas, and cystic lymphangiomas. The most common malignant neoplasm of the mesentery is lymphoma. On imaging, lymphoma may appear as bulky adenopathy, usually surrounding and not obstructing nearby structures. Lymphomas should not be resected, but they may require operative biopsy for diagnosis. Other malignant tumors of the mesentery include gastrointestinal stromal tumors, carcinoids, liposarcoma, leiomyosarcoma, malignant fibrous histiocytomas, lipoblastomas, or lymphangiosarcoma. Treatment of malignant mesenteric masses usually involves wide resection; however, given the proximity to mesenteric vessels, resection may not be feasible or require removal of large portions of bowel.

**RETROPERITONEUM**

**Surgical Anatomy**

The retroperitoneum is bound by the peritoneum anteriorly, the iliopsoas and lumbar muscles posteriorly, the diaphragm superiorly, and the levator ani muscles inferiorly. It is divided into the three spaces: the anterior pararenal space, the perirenal space, and the posterior pararenal space (Fig. 35-12). The anterior pararenal space refers to the area anterior to the renal fascia but posterior to the peritoneum. This area contains the ascending and descending colon, the duodenum, and the pancreas. Posterior to this space is the perirenal space, which houses the inferior vena cava, the aorta, kidneys, and adrenal glands. The posterior pararenal space is in continuity with preperitoneal fat of the anterior abdomen. Given the compliance of the anterior boundary of the retroperitoneum and the rigidity of other margins, tumors, hematomas, and abscesses tend to expand anteriorly toward the peritoneal cavity.

With the expansion of minimally invasive techniques in surgery, the retroperitoneoscopic approach has emerged as a potential modality for access to retroperitoneal organs. Patients are positioned in the prone or lateral decubitus positions. The retroperitoneoscopic approach allows access to the kidneys, adrenal glands, and retroperitoneal lymph nodes. One systematic review found no difference when comparing laparoscopic to retroperitoneoscopic adrenalectomy in terms of operative outcomes, complications, or postoperative recovery. Retroperitoneoscopic approach did, however, lead to shorter hospital stay likely because of reduced postoperative pain. The majority of the studies evaluating the retroperitoneoscopic approach are retrospective, and more randomized trials are needed to provide further guidance.

**Retroperitoneal Infections**

Infections of the retroperitoneum can be due to primary hematogenous spread of microbes or due to secondary infection from retroperitoneal or nearby organs. Examples include abscesses due to a perforated retrocecal appendix, diverticulitis, a contained perforated duodenal ulcer, iatrogenic perforation of the gastrointestinal tract, or pancreatitis. Patients may develop back,
flank, or groin pain and suffer from fevers or chills. Depending on severity, patients may present with fulminant sepsis. Abscesses may become quite large given the substantial retroperitoneal space. On physical exam, patients may present with erythema of the umbilicus or flank. Abscesses are usually found easily on CT scan of the abdomen with IV contrast, which can also show if the collection is loculated. Treatment of retroperitoneal abscesses includes source control via treatment of the underlying condition, drainage of well-defined collections, and IV antibiotics. Image guided drainage is preferred, but it may be unsuccessful if the abscess is multifocal or in an inaccessible area. For these cases, operative drainage may be required. Given the insidious nature of this disease and a lack of abdominal findings on physical exam, recognition of a retroperitoneal abscess may be delayed. Delays in diagnosis and insufficient drainage may lead to high morbidity and mortality. Depending on severity, mortality of retroperitoneal abscess can be as high as 25%. Rarely, patients may develop necrotizing fasciitis of the retroperitoneum, a condition with high mortality.

Retroperitoneal Fibrosis

Retroperitoneal fibrosis is a rare disease characterized by inflammation and fibrosis of the tissue of the retroperitoneum. It exists as a spectrum of disease with chronic periaortitis, which affects the retroperitoneal tissue near large arteries of the retroperitoneum. Fibrosis gradually expands, encasing the ureters, inferior vena cava, aorta, mesenteric vessels, or sympathetic nerves. Bilateral involvement is noted in up to 70% of cases. The condition may either be idiopathic or due to a secondary cause, including aortic aneurysms, pancreatitis, certain drugs (Ergot-derivatives, β-blockers, hydralazine, methyldopa, among others), malignancies (including lymphoma, carcinoids, sarcomas, colorectal, breast, and others), infections such as tuberculosis, radiation, retroperitoneal hematoma, surgery, asbestos, or tobacco use. Retroperitoneal fibrosis has been described in association with several autoimmune disorders including ankylosing spondylitis, systemic lupus erythematosus, Wegener’s granulomatosis, and polyarteritis nodosa.

Idiopathic retroperitoneal fibrosis accounts for 70% of cases and is thought to be an immune-mediated disorder. Theories regarding the pathogenesis of idiopathic fibrosis of the retroperitoneum include exaggerated local reactions to aortic or iliac atherosclerosis or autoimmune deposition of fibroinflammatory cytokines. Retroperitoneal fibrosis is also a manifestation of IgG4-related disease, a multisystem disease characterized by lymphocytic infiltrate and variable degrees of fibrosis of several affected organs. One study has associated retroperitoneal fibrosis with the HLA-DRB1*03 allele, which has been linked to autoimmune diseases such as systemic lupus erythematosus, type 1 diabetes mellitus, and myasthenia gravis.

Idiopathic retroperitoneal fibrosis is rare, with an incidence of 1.3 per 100,000 people per year. It most commonly affects individuals in the fourth to sixth decades of life. Some studies suggest a 2:1 male-to-female predominance, but others have not found a gender predilection. Patients with this condition present with nonspecific findings. Most patients are diagnosed after ureteral obstruction. Patients may complain of dull or acute back or flank pain. Systemic complaints include anorexia, weight loss, nausea, vomiting, fever, and malaise. Decreased urinary output may ensue if the ureters become fibrotic. Because of renal artery impingement, patients are commonly hypertensive on exam. Other physical exam findings include lower extremity edema and diminished lower extremity pulses if the vessels of the lower extremities are compressed. New hydrocele or varicocele can also be associated with fibrosis of the retroperitoneum. Laboratory analysis is also nonspecific. If there is renal obstruction, patients may have elevated serum blood urea nitrogen and creatinine. Erythrocyte sedimentation rate and C-reactive protein is elevated in most patients with retroperitoneal fibrosis. Antinuclear antibodies may also be elevated, highlighting the autoimmune nature of this disease.

Contrast-enhanced CT scan is the modality of choice to visualize the extent of disease. On CT scan, fibrotic retroperitoneum appears to have similar attenuation to muscle tissue (Fig. 35-13). The fibrotic mass can appear to encase the aorta and often compresses the inferior vena cava. If renal insufficiency predominates the use of IV contrast, MRI can be used to clarify the extent of fibrosis. Renal US can also be used to document ureteral compression and hydronephrosis. In some cases, a compressing mass lesion can also be found on abdominal US. Lower-extremity US may show deep venous thrombosis.

Pathologic examination of tissue is necessary to confirm the diagnosis and to rule out malignancy. Differential of retroperitoneal masses include lymphomas or sarcomas, desmoid

Figure 35-13. Computed tomography findings of retroperitoneal fibrosis. Arrows point to a soft tissue mass surrounding the aorta and common iliac arteries. Also present is hydronephrosis of the right kidney because of ureteral compression. (Reproduced with permission from Vaglio A, Salvarani C, Buzio C. Retroperitoneal Fibrosis, Lancet. 2006 Jan 21;367(9506):241-251.)
tumors, and infections such as tuberculosis. Biopsy can be obtained either via image-guided or surgical techniques, which may be performed laparoscopically or during an open procedure.

Once the diagnosis of retroperitoneal fibrosis is established, treatment is initiated. Surgical treatment is aimed at relieving ureteral obstruction either by ureterolysis or stenting. Surgery is reserved for patients who develop renal insufficiency as any surgical intervention in these patients carries several associated risks. In patients with secondary fibrosis, treatment is aimed at the underlying etiology or discontinuation of the causative medication. For patients with idiopathic retroperitoneal fibrosis, corticosteroids are the mainstay of treatment. Patients are usually started on one month of high-dose prednisone, followed by maintenance therapy and subsequent taper. Alternatively, tamoxifen can be used if patients have contraindications to steroids. Within days of instituting therapy, symptoms and inflammatory markers should improve. Over the course of a few weeks, renal function should improve. Interval imaging studies are usually obtained after 1 month to evaluate for efficacy. If steroids fail to cause regression of disease, patients can be started on immunosuppressant medications including methotrexate, azathioprine, cyclophosphamide, or mycophenolate mofetil. Recurrence of retroperitoneal fibrosis varies between studies from 10% to 30%. In patients who respond to steroid therapy, 5-year survival is as high as 90%.

BIBLIOGRAPHY

Entries highlighted in bright blue are key references.


Omentum


Mesentery


Avincos MO, Otani K, Kanzawa M, et al. Sclerosing mesenteritis: a real manifestation or histological mimic of


 **Retroperitoneum**


Sarcomas are a heterogeneous group of neoplasms that arise predominantly from cells of the embryonic mesoderm. While the majority of sarcomas are soft tissue sarcomas, other types of sarcoma include bone sarcomas (osteosarcoma, chondrosarcoma, and rare bone tumors like chordoma, angiosarcoma, and leiomyosarcoma of bone) and Ewing’s sarcoma/peripheral primitive neuroectodermal tumor, which can occur either in the bone or in the soft tissues. The primary focus of this chapter is soft tissue sarcomas. Most primary soft tissue sarcomas originate in an extremity (50–60%); the next most common sites are the trunk (19%), retroperitoneum (15%), and head and neck (9%). The anatomic site of a primary sarcoma influences treatment and outcome.\(^1\)

Soft tissue sarcomas include more than 70 histologic subtypes (Table 36-1). Historically, the most common subtypes in adults (excluding Kaposi’s sarcoma) were malignant fibrous histiocytoma (28%), liposarcoma (15%), leiomyosarcoma (12%), synovial sarcoma (10%), and malignant peripheral nerve sheath tumor (6%).\(^2\) Today, malignant fibrous histiocytoma is classified as either leiomyosarcoma, pleomorphic undifferentiated sarcoma, myxofibrosarcoma, or dedifferentiated liposarcoma based on cellular differentiation and genetics. Embryonal/alveolar rhabdomyosarcomas are the most common soft tissue sarcomas of childhood, whereas pleomorphic rhabdomyosarcoma occurs predominantly in adults, and although it shares part of the name, it has a different biology and should not be treated as a pediatric sarcoma.

During the past 25 years, patients with extremity sarcomas have been treated with a multimodality approach, which has led to some improvements in survival, local control, and quality of life.\(^3\) However, patients with abdominal sarcomas continue to have high rates of recurrence and poor overall survival.\(^4\) The overall 5-year survival rate for patients with all stages of soft tissue sarcoma is 50% to 60%. Of the patients who die of sarcoma, most succumb to lung metastasis, which 80% of the time occurs within 2 to 3 years after initial diagnosis.

In the United States in 2012, approximately 11,280 new cases of soft tissue sarcoma were diagnosed, and 3900 deaths were attributable to this disease.\(^5\) The incidence of soft tissue sarcomas increased 1.3% per year from 1995 to 2009, and subsequently leveled off while mortality rates increased slightly (0.5% per year) from 2001 to 2014.\(^6\) The true incidence of sarcoma is thought to be higher than reported, and gastrointestinal stromal tumors (GISTs) likely account for an additional 5000 new sarcoma cases per year.\(^1\) Overall, sarcomas affect 5 to 6 individuals per 100,000 inhabitants per year,\(^5\) accounting for less than 1% of all malignancies in adults and 15% of malignancies in children.\(^7\)
Key Points

1. Sarcomas are a heterogeneous group of tumors that can occur throughout the body and encompass more than 50 subtypes with distinct histologic lines of differentiation.

2. Approximately two-thirds of soft tissue sarcomas arise in the extremities; the remaining one-third is distributed between the retroperitoneum, trunk, abdomen, head, and neck.

3. Multimodality treatment, including surgical resection, radiation therapy, and, in selected cases, systemic chemotherapy, has been applied to patients with locally advanced, high-grade, extremity sarcomas.

4. Overall 5-year survival rate for patients with all stages of soft tissue sarcoma is 50% to 60%.

5. These rare tumors account for less than 1% of cancer in adults (estimated 10,000 cases per year in the United States) and represent 15% of cancers in children.

6. The treatment algorithm for soft tissue sarcomas depends on tumor stage, site, and histology.

7. Of the patients who die of sarcoma, most will succumb to metastatic disease in the lungs, which 80% of the time occurs within 2 to 3 years of the initial diagnosis.

8. Progress in the understanding of soft tissue sarcoma biology is crucial for the development of new treatments.

Epidemiology

Except for malignant peripheral nerve sheath tumors in patients with neurofibromatosis, sarcomas do not seem to result from progression or dedifferentiation of a benign soft tissue tumor. While most sarcomas are of unknown cause, a few sarcoma subtypes have been observed in settings suggesting etiology.

Radiation Exposure

External radiation therapy is a rare but well-established risk factor for soft tissue sarcoma that may be associated with radiation-induced mutations of the p53 gene.8 The incidence of sarcoma among patients who are often treated with radiation for cancer of the breast, cervix, ovary, testes, or lymphatic system is 8 to 50 times the general-population risk.9,10 In a review of 160 patients with postirradiation sarcomas, the most common histologic types were osteogenic sarcoma, pleomorphic undifferentiated sarcoma, angiosarcoma, and lymphangiosarcoma.9 The risk of developing a sarcoma increased with radiation dose, and the median time between radiation therapy and diagnosis of sarcoma was 10 years.9 A review of 44 patients with radiation-associated sarcomas identified between 1989 and 2009 noted that the average period from initial radiation treatment to diagnosis was 16 years and that radiation-associated sarcomas occurred most commonly in patients treated for breast cancer (36% of the patients in the series) and lymphoma (34% of the patients in the series).11 The 5-year overall survival rate for patients presenting without metastasis was 44%. A recent review of undifferentiated pleomorphic sarcoma characterizing outcomes between sporadic and radiation-associated presentation identified a higher incidence of local recurrence and worse overall and disease specific survival among patients who presented with radiation-associated undifferentiated pleomorphic sarcoma.12

Table 36-1

Relative frequency of histologic subtypes of soft tissue sarcoma

<table>
<thead>
<tr>
<th>HISTOLOGIC SUBTYPES</th>
<th>NO.</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liposarcoma</td>
<td>188</td>
<td>15</td>
</tr>
<tr>
<td>Leiomyosarcoma</td>
<td>148</td>
<td>12</td>
</tr>
<tr>
<td>Unclassified sarcoma</td>
<td>140</td>
<td>11</td>
</tr>
<tr>
<td>Synovial sarcoma</td>
<td>125</td>
<td>10</td>
</tr>
<tr>
<td>Malignant peripheral nerve sheath tumor</td>
<td>72</td>
<td>6</td>
</tr>
<tr>
<td>Rhabdomyosarcoma</td>
<td>60</td>
<td>5</td>
</tr>
<tr>
<td>Fibrosarcoma</td>
<td>38</td>
<td>3</td>
</tr>
<tr>
<td>Ewing sarcoma</td>
<td>25</td>
<td>2</td>
</tr>
<tr>
<td>Angiosarcoma</td>
<td>25</td>
<td>2</td>
</tr>
<tr>
<td>Osteosarcoma</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>Epithelioid sarcoma</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>Chondrosarcoma</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>Clear cell sarcoma</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>Alveolar soft part sarcoma</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Malignant hemangiopericytoma</td>
<td>5</td>
<td>0.4</td>
</tr>
</tbody>
</table>


Occupational Chemical Exposure

Exposure to herbicides such as phenoxyacetic acids and to wood preservatives containing chlorophenols has been linked to an increased risk of soft tissue sarcoma.13 Several chemical carcinogens, including thorium oxide (Thorotrast), vinyl chloride, and arsenic, have been associated with hepatic angiosarcomas.14

Trauma

Although patients with sarcoma often report a history of trauma, no causal relationship has been established. More often, a minor injury calls attention to a preexisting tumor.

Chronic Lymphedema

In 1948, Stewart and Treves first described the association between chronic lymphedema after axillary dissection and subsequent lymphangiosarcoma (Fig. 36-1).13 Lymphangiosarcoma has been estimated to occur in 0.07% of patients who undergo axillary node dissection.16 It also has been reported to occur after filarial infections and in the lower extremities of patients with congenital lymphedema.17,18 Lymphangiosarcoma is generally an aggressive tumor; average survival of patients with lymphangiosarcoma is 19 months.19
Figure 36-1. A 57-year-old with a chronic, progressive lymphedema of the left upper extremity developed lymphangiosarcoma 10 years after breast cancer treatment.

MOLECULAR PATHOGENESIS

Sarcomas can be broadly classified into three groups according to the genetic events underlying their development: specific translocations or gene amplification, defining oncogenic mutations, and complex genomic rearrangements.20 In general, sarcomas resulting from identifiable molecular events tend to occur in younger patients with histology suggesting a clear line of differentiation. The identifiable molecular events include point mutations, translocations causing overexpression of an autocrine growth factor, and oncogenic fusion transcription factor producing a cellular environment prone to malignant transformation. In contrast, sarcomas without identifiable genetic changes or expression profile signatures tend to occur in older patients and exhibit pleomorphic cytology and p53 dysfunction.21 Improved understanding of the molecular pathogenesis of sarcomas has revealed several potential targets against which investigators are working to develop subtype-specific targeted therapy.

Translocation-Associated Sarcomas

To date, translocations have been identified in 14 subtypes of soft tissue sarcoma, accounting for 20% to 30% of all sarcomas22 (Table 36-2). Translocations result in in-frame gene fusion, which in turn results in fused products encoding oncoproteins that function as transcriptional activators or repressors.23,24 The best characterized gene fusions are in Ewing’s sarcoma (EWS-FLI1), clear cell sarcoma (EWS-ATF1), myxoid/round cell liposarcoma (TLS-CHOP), alveolar rhabdomyosarcoma (PAX3-FKHR), desmoplastic small round cell tumor (EWS-WT1), and synovial sarcoma (SS18-SSX). Fusion gene–related sarcomas have been estimated to account for 30% or more of all sarcomas.25 Direct or indirect interactions between fusion transcripts and cell cycle regulators have been elucidated by several investigators and identify these transcripts as potentially promising molecular therapeutic targets.26 However, fusion genes in sarcoma have been successfully targeted in only a few cases, in which fusion resulted in overexpression of a growth factor or growth factor receptor. Several growth factors and their receptors (e.g., epidermal growth factor receptor) previously reported to play an important role in autocrine or paracrine stimulation of carcinoma growth have been associated with high histologic grade and poor prognosis in soft tissue sarcomas.

Amplification-Associated Sarcomas

Oncogenes are genes that can induce malignant transformation and tend to drive cell proliferation. Several oncogenes have been associated with soft tissue sarcomas, including MDM2, N-myrc, c-erbB2, and members of the ras family. These oncogenes produce specific oncoproteins that either play a role in nuclear function and cellular signal transduction or function as growth factors or growth factor receptors. This typically occurs in dedifferentiated liposarcoma, where the amplification of MDM2 drives the neoplastic process. Amplification of these genes has been shown to correlate with adverse outcome in several types of soft tissue sarcoma.23

Oncogenic Mutations

GISTs are the classic example of sarcomas in which tumorigenesis is primarily driven by a single activating mutation, in the

Table 36-2

Fusion transcripts in soft tissue sarcoma

<table>
<thead>
<tr>
<th>DIAGNOSIS</th>
<th>CHROMOSOMAL ABNORMALITY</th>
<th>GENES INVOLVED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alveolar rhabdomyosarcoma</td>
<td>t(2;13)(q35;q14)</td>
<td>PAX3-FKHR</td>
</tr>
<tr>
<td></td>
<td>t(1;13)(p36;q14)</td>
<td>PAX7-FKHR</td>
</tr>
<tr>
<td>Alveolar soft part sarcoma</td>
<td>t(X;17)(p11.2;q25)</td>
<td>TFE3-ASPL</td>
</tr>
<tr>
<td>Angiomatoid fibrous histiocyto</td>
<td>t(12;16)(q13:p11)</td>
<td>FUS-ATFI</td>
</tr>
<tr>
<td>Clear cell sarcoma</td>
<td>t(12;22)(q13;q12)</td>
<td>EWS-ATFI</td>
</tr>
<tr>
<td>Congenital fibrosarcoma/</td>
<td>t(12;15)(p13;q25)</td>
<td>ETV6-NTRK3</td>
</tr>
<tr>
<td>congenital mesoblastic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nephroma</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dermatofibrosarcoma</td>
<td>t(17;22)(q22;q13)</td>
<td>PDFGB-COL1A1</td>
</tr>
<tr>
<td>protuberans</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desmoplastic small round</td>
<td>t(11;22)(p13;q12)</td>
<td>EWS-WT1</td>
</tr>
<tr>
<td>cell tumor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endometrial stromal</td>
<td>t(7;17)(p15;q21)</td>
<td>JAZF1-JJAZ1</td>
</tr>
<tr>
<td>sarcoma</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ewing’s sarcoma/</td>
<td>t(11;22)(q24;q12)</td>
<td>EWS-FLI1</td>
</tr>
<tr>
<td>peripheral primitive</td>
<td>t(21;22)(q22;q12)</td>
<td>EWS-ERG</td>
</tr>
<tr>
<td>neuroectodermal tumor</td>
<td>t(7;22)(p22;q12)</td>
<td>EWS-ETV1</td>
</tr>
<tr>
<td></td>
<td>t(17;22)(q12;q12)</td>
<td>EWS-FEV</td>
</tr>
<tr>
<td></td>
<td>t(2;22)(q33;q12)</td>
<td>EWS-E1AF</td>
</tr>
<tr>
<td></td>
<td>t(16;21)(p11;q22)</td>
<td>FUS-ERG</td>
</tr>
<tr>
<td>Low-grade fibromyxoid</td>
<td>t(7;16)(q33;p11)</td>
<td>FUS-CREB3I2</td>
</tr>
<tr>
<td>sarcoma</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inflammatory myofibroblastic</td>
<td>t(2;12)(q22;p23)</td>
<td>TPM3-ALK</td>
</tr>
<tr>
<td>tumor</td>
<td>t(2;19)(p23;p13)</td>
<td>TPM4-ALK</td>
</tr>
<tr>
<td></td>
<td>t(2;17)(p23;q23)</td>
<td>CLTC-ALK</td>
</tr>
<tr>
<td>Myxoid liposarcoma</td>
<td>t(12;16)(q13;p11)</td>
<td>TLS-CHOP</td>
</tr>
<tr>
<td></td>
<td>t(12;22)(q13;q12)</td>
<td>EWS-CHOP</td>
</tr>
<tr>
<td>Myxoid chondrosarcoma</td>
<td>t(9;22)(q22;q12)</td>
<td>EWS-CHN</td>
</tr>
<tr>
<td></td>
<td>t(9;15)(q22;q21)</td>
<td>TFC12-CHN</td>
</tr>
<tr>
<td></td>
<td>t(9;17)(q22;q11)</td>
<td>TAF2N-CHN</td>
</tr>
<tr>
<td>Synovial sarcoma</td>
<td>t(x;18)(p11;q11)</td>
<td>SXX1-SYT</td>
</tr>
<tr>
<td></td>
<td>t(xx;2)(q23;q12)</td>
<td>SXX2-SYT</td>
</tr>
<tr>
<td></td>
<td>t(x;4)(q31;q12)</td>
<td>SXX4-SYT</td>
</tr>
</tbody>
</table>
gene encoding KIT receptor tyrosine kinase or platelet-derived growth factor receptor-α (PDGFRα). The majority of GISTs have mutations in either exon 11 or exon 9 of KIT and respond dramatically to the tyrosine kinase inhibitor imatinib mesylate, although this treatment rarely produces cure.

**Complex Genomic Rearrangements**

The largest group of sarcomas is the group with complex cytogenetic alterations, which includes high-grade spindle cell sarcomas and pleomorphic sarcomas. Many sarcomas in this group exhibit inactivation of tumor suppressor genes. The two genes most relevant to soft tissue sarcoma are retinoblastoma (Rb) and p53. Mutations or deletions in Rb can lead to retinoblastoma, the most common malignant ocular neoplasm of childhood. Survivors of retinoblastoma are at risk for developing soft tissue and bone sarcomas later in life. Patients with germline mutations in p53 (Li-Fraumeni syndrome) have a high incidence of soft tissue sarcomas. Mutant p53 expression is thought to correlate with poor overall survival.

*Neurofibromatosis type 1 (von Recklinghausen’s disease)* occurs in approximately 1 of every 3000 people and is due to various mutations in the *NF1* tumor suppressor gene, located on chromosome 17. Patients with neurofibromatosis type 1 have an estimated 3% to 15% additional risk of malignant disease compared with the general population lifetime risk, including malignant peripheral nerve sheath tumors (MPNST) and GIST. In turn, 25% to 50% of patients with MPNST have a mutation in *NF1*.

**INITIAL ASSESSMENT**

The clinical behavior of most soft tissue sarcomas is determined by anatomic location (depth in relation to the investing fascia), histologic subtype and grade of aggressiveness, and size. The dominant pattern of metastasis is hematogenous, primarily to the lungs. Lymph node metastasis is rare (affecting <5% of patients) except in a few histologic subtypes, including epithelioid sarcoma, pediatric rhabdomyosarcoma, clear cell sarcoma, angiosarcoma, and, more rarely, synovial sarcoma and myxofibrosarcoma.

**Clinical Presentation**

Soft tissue sarcomas most commonly present as an asymptomatic mass. Extremity sarcomas may present as a deep venous thrombosis, particularly in patients without significant risk factors for thrombosis. Tumors in the distal extremities are generally smaller, whereas tumors in the proximal extremities and retroperitoneum can grow quite large before becoming apparent. Tumors often grow centrifugally and can compress surrounding normal structures. Infrequently, tumor impingement on bone or neurovascular bundles produces pain, edema, and swelling. Less frequently, tumors cause obstructive gastrointestinal symptoms or neurologic symptoms related to compression of lumbar or pelvic nerves. Often an extremity mass is discovered after a traumatic event that draws attention to a preexisting lesion.

The differential diagnosis of a soft tissue mass should include consideration of lipoma (which is 100 times more common than sarcoma), lymphangioma, leiomyoma, neurinoma, primary or metastatic carcinoma, melanoma, and lymphoma. Superficial small lesions (<5 cm) that are new or that are not enlarging as indicated by clinical history can be observed. Enlarging masses and masses larger than 5 cm or deep to the fascia should be evaluated with a history, imaging, and biopsy.

**Diagnostic Imaging**

Diagnostic imaging of the primary should be performed before any invasive procedure to avoid the possibility of soft tissue swelling or hemorrhage complicating the image interpretation. Pretreatment diagnostic imaging is helpful for defining the size and anatomic location of a tumor and its proximity to adjacent structures; staging disease with respect to regional or metastatic spread; guiding percutaneous biopsy; and establishing whether a tumor is benign or malignant and low grade or high grade.

Radiographs are useful in the evaluation of primary bone tumors but not in the evaluation of soft tissue sarcomas of the extremities unless there is underlying bone involvement from an adjacent soft tissue tumor or mineralization patterns suggestive of histologic subtype. Magnetic resonance imaging (MRI) is the preferred imaging technique for soft tissue sarcomas of the extremities, whereas computed tomography (CT) is most useful for evaluating retroperitoneal, intra-abdominal, and truncal sarcomas. CT of the chest should be performed to assess for lung metastases in patients with high-grade tumors larger than 5 cm; while chest X-ray is sufficient for smaller or low-grade lesions. Abdominal/pelvic CT should be performed in patients with myxoid round cell liposarcomas, leiomyosarcomas, epithelioid sarcomas, or angiosarcomas because of their propensity to metastasize to the abdomen and/or pelvis. Total Spine MRI has been advocated for myxoid round cell liposarcoma. MRI of the brain should be considered for patients with alveolar soft part sarcomas and angiosarcomas because of their propensity to metastasize to the brain.

**Ultrasoundography.** Ultrasoundography may have a diagnostic role in patients with soft tissue sarcoma who cannot undergo MRI. Ultrasoundography can also be a useful adjunct to MRI when findings on MRI are indeterminate and for delineating adjacent vascular structures. Finally, ultrasoundography can be used for postoperative surveillance and to guide biopsies.

**Computed Tomography.** Chest CT should be performed to evaluate for lung metastasis at presentation and before any radical treatment. CT is also the preferred imaging technique for evaluating retroperitoneal sarcomas (Fig. 36-2).

![Figure 36-2. A 55-year-old man with a leiomyosarcoma involving the inferior vena cava. Note the displacement of the inferior vena cava to the right hemiabdomen adjacent to the liver (arrow).](image-url)
techniques can provide a detailed image of the abdomen and pelvis and can delineate adjacent organs and vascular structures. For extremity sarcomas, CT may be useful if MRI is not available or cannot be used. When histologic assessment of an extremity sarcoma reveals a myxoid liposarcoma, CT of the abdomen and pelvis should be done because this subtype is known to metastasize to the abdomen.32

**Magnetic Resonance Imaging.** MRI is the most useful imaging modality for extremity sarcomas because of its superior soft tissue contrast resolution and multiplanar capabilities. MRI accurately delineates muscle groups and distinguishes among bone, vascular structures, and tumor. Sagittal and coronal views allow evaluation of anatomic compartments in three dimensions (Fig. 36-3). Soft tissue sarcomas of the extremities usually present on MRI as a heterogeneous mass. Their signal intensity tends to be equal to or slightly higher than that of adjacent skeletal muscle on T1-weighted images and heterogeneous and high on T2-weighted images. Hemorrhagic, cystic, or necrotic changes may also be observed in the tumor. If adjacent vascular structures must be delineated, special MRI techniques may be performed, including magnetic resonance angiography. MRI may also be an important adjunct to cytologic analysis in distinguishing benign lesions such as lipomas, hemangiomas, schwannomas, neurofibromas, and intramuscular myxomas from their malignant counterparts. In patients undergoing preoperative chemotherapy, contrast-enhanced T1-weighted MRI can be useful in evaluating intratumoral necrosis.

MRI is also valuable for assessing tumor recurrence after surgery. A baseline image is usually obtained 3 months after surgery. Some clinicians forego routine postoperative imaging of the primary extremity tumor site in asymptomatic patients, citing the difficulties in detecting early recurrence in scarred, irradiated tissue.31 Others advocate routine imaging every 3 to 4 months for the first 2 years, every 6 months in years 3 through 5, and then annually.

**Positron Emission Tomography.** Positron emission tomography (PET) is a functional imaging modality that measures tumor uptake of the glucose analog [18F] fluorodeoxyglucose (FDG). PET imaging allows evaluation of the entire body. Although PET/CT may be useful in specific circumstances, FDG-PET is not currently recommended for the initial staging of patients with soft tissue sarcoma.

Roberge and colleagues compared FDG-PET/CT versus chest CT alone in the initial staging of 75 patients with soft tissue sarcoma and found that only one patient had disease upstaged as a result of PET, whereas two had false-positive findings and three had indeterminate findings with no subsequent development of metastasis.33 Previous studies that reported a marginal benefit of PET/CT for detecting metastasis at the time of sarcoma staging included patients with more heterogeneous tumors, such as osseous tumors, soft tissue osteosarcomas, Ewing’s sarcoma, and rhabdomyosarcoma.34-36

In patients with sarcoma, PET has primarily been used to assist with tumor grading and to assess response to chemotherapy.37-40 In 50 patients with resectable high-grade soft tissue tumors scheduled for preoperative chemotherapy and tumor resection, a 35% or greater reduction in tumor FDG uptake following an initial cycle of chemotherapy was associated with histopathologic tumor response defined as pathologic necrosis in 95% or more of the resected specimen.41 While this is helpful in determining response controversy exists as to whether this translates into a predictor of overall survival.

**Biopsy Techniques**

**Fine-Needle Aspiration.** At centers where cytopathologists have experience with evaluation of mesenchymal tumors, fine-needle aspiration is an acceptable method of diagnosing most soft tissue sarcomas, particularly when the results correlate closely with clinical and radiologic findings.42 Fine-needle aspiration of primary tumors has a lower diagnostic accuracy rate (60–90%) than core needle biopsy and is often not sufficient for establishing a specific histologic diagnosis and grade.43 However, fine-needle aspiration is the procedure of choice to confirm or rule out the presence of a metastatic focus or local recurrence.44 Although fine-needle aspiration of superficial lesions can often be done in the clinic, fine-needle aspiration of deeper tumors may need to be done by an interventional radiologist.

![Figure 36-3](image_url). A 62-year-old man presented with right thigh mass. Magnetic resonance imaging demonstrated an 18 × 15 cm² dedifferentiated liposarcoma within the posterior compartment. Note the atypical fatty mass (left) with a large necrotic and peripherally enhancing nodule (left).
under sonographic or CT guidance. Generally, a 21- to 23-gauge needle is introduced into the mass after appropriate cleansing of the skin and injection of local anesthetic. Negative pressure is applied, and the needle is moved back and forth several times in various directions. After the negative pressure is released, the needle is withdrawn, and the contents of the needle are used to prepare smears. A cytopathologist then examines the slides to determine whether sufficient diagnostic material is present.

**Core Needle Biopsy.** Core needle biopsy is safe, accurate, and economical, and has become the preferred technique for diagnosing soft tissue lesions. Dupuy and colleagues found that core needle biopsy had an accuracy of 93% in 221 patients with musculoskeletal neoplasms.

Image guidance (ultrasound or CT) can prevent sampling of nondiagnostic necrotic or cystic areas of the tumor and thus increase the positive yield rate. Image guidance also permits biopsy of tumors in otherwise inaccessible locations and tumors located near vital structures with less risk of injury or complication.

The tissue sample obtained from core needle biopsy is usually sufficient for several diagnostic tests, such as electron microscopy, cytogenetic analysis, and flow cytometry. The risk for needle track seeding is negligible, and the reported complication rate for core needle biopsy is less than 1%.

**Incisional Biopsy.** Historically, an open surgical biopsy was the gold standard for achieving adequate tissue for definitive and specific histologic diagnosis of bone or soft tissue sarcomas. Contemporary guidelines recommend incisional biopsy when core needle biopsy cannot produce adequate tissue for diagnosis or when findings on core needle biopsy are nondiagnostic.

The disadvantages of incisional biopsy include the need to schedule the procedure, the need for general anesthesia, and high costs. In addition, an inappropriately placed incision can necessitate more extensive definitive resection to incorporate the biopsy incision. In a series of 107 patients with soft tissue sarcoma, planned surgical treatments had to be changed because of poorly oriented biopsies in 25% of cases. Complication rates up to 17% have been reported after incisional biopsies. Potential complications include hematoma, infection, wound dehiscence, and tumor fungation, any of which can delay definitive treatment.

Incisional biopsies should be performed only by surgeons experienced in the management of soft tissue sarcoma, ideally in a center specializing in the treatment of sarcoma and by the surgeon who will perform the definitive surgery. The biopsy incision should be oriented longitudinally along the extremity to allow a subsequent wide local excision that encompasses the biopsy site, scar, and tumor en bloc minimizing the risk for increasing the complexity of the subsequent curative procedure. A poorly oriented biopsy incision often necessitates an excessively large surgical defect for a wide local excision, which in turn can result in a larger postoperative radiation therapy field to encompass all tissues at risk. Adequate hemostasis must be achieved at the time of biopsy to prevent dissemination of tumor cells into adjacent tissue planes by hematoma.

**Excisional Biopsy.** Excisional biopsy can be performed for easily accessible (superficial) extremity or truncal lesions smaller than 3 cm. However, excisional biopsy rarely provides benefits over other biopsy techniques. Excisional biopsies should not be performed for lesions involving the hands and feet because such biopsies may complicate definitive reexcision.

For sarcomas with initial diagnosis confirmed with excisional biopsy, microscopic residual disease has been reported in up to 69% of reexcision specimens; without reexcision, the reported rate of local recurrence is 30% to 40% when margins are positive or uncertain.

Wide en bloc excision is seldom performed as a diagnostic procedure. When en bloc excision is done for diagnosis, the margin status is often not adequately evaluated during pathologic assessment of the specimen. Unless detailed descriptions of the surgical procedure and the pathology specimen are provided, the margins should be classified as uncertain or unknown, a classification associated with the same prognosis as resection margins that are positive for tumor cells. In patients with uncertain or unknown margins, reexcision should be performed if possible to ensure negative margins. The biopsy site or tract (if applicable) should be included en bloc with the re-resected specimen.

**Pathologic Assessment and Classification**

Sarcoma is generally diagnosed by morphologic assessment based on microscopic examination of histologic sections by an experienced sarcoma pathologist. However, even expert sarcoma pathologists disagree on the specific histologic diagnosis and the tumor grade in 25% to 40% of cases.

Morphologic assessment can be supported by ancillary techniques, including conventional cytogenetics; immunohistochemistry; and molecular genetic testing, which is useful for classifying soft tissue sarcoma subtypes with multiple genetic aberrations. Other molecular diagnostic techniques include cytogenetic analysis, fluorescence in situ hybridization, and polymerase chain reaction–based methods. However, molecular genetic techniques are associated with significant technical limitations and should be interpreted in the context of the sarcoma’s morphologic features.

Some experts have suggested that pathologic classification of soft tissue sarcomas has more prognostic significance than does tumor grade when other pretreatment variables are taken into account. Tumors with limited metastatic potential include desmoid, atypical lipomatous tumor (also called well-differentiated liposarcoma), dermatofibrosarcoma protuberans, and solitary fibrous tumor. Tumors with an intermediate risk of metastatic spread usually have a large myxoid component and include myxoid liposarcoma, myxofibrosarcoma, and extraskelatal myxoid chondrosarcoma. Among the highly aggressive tumors with substantial metastatic potential are angiosarcoma, clear cell sarcoma, pleomorphic and dedifferentiated liposarcoma, leiomyosarcoma, MPNST, rhabdomyosarcoma, and synovial sarcoma.

It has recently been noted that malignant fibrous histiocytoma is not associated with a distinct gene cluster, suggesting that malignant fibrous histiocytoma is not a separate tumor entity but rather a common morphologic appearance of various sarcoma subtypes. For example, most tumors initially diagnosed as malignant fibrous histiocytoma in the retroperitoneum have been reclassified using genomic profiling as dedifferentiated liposarcoma, whereas those in the extremities have been reclassified as leiomyosarcoma, myxofibrosarcoma, or pleomorphic undifferentiated sarcoma.

Guidelines for the pathologic reporting of sarcoma have been established. Included in the report should be the primary diagnosis, anatomic site, depth, size, and histologic grade, presence or absence of necrosis, status of excision margins and lymph nodes, TNM stage, and additional features of the tumor (i.e., mitotic rate and presence or absence of vascular invasion).
Staging and Prognostic Factors

Soft tissue sarcoma is most commonly staged using either the American Joint Committee on Cancer (AJCC) system (generally used in the United States) or the World Health Organization system. A unique aspect of sarcoma staging is the inclusion of tumor grade, which is one of the most important prognostic factors.57

The seventh edition of the AJCC staging system for soft tissue sarcomas is based on histologic grade of aggressiveness, tumor size and depth, and the presence of nodal or distant metastases.58 This system does not apply to GIST, fibromatoses (desmoid tumor), Kaposi’s sarcoma, or infantile fibrosarcoma.

Histologic Grade of Aggressiveness. Histologic grade is the most important prognostic factor for patients with soft tissue sarcoma. For accurate determination of grade, an adequate tissue sample must be appropriately fixed, stained, and reviewed by an experienced sarcoma pathologist. The features that define grade are cellularity, differentiation (good, moderate, or poor/anaplastic), pleomorphism, necrosis (absent, <50%, or ≥50%), and amount of mitoses per high-power field (<10, 10–19, or ≥20). Tumor grade has been shown to predict metastasis and overall survival.59 Metastasis has been estimated to occur in 5% to 10% of low-grade lesions, 25% to 30% of intermediate-grade lesions, and 50% to 60% of high-grade lesions.

The number of grades varies according to the classification system used. The most common classification systems, those of the National Cancer Institute and the French Federation of Cancer Centers, use three-tier tumor grades.60 The National Cancer Institute system is based primarily on histologic subtype, location, and amount of necrosis. The French Federation of Cancer Centers system is based on tumor differentiation (good, moderate, or poor/anaplastic), number of mitoses per high-power field (<10, 10–19, or ≥20), and amount of tumor necrosis (absent, <50%, or ≥50%). A comparative analysis of the two systems suggested that the French Federation of Cancer Centers system has better prognostic capability, predicting 5-year survival rates of 90%, 70%, and 40% for grade 1, 2, and 3 tumors, respectively.60

Following the recommendation of the College of American Pathologists, the committee that developed the 2008 AJCC staging system changed the system from a four-grade to a three-grade system in which the grades are well differentiated (grade 1), moderately differentiated (grade 2), and poorly differentiated (grade 3).61 Grade 1 is considered low grade, and grades 2 and 3 are considered high grade.

Tumor Size and Location. Tumor size is an important prognostic variable in soft tissue sarcomas. Sarcomas have classically been stratified into two size groups; T1 lesions are 5 cm or smaller, and T2 lesions are larger than 5 cm. The staging of soft tissue sarcomas has changed considerably in the eighth edition of the AJCC Cancer Staging Handbook.62 Previously, all soft tissue sarcoma sites were staged as one. In the new edition, subsites have been created, and they include head and neck; extremity and trunk; gastrointestinal tract; genitourinary tract; viscera and peritoneum; gynecological sites; breast lung, pleura and mediastinum; and other histologies. Size criteria for head and neck tumors have been reduced from the classic criteria because of a worse prognosis at smaller sizes. Trunk and extremity sarcoma has been reclassified relative to size criteria for tumor stage, changing the four categories: T1 tumors are smaller than 5 cm; T2 tumors are 5 to 10 cm; T3 tumors are 10 to 15 cm; and T4 tumors are larger than 15 cm.62

Anatomic tumor location was incorporated into the AJCC staging system in 1998. Soft tissue sarcomas above the superficial investing fascia of the extremity or trunk are designated “a” lesions within the T category, whereas tumors invading or deep to the fascia and all retroperitoneal, mediastinal, and visceral tumors are designated “b” lesions. However, in this most recent AJCC staging system, the superficial versus deep distinction is less important and has been eliminated.62

Nodal Metastasis. Overall, lymph node metastases arising from soft tissue sarcomas are rare,29 but the incidence of nodal involvement is higher for epithelioid sarcoma, pediatric rhabdomyosarcoma, clear cell sarcoma, synovial sarcoma, myxofibrosarcoma, and angiosarcoma. In the seventh edition of the AJCC staging system, sarcoma associated with nodal metastases was reclassified as stage III rather than stage IV because several studies reported better survival for patients with isolated regional lymph node metastases treated with radical lymphadenectomy than for patients with distant metastases.28,63-65 However, in the eighth edition of the AJCC, nodal disease has been revisited: N1 disease behaves similarly between stages III and IV and is now captured as stage IV.62 Patients with clinically or radiologically suspicious regional nodes should have metastases confirmed or ruled out by either fine-needle aspiration or core biopsy before radical lymphadenectomy.

Distant Metastasis. Distant metastases occur most often in the lungs (Fig. 36-4). Selected patients with pulmonary metastases may survive for long periods after surgical resection and chemotherapy. Other potential sites of metastasis include bone (Fig. 36-5), brain (Fig. 36-6), and liver (Fig. 36-7). Visceral and retroperitoneal sarcomas have a higher incidence of liver and peritoneal metastases.

Prognostic Factors. Prognostic variables in soft tissue sarcoma include primary tumor size, grade, and depth, all of which are incorporated into the staging system, as well as histology.
tumor site, and presentation (local recurrence or initial diagnosis). Patient factors such as older age and gender have also been associated with recurrence and mortality in several studies. A positive microscopic margin and early recurrence after resection of an extremity sarcoma have been shown to be associated with decreased survival.

Several groups have reported that Ki-67, a proliferation marker, is correlated with a poor clinical outcome in high-grade extremity sarcomas. E-cadherin and catenins, proteins essential for intercellular junctions, have been associated with poor outcome in patients with soft tissue sarcoma. Similarly, higher CD100 expression has been shown to correlate with higher proliferative potential and poorer outcome.

Prognostic Nomograms. Prognostic nomograms for soft tissue sarcoma have been introduced for use in patient counseling, selecting appropriate surveillance strategies, and selecting patients for clinical trials. One such nomogram, developed by Kattan and colleagues at Memorial Sloan-Kettering Cancer Center, considers age, histology, grade, location, depth, and size to determine the likelihood of 12-year sarcoma-specific survival. Two validation studies using the nomogram demonstrated good predictive value. More recently, the same group of investigators developed histology subtype-specific nomograms for patients with liposarcoma, synovial sarcoma, and GIST and demonstrated that they were accurate in predicting disease-specific survival. Other investigators have just developed a site-specific nomogram for patients with retroperitoneal sarcoma, demonstrating an accurate prediction of survival and disease recurrence.

TREATMENT OF EXTREMITY AND TRUNK WALL SARCOMA

The goals of treatment of soft tissue sarcoma are to maximize the likelihood of long-term recurrence-free survival while minimizing morbidity and maximizing function. In the past two decades, a multimodality treatment approach with optimal sequencing of treatments for individual patients has been shown to improve survival. Furthermore, patients with soft tissue sarcoma treated at high-volume centers have been shown to have improved survival and functional outcomes. Care at such centers is particularly important for patients with high-risk and advanced disease.

The overall 5-year survival rate for patients with all stages of soft tissue sarcoma is 50% to 60%. For patients with extremity sarcomas, a multidisciplinary treatment approach has resulted in local control rates exceeding 90% and 5-year survival rates exceeding 70%. Most patients who die of soft tissue sarcoma die of metastatic disease, which becomes evident within 2 to 3 years of initial diagnosis in 80% of cases.

Recommendations for evaluation and treatment of patients presenting with soft tissue masses are summarized in Table 36-3.
SOFT TISSUE SARCOMAS

CHAPTER 36

Table 36-3
Recommendations for the management of soft tissue masses

1. Soft tissue tumors that are enlarging or greater than 3 cm should be evaluated with radiologic imaging (ultrasonography or computed tomography [CT]), and a tissue diagnosis should be made using core needle biopsy.
2. Once a sarcoma diagnosis is established, obtain imaging (magnetic resonance imaging for extremity lesions and CT for other anatomic locations) and evaluate for metastatic disease with chest CT for intermediate- or high-grade (grade 2 or 3) or large (T2) tumors.
3. A wide local excision with 1- to 2-cm margins is adequate therapy for low-grade lesions and T1 tumors.
4. Radiation therapy plays a critical role in the management of large (T2), intermediate- or high-grade tumors.
5. Patients with locally advanced high-grade sarcomas or distant metastases should be evaluated for chemotherapy.
6. An aggressive surgical approach should be taken in the treatment of patients with an isolated local recurrence or resectable distant metastases.

Surgery

Primary tumors with no evidence of distant metastasis are managed with surgery alone or, when wide pathologic margins cannot be achieved because of anatomic constraints and/or the grade is high, surgery plus radiation therapy. The type of surgical resection is determined by several factors, including tumor location, tumor size, depth of invasion, involvement of nearby structures, need for skin grafting or autogenous tissue reconstruction, and the patient’s performance status. In 1985, the National Institutes of Health developed a consensus statement recommending limb-sparing surgery for most patients with high-grade extremity sarcomas. However, for patients with primary or recurrent tumors that cannot be grossly resected with a limb-sparing procedure and preservation of function (<5% of patients), amputation remains the treatment of choice.

Margin status after surgical resection has been shown to be an independent prognostic factor. The goal of surgical resection is to achieve a complete resection because microscopically positive or grossly positive resection margins are associated with increased risk of local recurrence and death. If an unexpected positive margin is found on pathologic examination of the resection specimen, reexcision should be performed if feasible. In patients with a positive margin, particularly in patients with macroscopic residual disease, local control is unlikely even with the addition of postoperative radiation therapy, emphasizing the importance of a well-planned initial operation.

Wide Local Excision. The preferred treatment for extremity sarcomas is wide local excision that includes resection of the biopsy site. The goal of wide local excision is to remove the tumor with approximately 1 to 2 cm of surrounding normal soft tissue, but narrower margins may be necessary to preserve uninvolved critical neurovascular structures and may be adequate for patients undergoing radiation therapy. Dissection should proceed through grossly normal tissue planes not abutting the tumor. Soft tissue sarcomas are generally surrounded by a zone of compressed reactive tissue that forms a pseudocapsule; this pseudocapsule should not be used to guide resection (enucleation) as microscopic disease exists within this reactive zone. If the tumor is adjacent to or displacing major neurovascular structures, these do not need to be resected, but the adventitia or perineurium should be removed. For some massive tumors of the extremities, wide local excision entails a radical or complete anatomic compartment resection. Surgical clips should be placed to delineate the extent of the resection bed for patients likely to require postoperative radiation therapy.

Recent reports demonstrate encouraging results following radical en bloc resection with vascular reconstruction in the lower extremities. While en bloc resection with vascular reconstruction has been associated with increased rates of postoperative complications, reported local recurrence and 5-year survival rates are similar to those for patients not requiring vessel resection. Similarly, studies have shown acceptable functional outcomes with resection of the sciatic, tibial, and peroneal nerves with appropriate reconstruction and rehabilitation.

Bone invasion from extremity soft tissue sarcoma, which can generally be identified using high-quality cross-sectional imaging such as MRI, has been estimated to occur in about 5% of patients and is associated with reduced overall survival. In cases of bone invasion, bone resection is required to obtain an adequate surgical margin and to achieve local control. Although tumor resection and repair of skeletal defects are possible, the likelihood of postoperative complications may be increased, and functional outcomes may be less favorable. Lin and colleagues recently analyzed 55 patients with soft tissue sarcomas abutting bone and reported that in the absence of frank cortical bone penetration, periosseous was an adequate surgical margin in patients treated with wide local excision and radiation.

Soft tissue sarcomas arising in the distal extremities, particularly the hands and feet, present unique technical challenges. While distal-extremity tumors are often detected at a smaller size (<5 cm) than proximal-extremity tumors, resection and reconstruction techniques are often more complex for distal-extremity tumors, and preoperative planning is critical to obtain favorable functional outcomes. Identifying the proximity of the tumor to underlying critical structures (e.g., bone, tendon, or neurovascular structures) using MRI is essential for surgical planning. In a reported series of patients with sarcomas of the hands or feet treated with limited surgery only, 32% of patients had local recurrences. Preservation of function and acceptable recurrence rates with limited surgery and adjuvant radiation therapy for soft tissue sarcomas of the distal extremities have been reported. For locally advanced tumors, repair of bone defects, vascular reconstruction, tendon transfers, and soft tissue reconstruction using regional or free flaps have resulted in good functional outcomes. Amputation remains a reasonable option for patients with soft tissue sarcomas of the distal extremities when acceptable oncologic or functional outcomes cannot be achieved using available limb salvage techniques.

In an interesting study conducted in Ontario and Quebec, investigators found patients expecting a difficult recovery and patients with uncertain expectations had worse functional outcomes than patients anticipating an easy recovery, indicating that preoperative education including consultation with rehabilitation services may optimize outcomes. Furthermore, all patients undergoing resection of extremity sarcomas should undergo physical therapy beginning immediately after surgery and continuing until maximum function is achieved.
Locoregional Lymphadenectomy. Several studies have reported improved survival for patients with isolated regional lymph node metastases treated with radical lymphadenectomy.²⁶,₆₃-₆₅ Patients with clinically or radiologically suspicious regional nodes should have metastases confirmed by biopsy before radical lymphadenectomy. An ultrasound-guided fine-needle aspiration or core biopsy of lymph nodes in selected patients with suspicious clinical or radiologic findings. The utility of sentinel lymph node biopsy has remained controversial despite the recognition that several histologic subtypes of high-grade sarcoma are known to have a propensity for lymph node metastasis. However, there have been no prospective studies of the sensitivity and specificity of sentinel lymph node biopsy for such tumors, and as such, sentinel node biopsy for sarcoma should be performed in either highly selected patients or in the setting of a clinical trial.

Amputation. Amputation is the treatment of choice for the 5% of patients with primary or recurrent extremity tumors whose tumors cannot be grossly resected with limb-sparing procedures and preservation of function. Historically, local excision of large, high-grade soft tissue sarcomas resulted in local failure rates of 50% to 70%, even when a margin of normal tissue around the tumor was excised; consequently, radical resection or amputation was recommended. Today, however, the addition of radiation therapy to less radical surgical resection has made limb salvage possible in most cases.

A comparison of amputation versus limb-sparing surgery followed by adjuvant radiation therapy performed by the National Cancer Institute between 1975 and 1981 demonstrated no significant difference between the two groups in local recurrence or overall survival rate.⁹³ Potter and colleagues⁵⁰ later reviewed the entire National Cancer Institute experience with 123 patients treated with conservative surgery plus radiation therapy and 83 treated with amputation. The local recurrence rate was significantly higher in the surgery and adjuvant radiation therapy group: 8% versus 0% in the amputation group. However, survival rates did not differ between the groups. Several large single-institution studies have since also reported favorable local control rates with conservative resection plus radiation therapy.⁹⁴-⁹⁶

Isolated Regional Perfusion. Isolated regional perfusion is a limb-sparing technique in which a soft tissue sarcoma is perfused with high concentrations of tumor necrosis factor-alpha (TNF-α) and melphalan under hyperthermic conditions. The use of TNF-α is not approved by the U.S. Food and Drug Administration (FDA) and is used only in European countries. The technique is generally used for locally advanced, multifocal, or locally recurrent disease; it has also served as a palliative treatment to achieve local control for patients with distant metastases.

Limb perfusion requires isolating the main artery and vein of the perfused limb from the systemic circulation. The anatomic approach is determined by tumor site: external iliac vessels are used for thigh tumors, femoral or popliteal vessels for calf tumors, and axillary vessels for upper extremity tumors. The vessels are dissected, and all collateral vessels are ligated. The main artery and vein are then cannulated and connected to a pump oxygenator similar to that used in cardiopulmonary bypass. Either a tourniquet or an Esmarch band is applied to the limb to achieve complete vascular isolation. Chemotherapeutic agents are then added to the perfusion circuit and circulated for 90 minutes. Systemic leakage from the perfused limb is monitored continuously with Tc-radiolabeled human serum albumin injected into the perfusate, and radioactivity above the precordial area is recorded with a Geiger counter. During the entire procedure, hyperthermia of the perfused limb is maintained by external heating and by warming the perfusate to 40°C. At the end of the procedure, the limb is washed out, the cannulas are extracted, and the blood vessels are repaired.

Despite the 40-year history of using isolated limb perfusion to treat extremity sarcomas, many questions about this technique remain to be answered. The optimal chemotherapeutic agent in the perfusion circuit, the benefits of hyperthermia, and the effectiveness of hyperthermic perfusion as neoadjuvant or adjuvant treatment remain to be elucidated. Studies published to date have involved heterogeneous patient groups and various chemotherapeutic agents. Despite these limitations, response rates from 18% to 80% and overall 5-year survival rates from 50% to 70% have been reported.⁹⁷,⁹¹ However, survival outcomes following isolated limb perfusion have not yet been directly compared with survival outcomes after more conventional treatment approaches.

In the initial report of isolated regional perfusion for extremity sarcomas, published in 1974, McBride reported results in 79 patients with extremity sarcomas who had been treated with isolated limb perfusion during the previous 14 years.⁹⁷ All patients received melphalan and dactinomycin. The overall 5-year survival rate was 57%, and only 13 patients had subsequent amputation for recurrent disease. Over the next 20 years, isolated perfusion for treatment of extremity sarcoma fell out of favor for several reasons. Most notably, improved survival and decreased local recurrence rates could be obtained with less radical therapy, including conservative surgical excision combined with radiation to allow limb sparing in patients who were previously thought to require amputation.

A 1992 report by Lienard and colleagues¹⁰¹ renewed interest in isolated limb perfusion for extremity tumors. Those investigators reported a 100% response rate among patients with extremity melanomas and sarcomas treated with high-dose recombinant TNF-α plus interferon-γ and melphalan in an isolated perfusion circuit. This report led to larger studies geared specifically to patients with sarcoma. The largest of these studies, the European Multicenter Study, was reported by Eggermont and colleagues in 1996.⁹⁹ In that study of 186 patients, the overall tumor response rate was 82%, and the clinical and pathologic complete response rate was 29%. Although all of the study participants were reported to initially be candidates for amputation, the rate of limb salvage following isolated limb perfusion was 82%.⁹⁹ Subsequent studies have shown high local response and limb salvage rates and acceptable local and systemic toxic effects.¹⁰²

However, results in the United States have been inferior to those reported in Europe. In a study by Fraker and colleagues, the complete response rate was 26%, and an additional 30% of patients had a partial response. Fourteen patients (32%) underwent amputation for progressive tumors, while the remaining 30 patients (68%) were able to undergo limb-sparing surgery after isolated limb perfusion.¹⁰⁰ The inferior results in the U.S.-based studies are thought to be due to patient selection biases and the degree of treatment before limb perfusion.

While isolated limb perfusion for extremity sarcoma has fallen out of favor, recent reports of isolated limb infusion from the H. Lee Moffitt Cancer Center have shown promising
Patients with extremity sarcoma and who are considered for amputation are offered the option of isolated limb infusion with high-dose melphalan and actinomycin-D on protocol. Isolated limb infusion is a less invasive technique that can be repeated. Percutaneous cannulas are placed prior to infusion, the extremity is isolated similar to limb perfusion with an Esmarch band or tourniquet, and the perfusion is normothermic and acidic. Results from a multicenter retrospective study have demonstrated an overall response rate of 58%, and after a median follow-up of 21 months, there was an overall limb salvage rate of 78%. The benefits over limb perfusion remain the ability to repeat the technique in patients with disease response and less operative morbidity and risk of vascular injury from open surgery and cannulation of the iliac vessels. Although to date the technique has been well established for patients with locally advanced extremity disease for melanoma, its application for advanced, locally recurrent extremity sarcoma deserves further study.

**Radiation Therapy**

Radiation therapy is part of the standard treatment for high-grade extremity and trunk wall soft tissue sarcomas either in the pre- or postoperative setting. Patients with low-grade tumors or small, superficial high-grade tumors that have been resected with adequate margins may safely avoid radiation therapy.

The evidence supporting adjuvant radiation therapy for patients eligible for conservative surgical resection comes from two randomized trials104,105 and three large single-institution reports.106-108 In a randomized trial by the National Cancer Institute, 91 patients with high-grade extremity tumors were treated with limb-sparing surgery followed by chemotherapy alone or radiation therapy plus chemotherapy. The 10-year local control rate was 98% for patients receiving radiation therapy compared with 70% for those not receiving radiation therapy ($P = .0001$).104 Similarly, in a randomized trial from Memorial Sloan-Kettering Cancer Center, 164 patients underwent conservative surgery followed by observation or brachytherapy. For patients with high-grade tumors, the 5-year local control rate was 66% in the observation group and 89% in the brachytherapy group ($P = .003$).105 For patients with low-grade tumors, no significant difference was observed between treatment groups.109

Until recently, the standard treatment guidelines required radiation therapy after surgery for all patients with intermediate- or high-grade tumors of any size. However, small tumors (≤5 cm) have not generally been associated with local recurrence, and radiation therapy for such tumors may not be necessary.105 In a series of 174 patients reported by Geer and colleagues, postoperative radiation therapy did not improve 5-year local recurrence or overall survival rates for patients with small soft tissue sarcomas.109 Karakousis and colleagues reported a 5-year local recurrence rate of 6% for 80 patients with extremity sarcomas treated with wide local excision and observation, a rate similar to that for the 64 patients who underwent resection with narrower surgical margins and postoperative radiation therapy.110

The optimal mode of radiation therapy (external-beam radiation therapy, brachytherapy, or intensity-modulated radiation therapy [IMRT]) and timing of radiation therapy (preoperative, intraoperative, or postoperative) have yet to be defined. External-beam radiation therapy can be delivered using photons or particle beams (electrons, protons, pions, or neutrons). Conventional fractionation is usually 1.8 to 2 Gy per day. CT is an integral part of radiation therapy, used to define the gross tumor volume and to estimate the margin of tissue at risk for microscopic tumor involvement. The optimal radiation margin is not well defined: a margin of 5 to 7 cm is standard, but some centers advocate wider margins for tumors larger than 15 cm. At most institutions, the typical preoperative dose is 50 Gy given in 25 fractions, and resection is performed 4 to 8 weeks after completion of radiation therapy to allow acute radiation changes to subside. Postoperative radiation therapy planning is based on tumor site, tumor grade, surgical margins, and institutional preferences. The entire surgical scar and drain sites should be included in the field so that a near-full dose can be administered to the superficial skin. Metallic clips placed in the tumor bed during surgery can help define the limits of the resection and aid in radiation therapy planning. Doses of 60 to 70 Gy are usually necessary for postoperative treatment.

No consensus exists on the optimal sequence of radiation therapy and surgery. The available data come largely from single-institution, nonrandomized studies. Proponents of preoperative radiation therapy note that multidisciplinary planning with radiation oncologists, medical oncologists, and surgeons is easier early in the course of therapy. In addition, for some radiosensitive histologic subtypes, such as myxoid liposarcoma, preoperative radiation therapy may shrink the tumor, facilitating resection with negative margins. Furthermore, a tissue bed undisturbed by resection has better tissue oxygenation and can be successfully treated with lower doses of radiation. In addition, Nielsen and colleagues112 demonstrated that preoperative radiation fields are smaller than postoperative radiation fields and that the average number of joints included in the field is lower with preoperative than postoperative radiation therapy, which may result in improved functional outcome. Critics of preoperative radiation therapy cite the difficulty of pathologic assessment of margins and the increased rate of postoperative wound complications.113 However, reconstructive surgical techniques with advanced tissue transfer procedures are being used more often in these high-risk wounds and reportedly result in better outcomes. The higher doses generally required for postoperative radiation therapy have also been shown to be associated with greater long-term functional impairment.

The only randomized comparison of preoperative and postoperative radiation therapy to date was performed by the National Cancer Institute of Canada Clinical Trials--Canadian Sarcoma Group.114 This trial was designed to examine complications and functional outcome. The 190 patients enrolled from October 1994 to December 1997 were randomized to preoperative radiation therapy (50 Gy) or postoperative radiation therapy (66 Gy). With a median follow-up time of 3.3 years, the recurrence and progression-free survival rates were similar in the two groups. The incidence of wound complications was higher in the preoperative group versus the postoperative group (35% vs. 17%), and the incidence of wound complications was significantly higher for tumors of the lower extremity (43%) than for those of the upper extremity (5%).114 Late radiation toxic effects (e.g., fibrosis, joint stiffness, and edema) were more common with postoperative than preoperative radiation therapy (48% vs. 32%) because of higher postoperative radiation doses and larger treatment field sizes.115

Brachytherapy involves the placement of multiple radioactive seeds through catheters inserted in the tumor resection bed. The primary benefit of brachytherapy is the shorter overall treatment time of 4 to 6 days, compared to the 4 to 6 weeks generally required for preoperative or postoperative...
radiation therapy regimens. A cost-analysis comparison of adjuvant brachytherapy versus adjuvant external-beam irradiation for soft tissue sarcomas showed that costs were lower with brachytherapy.\textsuperscript{116} The implications of cost have been studied after radical resections and immediate adjuvant brachytherapy with either staged or immediate reconstruction noting a lower cost with staged reconstruction and tissue transfer techniques.\textsuperscript{117} Brachytherapy can also be used for recurrent disease previously treated with external-beam radiation. Guidelines established at Memorial Sloan-Kettering Cancer Center recommend spacing the afterloading catheters in 1-cm increments while leaving a 2-cm margin around the surgical bed.\textsuperscript{105} After adequate wound healing is confirmed, usually after the fifth postoperative day, the catheters are loaded with seeds containing iridium-192 that deliver 42 to 45 Gy of radiation to the tumor bed over 4 to 6 days. Subsequent studies at the H. Lee Moffitt Cancer Center that sought to determine outcomes between staged immediate and delayed reconstruction noted advantages in improved local control, wound healing, and less radiation-associated toxicity with staged reconstruction.\textsuperscript{118} The primary disadvantage of brachytherapy is that it requires significant expertise, extended inpatient hospital stays, and bed rest.

IMRT delivers radiation more precisely to the tumor than external-beam irradiation while minimizing the volume of surrounding tissues exposed to high radiation doses. The proposed benefits of preoperative IMRT include reduced risk of postoperative wound infections because of minimization of the dose to the skin\textsuperscript{119} and protection of underlying bone (e.g., femur) as a result of concave dose distributions.\textsuperscript{120} There have been no prospective randomized trials comparing the long-term outcomes following IMRT versus other types of radiation therapy. In a retrospective analysis of IMRT, patients with negative and positive/close (within 1 mm) margins were found to have 5-year local control rates of 94%.\textsuperscript{121} In addition, the rates of posttreatment edema and joint stiffness with IMRT were lower than the expected rates with conventional radiation therapy.

Local toxic effects of radiation therapy vary according to radiation dose, field size, and timing (preoperative or postoperative). With preoperative radiation therapy, the most frequent wound complications are wound dehiscence, wound necrosis, persistent drainage, infection, seroma formation, ulceration, and cellulitis.\textsuperscript{114} Postoperative irradiation of free flaps is often associated with wound complications, and patients should be advised that secondary surgical repair may be necessary; therefore, consideration for preoperative radiation rather than postoperative radiation for larger tumors requiring flap reconstruction is a logical and sound approach. Wound complication rates of 13% to 37% have been reported for preoperative radiation therapy, compared to 5% to 20% for postoperative radiation therapy.\textsuperscript{121} If catheters are loaded after the fifth postoperative day and/or if staged reconstruction is used, rates of wound complications after brachytherapy are improved if not similar to those after postoperative radiation therapy.\textsuperscript{118}

Long-term (chronic) effects of radiation therapy (those occurring >1 year after completion of therapy) are generally related to fibrosis/contractures, lymphedema, neurologic injury, osteitis, and fractures, all of which can cause substantial functional impairment.\textsuperscript{122} Variables associated with poorer functional outcome after radiation therapy include larger tumors, higher doses of radiation (>63 Gy), longer radiation fields (>35 cm), poor radiation technique, neural sacrifice, postoperative fractures, and wound complications.\textsuperscript{115,121} Additionally, complications of any kind are less likely after treatment for upper extremity sarcoma than after treatment for lower extremity sarcoma.\textsuperscript{113,114}

Definitive radiation therapy that delivers maximal-tissue-tolerance doses of radiation may be appropriate for selected patients with unresectable soft tissue sarcomas. In a study of 112 patients with unresectable soft tissue sarcomas, tumor size and radiation dose were found to influence local control and survival.\textsuperscript{124} The local control rate was 51% for tumors smaller than 5 cm and 9% for tumors larger than 10 cm, and patients who received at least 64 Gy had better local control and survival.

**Systemic Therapy**

Despite improvements in local control rates, metastasis and death remain significant problems for patients with high-risk soft tissue sarcomas. Patients considered at high risk of death from sarcoma include those presenting with metastatic disease, localized sarcomas at nonextremity sites, or sarcomas of intermediate- or high-grade histology larger than 5 cm.\textsuperscript{59,105}

**Standard Chemotherapy.** For most patients with sarcoma, results of conventional chemotherapy regimens have been poor. The chemosensitivity of soft tissue sarcoma varies by histologic subtype.\textsuperscript{30} Synovial sarcoma, myxoid/round cell liposarcoma, and uterine leiomyosarcoma are sensitive to chemotherapy,\textsuperscript{125} whereas pleomorphic liposarcoma, myxofibrosarcoma, epithelioid sarcoma, leiomyosarcoma, MPNSTs, angiosarcoma, and desmoplastic round cell tumors have intermediate sensitivity to chemotherapy. Relatively chemoresistant histologic subtypes include clear cell sarcoma, endometrial stromal sarcoma, alveolar soft part sarcoma, and extraskeletal myxoid chondrosarcoma. Considering the variability of responses by histologic subtype, it is not surprising that clinical trials of standard chemotherapy, which often include heterogeneous populations with respect to tumor grade and histology, have demonstrated no overall survival benefit.

Doxorubicin and ifosfamide are the two most active agents against soft tissue sarcoma, with consistently reported response rates of 20% or greater and positive dose-response curves.\textsuperscript{126,127} The European guidelines recommend doxorubicin 75 mg/m\textsuperscript{2} every 3 weeks as first-line treatment for advanced disease.\textsuperscript{30} Treatment duration is based on response, but a maximum of six cycles is generally recommended because of the risk of cumulative cardiotoxicity. Ifosfamide is the recommended second-line treatment and is recommended for first-line treatment in patients with cardiac morbidity. The standard dose of ifosfamide is 9 to 10 g/m\textsuperscript{2}; however, single-institution series using higher-dose regimens (>10 g/m\textsuperscript{2}) or standard-dose ifosfamide combined with doxorubicin have shown response rates of 20% to 60%.\textsuperscript{127} Synovial sarcomas have been shown to be particularly sensitive to ifosfamide. Ifosfamide-associated toxic effects include hemorrhagic cystitis, neurotoxicity, and renal tubular acidosis. Historically, combination therapy with doxorubicin plus ifosfamide, dacarbazine, or both has resulted in increased response rates but no improvement in overall survival.\textsuperscript{128} Dacarbazine as a single agent has also demonstrated activity in clinical trials.

Over the past decade, several additional chemotherapeutic agents, including gemcitabine, taxanes, and trabectedin, have been noted to be active against soft tissue sarcomas. Gemcitabine as a single agent was reported to produce responses in 18% of patients with advanced sarcoma.\textsuperscript{129} Gemcitabine combined with docetaxel has been reported to produce response rates as high as 53% in patients with uterine leiomyosarcoma.\textsuperscript{129,130} Gemcitabine
combined with vinorelbine has also been associated with clinical benefit in patients with advanced sarcomas.\textsuperscript{131} The taxanes (docetaxel and paclitaxel) have been found to be active against angiosarcomas, particularly of the face and scalp, likely because of their potent antiangiogenic effects.\textsuperscript{122,133}

**Novel Chemotherapeutic Agents.** Aldoxorubicin is a doxorubicin derivative that serves as a produg of doxorubicin that covalently binds to albumin in the blood until reaching the acidic tumor environment releasing doxorubicin into the tissue. A recent international, multicenter, phase 2b, open-label, randomized study enrolled 126 patients from 2012 to 2013. Single-agent aldoxorubicin therapy showed superior efficacy over doxorubicin by prolonging progression-free survival and improving rates of 6-month progression-free survival and tumor response, warranting further study.\textsuperscript{134}

Trabectedin, a marine-derived alkaloid that binds DNA, affecting transcription and inducing the formation of DNA double-strand breaks, has shown benefit in the treatment of advanced soft tissue sarcomas, particularly leiomyosarcoma, myxoid liposarcoma, and other translocation-related sarcomas.\textsuperscript{135} Trabectedin is generally well tolerated but can be associated with prolonged and severe neutropenia, thrombocytopenia, and hepatic toxic effects.

Palifosfamide is a stabilized formulation of the active metabolite of ifosfamide that has been reported to be better tolerated than ifosfamide.\textsuperscript{136} Early trials have suggested anti-tumor activity comparable or superior to that of ifosfamide without nephrotoxicity; however, recent negative results of the PICASSO III Trial, a phase 3, placebo-controlled study of doxorubicin with or without palifosfamide in patients with metastatic soft tissue sarcoma, have neatly secured the fate of palifosfamide in the treatment of soft tissue sarcoma.\textsuperscript{137}

**Targeted Therapies.** Several targeted agents are being investigated for the treatment of soft tissue sarcomas. Among these are tyrosine kinase inhibitors (e.g., imatinib, sunitinib, sorafenib, and dasatinib) that have been developed and approved for treatment of GIST. Clinical data accumulated in phase 2 trials also support the use of tyrosine kinase inhibitors (e.g., imatinib, sorafenib, and sunitinib) in the management of other advanced sarcomas.\textsuperscript{138} Anti-vascular endothelial growth factor antibodies such as bevacizumab have demonstrated activity in patients with metastatic or unresectable angiosarcoma, solitary fibrous tumor, and epithelioid hemangioendothelioma.\textsuperscript{139} Pazopanib is an oral angiogenesis inhibitor that targets vascular endothelial growth factor receptors, platelet-derived growth factor receptor (PDGFR), and c-kit. In a recent phase 3 study, pazopanib showed efficacy against placebo in second or further line of therapy in patients with advanced soft tissue sarcoma.\textsuperscript{139} Inhibitors of the mammalian target of rapamycin pathway, including temsirolimus, everolimus, and ridaforolimus, have also shown activity against some soft tissue sarcomas (i.e., PEComas).\textsuperscript{140}

Olaratumab is a human antiplatelet-derived growth factor receptor α monoclonal antibody that has antitumor activity in human sarcoma xenografts. Recently, a phase 1b and randomized phase 2 study of olaratumab and doxorubicin versus doxorubicin alone has demonstrated improvements in both objective response rates (18.2% vs. 11.9%, $P = 0.34$) and median overall survival (26.5 vs. 14.7 months [stratified hazard ratio 0.46, 0.30–0.71, $P = 0.0003$]). Additional studies are warranted for this promising combination of agents.\textsuperscript{141}

**Benefits of Systemic Therapy.** The use of adjuvant and neoadjuvant chemotherapy for soft tissue sarcomas remains controversial. More than a dozen individual randomized trials of adjuvant chemotherapy have failed to demonstrate improvement in disease-free or overall survival for patients with soft tissue sarcoma. However, several limitations of these individual trials may explain the lack of observed improvement. First, the chemotherapy regimens used were suboptimal, consisting of single-agent therapy (most commonly with doxorubicin) and insufficiently intensive dosing schedules. Second, the patient groups were not large enough to reveal clinically significant differences in survival rates. Finally, most studies included patients at low risk of metastasis and death, namely those with small (≤10 cm) and low-grade tumors.

The Sarcoma Meta-Analysis Collaboration analyzed 1568 patients from 14 trials of doxorubicin-based adjuvant chemotherapy to evaluate the effect of adjuvant chemotherapy on localized, resectable soft tissue sarcomas.\textsuperscript{142} At a median follow-up time of 9.4 years, doxorubicin-based chemotherapy significantly improved the time to local and distant recurrence and recurrence-free survival rates. However, the absolute benefit in overall survival was only 4%, which was not significant ($P = .12$). In a subset analysis, patients with extremity tumors had a 7% benefit in terms of overall survival ($P = .029$).\textsuperscript{142}

After this meta-analysis, randomized controlled trials of more contemporary anthracycline/ifosfamide dosing combinations with relatively small numbers of patients have yielded conflicting results. In an Italian cooperative trial, adjuvant chemotherapy improved median disease-free and overall survival times in patients with high-risk extremity soft tissue sarcomas.\textsuperscript{143} In that study, 104 patients with high-grade tumors 5 cm or larger were randomized to definitive surgery or surgery plus adjuvant chemotherapy consisting of epirubicin (60 mg/m$^2$ per day on days 1 and 2) and ifosfamide (1.8 g/m$^2$ per day on days 1 through 5) for five cycles. With a median follow-up time of almost 5 years, disease-free survival times were 16 months in the surgery-alone group and 48 months in the combined-treatment group ($P = .04$), and median overall survival times were 46 months in the surgery-alone group and 75 months in the combined-treatment group ($P = .03$).\textsuperscript{143} However, several years later, the surgery-alone and combined-treatment groups had equivalent relapse rates and deaths, which resulted in statistically similar overall survival.\textsuperscript{144}

In an effort to further assess the role of chemotherapy in patients with stage III extremity sarcoma, a cohort analysis of the combined databases of The University of Texas MD Anderson Cancer Center and Memorial Sloan-Kettering Cancer Center was performed. Data on 674 patients with stage III extremity sarcoma who received either preoperative or postoperative doxorubicin-based chemotherapy were reviewed. The 5-year disease-specific survival rate was 61%.\textsuperscript{145} Cox regression analysis showed a time-varying effect of chemotherapy with an associated benefit during the first year while receiving chemotherapy. However, the clinical benefits of chemotherapy in patients with stage III sarcomas were not sustained beyond 1 year. Grobmyer and colleagues compared the outcomes of patients treated at two institutions (1990–2001) with surgery only or surgery plus preoperative chemotherapy containing doxorubicin and ifosfamide. In this analysis, chemotherapy was associated with an improvement in the 3-year disease-specific survival rate that was most pronounced in patients with tumors larger than 10 cm (62% for surgery alone vs. 83% for neoadjuvant chemotherapy and surgery).\textsuperscript{146}
More recently, the European Organization for Research and Treatment of Cancer (EORTC) completed a phase 3 randomized study (trial EORTC-62931; conducted from 1995 through 2003) comparing surgery alone versus surgery plus adjuvant ifosfamide (5 g/m²) plus doxorubicin (75 mg/m²) with growth factor support (lenograstim) every 21 days for five cycles in 351 patients with resected grade II or III soft tissue sarcoma at any site. The estimated relapse-free survival rate was 52% in both arms, and the overall survival rate was better in the control arm (69% vs. 64%). Although most individual studies are underpowered, data from all of these studies suggest that chemotherapy regimens that incorporate ifosfamide may provide some disease-free survival benefit but do not improve long-term overall survival for the majority of patients with soft tissue sarcoma.

In 2008, two updates to the 1997 Sarcoma Meta-Analysis Collaboration were published. O’Connor and colleagues included all of the trials in the original meta-analysis and added data from four additional trials, for a total of 18 trials with 2170 patients. The results showed a benefit of chemotherapy in terms of disease-free survival at 5 years and recurrence-free survival at 10 years but again failed to demonstrate a benefit in terms of long-term overall survival. The second update, by Pervaiz and colleagues, which did not include the EORTC-62931 trial, showed that adjuvant chemotherapy was associated with a significant decrease in the risk of death (hazard ratio, 0.77; P = .01).

Because the evidence regarding adjuvant systemic therapy for stage III soft tissue sarcoma is inconclusive, considerable variation still exists in treatment recommendations even though patients with large, stage II or stage III soft tissue sarcomas are at high risk for recurrence and metastasis. Chemotherapy may be considered to downstage large tumors to enable limb-sparing at high risk for recurrence and metastasis. Chemotherapy may be spared the toxic effects of prolonged adjuvant chemotherapy. The theoretical advantages notwithstanding, concurrent chemoradiation therapy decreases the total treatment time for patients with high-risk sarcoma. This decrease represents a substantial advantage over current sequential combined-modality treatment approaches, for which the total duration of radiation therapy, chemotherapy, surgery, and rehabilitation frequently exceeds 6 to 9 months.

**Concurrent Chemoradiation Therapy**

Treatment approaches that combine systemic chemotherapy with radiosensitizers and concurrent external-beam radiation therapy may improve disease-free survival by treating microscopic disease and enhancing the treatment of macroscopic disease. Concurrent chemoradiation therapy with doxorubicin-based regimens reportedly produces favorable local control rates for patients with sarcoma. Since those findings were published, several groups have evaluated routes of administration, alternative chemotherapeutic agents, and the toxicity of combined therapies.

The National Comprehensive Cancer Network (NCCN) recommends a history and physical and chest CT or radiography every 3 to 6 months for 2 to 3 years after completion of treatment. Because most cases of distant metastasis occur within 2 to 3 years of initial diagnosis, the NCCN guidelines indicate that follow-up intervals can be lengthened to every 6 months, and imaging can be done annually during years 2 through 5. Consideration should also be given to imaging the primary tumor site; most experts recommend that the tumor site be evaluated every 6 months with MRI for extremity tumors or CT for intra-abdominal or retroperitoneal tumors. Guidelines have been established for using MRI to distinguish recurrences from typical postsurgical changes: a discrete nodule with low signal intensity on T1-weighted images and higher signal intensity on T2-weighted images that enhances after administration of intravenous contrast material is strongly suggestive of recurrence and should be biopsied. Ultrasonography may be an alternative to MRI or CT for assessing for recurrence in the extremities.
Recurrence is common after surgery for abdominal soft tissue sarcomas. CT is useful for detecting recurrences at primary and distant anatomic sites in the abdomen and pelvis. After surgery, CT every 3 to 6 months during the first 2 years and every 6 months for 3 years thereafter has been recommended. However, today many experienced surgeons are advocating less aggressive imaging for asymptomatic patients, particularly after a second recurrence of retroperitoneal sarcoma, arguing that there is insufficient evidence to suggest that survival is improved by earlier detection.

Whooley and colleagues reviewed the efficacy of the surveillance strategy used at Roswell Park Cancer Institute for 174 patients with soft tissue sarcomas of the extremities.¹⁶⁰ Patients were evaluated every 3 months for the first 2 years, every 4 to 6 months during year 3, and every 6 months during years 4 and 5. Local recurrence occurred in 18% of patients at a median time after completion of treatment of 14 months, and all but one of the recurrences were detected with physical examination alone. Fifty-seven patients had distant recurrences (at a median of 18 months after treatment), of which 36 were asymptomatic and diagnosed by surveillance imaging. The investigators determined that the positive predictive value of chest radiography during follow-up was 92%.¹⁶⁰ However, evaluation of the primary tumor site by CT or MRI was ineffective in detecting recurrences. The authors recommended that patient characteristics, location of the primary tumor, previous treatment, and physician familiarity with changes after surgery and radiation therapy should all be considered in determining the need for radiographic imaging.

Management of Recurrent Sarcoma

Up to 20% of patients with extremity sarcoma develop locally recurrent disease, which is often accompanied by distant metastases; thus, all patients with recurrent extremity sarcoma should undergo a full staging assessment. Patients with microscopically positive surgical margins are at increased risk of local recurrence. In a series of 179 patients with locally recurrent extremity soft tissue sarcoma at Memorial Sloan-Kettering Cancer Center, the median interval to local recurrence was 16 months; 65% of patients developed a local recurrence by 2 years, and 90% by 4 years.¹⁵⁹ The majority of patients (89%) were treated with additional limb-sparing surgery, and 73% received additional adjuvant therapy; the disease-specific survival after treatment of first local recurrence was 55% at 4 years. Independent prognostic factors for disease-specific survival after local recurrence included tumor grade, local recurrence size, and local recurrence-free interval. These data indicate that an isolated local recurrence should be treated aggressively with resection with negative margins.

For patients with extremity sarcomas, achieving negative margins on resection of recurrent disease frequently requires amputation. However, in some patients with recurrent extremity sarcoma, function-preserving resection combined with additional radiation therapy, with or without chemotherapy, can produce acceptable rates of local control.¹⁶¹-¹⁶³ Nori and colleagues reported a local control rate of 69% among 40 patients with recurrent tumors treated with reexcision and brachytherapy to a median dose of 45 Gy.¹⁶³ In a similar series, Midis and colleagues reported that limb-sparing surgery was possible in 66% of patients, and the 5-year local recurrence-free survival rate was 72% in those patients.¹⁶⁰

The primary determinant of survival in patients with soft tissue sarcoma is the development of distant metastases.

Patients with extremity sarcomas generally develop pulmonary metastases.¹⁶⁰ Less common sites of metastasis for soft tissue sarcomas include bone (7%), liver (4%),⁴⁰ and lymph nodes (5–7%).²⁸ Myxoid liposarcoma of the extremity is known to metastasize to the abdomen and pelvis; therefore, staging CT of these regions must be performed before definitive local therapy is administered.³²

Management of Recurrent and Distant Metastatic Sarcoma.

In selected individuals with distant metastatic disease, surgical resection of a primary soft tissue sarcoma may be appropriate as a palliative procedure. The decision should be based on the patient’s symptoms, which often include pain; ability to achieve local tumor control; comorbidities; anticipated morbidity of the surgical procedure; and the extent of metastases.

The most common initial site of distant metastasis of soft tissue sarcomas is the lung. Selected patients with a limited number of pulmonary nodules (less than four nodules), long disease-free intervals, and no endobronchial invasion may become long-term survivors after pulmonary resection (Fig. 36-8); 15% to 40% of patients with complete resection of metastatic disease confined to the lung are long-term survivors.¹⁵⁹,¹⁶⁰,¹⁶⁴ In a retrospective multi-institutional study of 255 patients with lung metastases, the 5-year overall survival rate after metastasectomy was 38%.¹⁵⁷ Favorable prognostic factors in that study included microscopically tumor-free margins, age younger than 40 years, and grade 1 or 2 tumor.¹⁵⁷ For patients who are surgical candidates, pulmonary resection alone can be more cost-effective than watchful waiting, chemotherapy, or chemotherapy plus surgery.¹⁵⁸

Chemotherapy for Distant Metastatic Sarcoma.

Doxorubicin, either alone or combined with other agents, has been the primary treatment modality for patients with advanced or distant metastatic sarcomas for several decades.¹²⁸ Although most patients with metastatic disease are not curable, some
patients with limited disease experience stabilization of disease with multidisciplinary treatment, which often includes surgery and radiation therapy in addition to chemotherapy. Several factors predict better outcome for patients with recurrent metastatic sarcoma undergoing chemotherapy, including good performance status, previous response to chemotherapy, younger age, absence of hepatic metastases, low-grade tumor, and long disease-free interval.\textsuperscript{163} Isolated liver metastases, if stable over several months, may be amenable to resection,\textsuperscript{166} radiofrequency ablation,\textsuperscript{167} or chemoembolization.\textsuperscript{168}

As data accumulate regarding the sensitivity of sarcoma subtypes to particular chemotherapies, it is critical that histology-driven treatment approaches be used. New therapies are also being identified based on the unique molecular signatures of sarcomas.\textsuperscript{128}

**Palliative Radiation Therapy.** Definitive radiation therapy can be considered when no acceptable surgical option is available (e.g., in patients with significant medical comorbidities). In this setting, radiation doses greater than 63 Gy yielded superior tumor control, but doses greater than 68 Gy resulted in increased rates of major complications.\textsuperscript{169}

**SPECIAL CLINICAL SITUATIONS**

**Myxoid Liposarcoma**

Myxoid liposarcomas belong to the group of soft tissue sarcomas with lipomatous differentiation. However, myxoid liposarcomas differ from the other liposarcoma subtypes with respect to morphology (i.e., myxoid stroma and lipomatous differentiation) and clinical behavior. Myxoid liposarcomas frequently present as slow-growing, deep tumors in the lower extremity and can metastasize to other soft tissue locations, including the retroperitoneum and extremities.\textsuperscript{170,171} For this reason, CT of the chest, abdomen, and pelvis is recommended for adequate staging and surveillance of myxoid liposarcoma.

**Retroperitoneal Sarcoma**

Most retroperitoneal tumors are malignant, and about one-third are soft tissue sarcomas. Also to be considered in the differential diagnosis of a retroperitoneal tumor are primary germ cell tumors, lymphoma, and metastatic testicular cancer. Approximately 1000 new cases of retroperitoneal sarcoma are diagnosed annually in the United States, and these tumors account for 10% to 15% of all adult tissue sarcomas. Approximately two-thirds of retroperitoneal sarcomas are high grade (either grade 2 or 3), and liposarcoma and leiomyosarcoma are the most common histologies.

Retroperitoneal sarcomas generally present as large masses: 70% are larger than 10 cm at diagnosis.\textsuperscript{172} They typically do not produce symptoms until they grow large enough to compress or invade contiguous structures, although pain, early satiety, and obstructive gastrointestinal symptoms may occur early in the disease course in some patients. Evaluation of a patient with a retroperitoneal mass begins with an accurate history that should exclude signs and symptoms associated with lymphoma (e.g., fever and night sweats). A complete physical examination, with particular attention to all nodal basins and with a testicular examination in men, is critically important. Laboratory assessment can be helpful; elevated lactate dehydrogenase levels may suggest lymphoma, and elevated β-human chorionic gonadotropin levels or α-fetoprotein levels may indicate a germ cell tumor.

Although the general principles of evaluation and management for retroperitoneal sarcomas are similar to those for extremity sarcomas, there are some differences. Contrast-enhanced CT of the abdomen and pelvis is used to define the extent of the tumor and its relationship to surrounding structures, particularly vascular structures, for surgical planning; contrast-enhanced CT can also often distinguish between well-differentiated and dedifferentiated liposarcoma. CT imaging is also done to evaluate the liver for the evidence of metastases, the peritoneal cavity for evidence of discontiguous disease, and the kidneys for assessment of function. Angiography or magnetic resonance arteriography/venography can also be used to delineate vascular anatomy when involvement of critical vascular structures is suspected. Thoracic CT should be performed to evaluate for potential lung metastases because 11% of patients with retroperitoneal sarcoma present with synchronous metastatic disease. CT-guided core needle biopsy is appropriate to provide a tissue diagnosis; however, well-differentiated liposarcoma may be diagnosed with CT imaging alone, and negative biopsy findings should not delay operative intervention.

Complete surgical resection is the most effective treatment for primary or recurrent retroperitoneal sarcoma (Fig. 36-9). En bloc resection often necessitates sacrificing contiguous structures such as the colon, kidney, spleen, pancreas, psoas muscle, small bowel, inferior vena cava, and aorta.\textsuperscript{173} In a review of 25 patients who underwent resection of retroperitoneal sarcoma with major blood vessel involvement in a 16-year time span, postoperative morbidity and mortality rates were 36% and 4%, respectively. Vessel patency rates were greater than 88% with a median follow-up time of 19.3 months.\textsuperscript{173,174} Local control and survival rates were favorable in patients with tumor-free resection margins. The authors concluded that vascular resection is the treatment of choice in sarcomas that involve major blood vessels in the retroperitoneum.\textsuperscript{83} Similar considerations were made by other groups reporting specifically on inferior vena cava resection in the context of multivisceral resection for retroperitoneal sarcoma and on surgical morbidity after extended surgical resection of retroperitoneal sarcoma. Extended procedures, including also vessels, are feasible and safe if carried out in experienced centers. While the goal of sarcoma resection is

![Figure 36-9. A 50-year-old man with a large right dedifferentiated liposarcoma. Note the atypical fat surrounding the right kidney and displacing the viscera to the left hemiabdomen and the large dedifferentiated mineralized solid nodule lateral to the right kidney.](image-url)
wide excision, this is unlikely to be achievable in most patients with retroperitoneal sarcomas. Surgery is considered marginal in most cases, even when macroscopically complete, but every attempt should be made to minimize this marginality by liberally resecting surrounding organs when involved. The extension of surgery should then take into consideration a trade-off between expected morbidity and benefit and should be best carried out at high-volume centers, where technical skills and knowledge of the natural history of this very rare disease can be found.

In an analysis of 500 patients with retroperitoneal soft tissue sarcoma treated at Memorial Sloan-Kettering Cancer Center, the median survival time was 103 months for those who underwent complete resection versus 18 months for those who underwent incomplete resection or observation without resection.172 In general, surgical resection should not be offered unless radiographic evidence indicates the potential for complete resection; however, palliative surgical resection may be considered to reduce symptoms of intestinal obstruction, pain, or bleeding.173 In particular, in patients with atypical lipomatous tumors, an aggressive surgical approach including incomplete resection or debulking is justified to palliate symptoms and may provide a potential survival benefit.174 Such an approach is not justified for dedifferentiated liposarcomas or other high-grade retroperitoneal sarcomas because these tumors have high rates of distant metastasis and local recurrence.

Adjuvant Therapy. Most studies have failed to show a survival benefit from adjuvant chemotherapy for retroperitoneal sarcoma.177-179 Because of the high rates of local recurrence, radiation therapy has been proposed for treating microscopic residual disease as an adjunct to surgical resection. However, the optimal technique and timing of radiation therapy have not been established, and the potential benefits of radiation therapy must be weighed against the increased risk of treatment-related toxic effects.

Radiation treatment of retroperitoneal sarcomas is complex because tumors are usually large, which necessitates large treatment fields close to radiosensitive structures (e.g., bowel). Several techniques have been used, including preoperative and postoperative external-beam radiation therapy, intraoperative radiation therapy, and brachytherapy.180 Preoperative radiation therapy is feasible and well tolerated. Toxic effects may be less severe with preoperative radiation therapy given that the tumor borders are definable, the tumor displaces radiosensitive viscera away from the treatment field, and effective doses of radiation may be lower preoperatively.181

Several studies have shown favorable local control rates for intermediate- and high-grade retroperitoneal sarcoma treated with preoperative radiation therapy and complete resection.182 However, most studies have failed to show a survival benefit.183 This situation prompted the initiation of a multicenter, randomized trial sponsored by the American College of Surgeons Oncology Group (ACOSOG) comparing surgery to surgery with preoperative radiation (ACOSOG Z9031). Unfortunately, the study was closed prematurely in 2006 because of low patient accrual. A similar study is now ongoing in Europe, sponsored by the Soft Tissue and Bone Sarcoma Group (STBSG) of the EORTC.

Current recommendations for radiation therapy for patients with retroperitoneal sarcoma at high volumes centers are based on disease characteristics at presentation.184 For high-risk patients, defined as those with large, high-grade tumors or recurrent low-grade tumors, preoperative radiation therapy to a total dose of 50 Gy followed by surgical resection is considered. Postoperative radiation is discouraged unless the resected tumor bed is clearly away from dose-limiting structures.

Treatment of Recurrence. Retroperitoneal sarcomas recur more often than extremity and trunk wall ones. Retroperitoneal leiomyosarcomas, in addition to recurring locally in the tumor bed and metastasizing to the lungs, frequently spread to the liver. Retroperitoneal sarcomas can also recur diffusely throughout the peritoneal cavity (sarcomatosis). Resection of recurrent retroperitoneal sarcoma is similar to resection of recurrent extremity sarcoma. However, the likelihood that a recurrent retroperitoneal sarcoma will be resectable declines precipitously with each recurrence. In a large series of patients treated at Memorial Sloan-Kettering Cancer Center, the authors were able to resect recurrent tumors in 57% of patients with a first recurrence but only 20% of patients with a second recurrence and 10% of patients with a third recurrence.185 In up to 25% of patients, well-differentiated retroperitoneal liposarcoma recurs in a poorly differentiated form or recurs with areas of dedifferentiation. Dedifferentiated retroperitoneal liposarcoma is more aggressive than its well-differentiated precursor and has a greater propensity for distant metastasis.

Gastrointestinal Sarcoma

Patients with gastrointestinal sarcoma most often present with nonspecific gastrointestinal symptoms that are determined by the site of the primary tumor. In a series from Memorial Sloan-Kettering Cancer Center, early satiety and dyspepsia were noted in patients with tumors of the upper gastrointestinal tract, whereas tenesmus and changes in bowel habits were common in patients with tumors of the lower gastrointestinal tract.184 In a series of 80 patients with various smooth-muscle tumors of the gastrointestinal tract, Chou and colleagues185 identified the most common presenting symptoms and signs as gastrointestinal bleeding (44%), abdominal mass (38%), and abdominal pain (21%).

Establishing the diagnosis of a gastrointestinal sarcoma preoperatively is often difficult. Radiologic assessment, including CT of the abdomen or pelvis, is sometimes useful to determine the anatomic location, size, and extent of disease. Patients with localized disease frequently present with a large intra-abdominal mass. However, there is no radiographic evidence of regional lymph node metastases, which would be typical of an adenocarcinoma of similar size and anatomic location. In patients with advanced gastrointestinal sarcoma, CT may demonstrate disseminated intra-abdominal masses with or without concomitant ascites and invasion of tissue planes.

Endoscopy (esophagogastroduodenoscopy or colonoscopy) has become the mainstay for evaluating symptoms related to the gastrointestinal tract. For tumors involving the stomach, upper endoscopy with endoscopic ultrasonography and biopsy are important diagnostic tests used to distinguish gastrointestinal sarcoma from adenocarcinoma of the stomach. Endoscopic biopsy of these tumors is preferred over CT-guided biopsy if feasible. This distinction is clinically significant because the extent of resection (local excision versus gastrectomy) and the role of regional lymphadenectomy differ for these two conditions. For gastrointestinal sarcomas, lymphatic spread is not the primary route of metastasis; consequently, lymphadenectomy is not routinely performed as part of resection. The general recommendation for gastrointestinal sarcoma, based on published data and the primary pattern of distant (vs. local) failure, is to
resect the tumor with a 2- to 4-cm margin of normal tissue. However, some cases may be technically challenging because of the tumor’s anatomic location or size. For example, for gastric tumors located near the gastroesophageal junction, achieving adequate surgical margins may not be possible without a total or proximal subtotal gastrectomy. This recommendation is much different when considering resection of gastrointestinal stromal tumors (GIST) where a gross margin negative resection is recommended are rarely is a total gastrectomy required. Similarly, large leiomyosarcomas arising from the stomach with invasion of adjacent organs should be resected together with the adjacent involved viscera en bloc.

For sarcomas of the small or large intestine, segmental bowel resection is the standard treatment. For sarcomas of the jejunum, ileum, and colon, the tumor is excised en bloc with the involved segment of intestine and its mesentery; radical mesenteric lymphadenectomy is not attempted. For sarcomas originating in the rectum, the tumor resection technique is based on the anatomic location and size of the tumor. For small, low rectal lesions, clear margins may be achievable with a transanal excision. Large or locally invasive lesions may require more extensive operations for complete tumor extirpation.\textsuperscript{186,187}

**Breast Sarcoma**

Sarcomas of the breast are rare tumors, accounting for less than 1% of all breast malignancies and less than 5% of all soft tissue sarcomas. A variety of histologic subtypes have been reported within the breast, including angiosarcoma, stromal sarcoma, fibrosarcoma, and malignant fibrous histiocytoma.

Angiosarcoma of the breast accounts for about 50\% of all sarcomas of the breast and has increasingly been associated with radiation therapy for treatment of primary breast cancer.\textsuperscript{10} The period between radiation therapy and diagnosis of radiation-associated breast sarcoma has been reported to range from 3 to 20 years, with an incidence of 0.3\% at 10 years and 0.5\% at 15 years.\textsuperscript{188} In a retrospective study of 55 patients with angiosarcoma of the breast, patients with radiation-associated angiosarcoma were on average 30 years older and were less likely to present with distant metastases than radiation-naive patients. Clinically, radiation-associated angiosarcoma of the breast may occur in the irradiated chest wall after mastectomy or in the irradiated breast following segmental resection. The findings at presentation of a patient with cutaneous angiosarcoma often include an expanding erythematous patch, red papular eruptions, bluish-black lesions, or bruise-like discoloration overlying an area of induration. Mammography is often nonspecific, and diagnosis requires punch or incisional biopsy.

Cystosarcoma phylloides are generally not considered to be sarcomas because these tumors are thought to originate from hormonally responsive stromal cells of the breast and are usually benign. In patients with these tumors, infiltrating tumor margins, severe stromal overgrowth, atypia, and cellularity have all been identified as risk factors for metastases.\textsuperscript{189}

As with sarcomas at other anatomic sites, histopathologic grade and tumor size are important prognostic factors for sarcomas of the breast. The likelihood of local recurrence increases as tumor size increases; tumors smaller than 5 cm are associated with better overall survival. Local and distant recurrences are more common in patients with high-grade lesions. Complete excision with negative margins is the primary therapy. Simple mastectomy confers no additional benefit if complete excision can be accomplished by segmental mastectomy. Because of low rates of regional lymphatic spread, axillary dissection is not routinely indicated. Neoadjuvant chemotherapy or radiation therapy may be considered for patients with large, high-risk tumors.

**Uterine Sarcoma**

Sarcomas account for less than 5\% of uterine malignancies. Uterine sarcomas have been classified into four histologic subgroups: uterine leiomyosarcoma, endometrial stromal sarcoma, malignant mixed Müllerian tumor (carcinosarcoma), and undifferentiated endometrial sarcoma. Five-year overall survival rates for patients with uterine sarcoma are 30\% to 50\%.\textsuperscript{190} Total abdominal hysterectomy (TAH) is recommended for localized disease. Bilateral salpingo-oophorectomy is mandatory only in endometrial stromal sarcoma. Because uterine sarcomas are rare, the benefits of adjuvant therapy (e.g., chemotherapy, hormonal therapy) have not been adequately evaluated. Pelvic postoperative irradiation has been studied instead in a randomized fashion. The results of such study have been reported, showing no benefit in survival in favor of radiation therapy.\textsuperscript{191}

Uterine leiomyosarcomas are smooth-muscle tumors and account for 35\% to 40\% of uterine sarcomas. Leiomyosarcoma can affect women in their twenties, although it is more commonly diagnosed between 50 and 60 years of age. Standard treatment is TAH with or without ovarian preservation depending on the patient’s wishes and menopausal status. Lymph node metastasis is present in less than 5\% of patients at diagnosis, and lymphadenectomy is not recommended. Adjuvant pelvic radiation therapy can be considered for selected high-risk patients. Adjuvant chemotherapy is controversial. Gemcitabine plus docetaxel has been noted to be well tolerated and highly active, with a response rate of 53\% in patients with unresectable uterine leiomyosarcoma.\textsuperscript{180} Doxorubicin and trabectedin have also demonstrated activity when used as first- or second-line therapy.

Endometrial stromal sarcomas account for approximately 7\% to 10\% of uterine sarcomas. Mitotic count is used to classify endometrial stromal sarcomas as low grade (<10 mitoses per 10 high-power fields) or high-grade (>10 mitoses per 10 high-power fields). In general, low-grade tumors demonstrate an indolent clinical course, while high-grade tumors are more aggressive with a poorer prognosis. Unlike other uterine sarcomas subtypes, endometrial stromal sarcomas express progesterone receptors and have been found to be responsive to hormonal manipulation as an adjuvant therapy or for treatment of recurrent disease.\textsuperscript{192,193} Surgical treatment for these tumors includes TAH and bilateral salpingo-oophorectomy in premenopausal women; postoperative hormone replacement therapy is contraindicated.\textsuperscript{194} Recurrent or advanced disease may respond to antiestrogen therapy. Tamoxifen is not recommended because it may be proestrogenic in this setting.

Malignant mixed müllerian tumor accounts for 50\% of uterine sarcomas and arises predominantly in postmenopausal women. This tumor is regarded as epithelial and is treated not with agents typically used to treat sarcoma but with agents used to treat ovarian and endometrial cancers.

Undifferentiated endometrial sarcoma is an aggressive malignancy that does not express estrogen or progesterone receptors. It is associated with a poor prognosis even in patients presenting with localized disease. TAH with or without preservation of the ovaries is recommended; postoperative pelvic radiation therapy may also be administered. Systemic agents for other soft tissue sarcomas are used for recurrent and/or metastatic disease.
GASTROINTESTINAL STROMAL TUMORS

GISTs, which account for the majority of gastrointestinal sarcomas, have distinctive molecular features that have been characterized over the last decade. These tumors share phenotypic similarities with the intestinal pacemaker cells known as the interstitial cells of Cajal; interstitial cells of Cajal and GIST cells express the hematopoietic progenitor cell marker CD34 and the growth factor receptor c-Kit. Expression of the c-Kit gene product, CD117, has emerged as an important defining feature of GISTs. Using these diagnostic criteria, the incidence of GIST has been estimated to be 6 to 15 cases per million individuals per year. Until recently, systemic treatment for patients with unresectable or metastatic GIST was of little benefit because these tumors were resistant to conventional chemotherapy. Since the recognition that KIT activation occurs in most GISTs, KIT inhibition has emerged as an adjunct to surgery in select patients with resectable disease and as a primary treatment modality for patients with stage IV disease.

Approximately 80% of GISTs have a mutation in the gene encoding the KIT receptor tyrosine kinase, and 5% to 10% have a mutation in the gene encoding the related PDGFRA receptor tyrosine kinase; such mutations result in the expression of mutant proteins with constitutive tyrosine kinase activity. The remaining GISTs do not have a detectable mutation, but lack of a mutation does not preclude a diagnosis of GIST if the tumor is morphologically typical of GIST. The presence and type of KIT (exon 11 or exon 9) or PDGFRA (exon 18) mutation has been found to predict tumor response to imatinib. In a phase 2 trial, patients with KIT exon 11 mutations had better response rates (83.5% vs. 47.8%) and survival than those with KIT exon 9 mutations or those without KIT or PDGFRA mutation. These findings have subsequently been confirmed in two additional phase 3 trials conducted by the EORTC–Italian Sarcoma Group–Australasian Gastrointestinal Trials Group (EORTC-62005). The most common locations for GISTs are the stomach (60%) and small intestine (30%), but GISTs can arise anywhere along the gastrointestinal tract. Gastric GISTs have been shown to be associated with a more favorable prognosis than GISTs at other sites. GISTs are most commonly diagnosed by upper endoscopy and/or CT of the abdomen as an incidental finding in an asymptomatic patient or in a patient being evaluated for symptoms of early satiety, abdominal pain, or gastrointestinal bleeding. GIST most frequently metastasizes to the liver and/or abdominal cavity.

**Radiologic Assessment**

Standard imaging techniques apply for GIST as for other intraabdominal sarcomas. In general, oral and IV contrast enhanced spiral CT is the staging modality of choice for GIST. CT scan of the abdomen and pelvis allows for assessment of the primary lesion and the presence or absence of intraabdominal disseminated disease or metastatic disease to the liver (the two most common locations for distant metastasis of GIST). FDG-PET has been reported to be useful for preoperative staging of GISTs because it may reveal early metastases and establish baseline metabolic activity and may be considered in select patients where equivocal findings are identified on CT or in the setting of following metabolic response to therapy. PET has been shown to be highly sensitive in detecting early response to imatinib treatment and in predicting long-term response in patients with metastatic GIST. If PET is to be used for monitoring response to therapy, baseline PET should be performed before initiation of treatment. Useful and effective CT-based criteria by Choi et al for detection of GIST and for predicting prognosis of GIST have also been proposed and may be used readily without incurring the cost and radiation exposure of PET/CT.

**Management of Localized Disease**

Complete surgical resection with negative margins is the recommended treatment for localized GISTs. Extended anatomic resection, wide margins, and lymphadenectomy are not required; therefore, total gastrectomy for gastric primaries is rarely required even with the largest of lesions. Resection of even locally advanced tumors is associated with improved survival. The 5-year survival rate for all patients with GISTs ranges from 20% to 44%, and the 5-year survival rate for patients with completely excised early-stage tumors is up to 75%. An analysis of 200 patients by DeMatteo and colleagues found a disease-specific survival rate of 54% for patients with grossly complete resection of primary GIST, and the median survival duration for patients with metastatic disease was only 20 months.

As for other soft tissue sarcomas, tumor size has consistently been identified as an important prognostic factor for GIST. Mitotic activity has also been identified as an important prognostic factor and is generally categorized as fewer than 5, 5 to 10, or more than 10 mitoses per high-power field. The National Institutes of Health and the Armed Forces Institute of Pathology have proposed prognostic criteria for risk stratification of surgically treated, localized primary GIST. Both groups take into account tumor size and mitotic count; the Armed Forces Institute of Pathology also includes tumor site as a prognostic variable. Accurate risk stratification is essential for selecting patients most likely to benefit from adjuvant treatment.

**Management of Locally Advanced or Metastatic Disease**

Treatment with imatinib mesylate (Gleevec, ST1571), a selective inhibitor of the KIT protein tyrosine kinase, has resulted in impressive clinical responses in a large percentage of patients with unresectable or metastatic GISTs. On the basis of the initial results in a single patient with metastatic GIST, the EORTC Soft Tissue and Bone Sarcoma Group initiated a phase 1 study to test the safety and efficacy of imatinib. In that study, 53% of patients with GISTs had confirmed partial responses; investigators concluded that imatinib is safe and effective against this disease. A multicenter, international trial of imatinib for GIST was begun in July 2000 at four treatment centers: Dana-Farber Cancer Institute, Oregon Health Sciences University, Fox Chase Cancer Center, and University Hospital of Helsinki, Finland. A total of 147 patients with unresectable or metastatic GISTs were randomized to 400 or 600 mg of imatinib daily for up to 24 months. Objective response was demonstrated in 79 patients (54%); all had partial responses, and there was no significant difference in response rate between imatinib doses. Fourteen percent of patients experienced disease progression. The toxicity profile was acceptable; the predominant effects were gastrointestinal effects (diarrhea, nausea), periorbital edema, muscle cramps, and fatigue. However, 21% of patients experienced serious (grade 3 or 4) adverse events, including gastrointestinal bleeding in 5% of patients, most likely related to the rapid tumor response of mural lesions.
A phase 3 randomized Intergroup trial was simultaneously performed to assess the clinical activity of imatinib at two dose levels for patients with unresectable or metastatic GIST expressing the c-Kit tyrosine kinase. From December 15, 2000, to September 1, 2001, 746 patients were accrued and randomized to low-dose (400 mg/d) or high-dose (800 mg/d) imatinib. The primary endpoint of the trial was survival. Preliminary toxicity data from 325 patients revealed a 23% incidence of grade 3 or 4 adverse events, including nausea and vomiting, gastrointestinal bleeding, abdominal pain, edema, fatigue, and rash.

In February 2002, the FDA approved imatinib for treatment of GIST based on the results of these promising clinical trials. Both the Intergroup trial mentioned in the preceding paragraph and a separate phase 3 trial compared the efficacy of low-dose (400 mg/d) and high-dose (800 mg/d) imatinib in patients with metastatic or unresectable GISTs. Both studies showed equivalent response rates and overall survival for the two doses but increased toxicity for the 800-mg/d dose. Current recommendations include consideration of dose escalation to 800 mg/d for patients who experience disease progression at a dose of 400 mg/d and for patients with advanced GIST and KIT exon 9 mutations.

The optimal duration of imatinib treatment, the duration of benefit from imatinib, and the long-term toxicity of imatinib have not been established. When feasible, imatinib should be continued in the absence of disease progression. A randomized trial reported worse median progression-free survival in patients who stopped imatinib after 1 year than in patients who continued beyond 1 year (progression-free survival of 6 months vs. 18 months). Less than 4% of patients with GISTs have experienced serious adverse events with imatinib. Mild gastrointestinal toxicity is the most frequently reported adverse event, but gastrointestinal tract hemorrhage, presumably from rapid tumor necrosis, has also been reported. Thus, all patients with GISTs treated on clinical protocols should be evaluated and followed by a team of medical professionals that includes a surgeon.

Many patients with GIST develop resistance to imatinib. Primary resistance is defined as clinical progression that develops during the first 6 months of treatment and is most commonly seen in patients with KIT exon 9 mutation, PDGFRα exon 18 mutation, or no mutations. Secondary resistance is defined as progression that develops more than 6 months after the start of treatment in a patient with an initial response. Imatinib resistance should be managed by either dose escalation or transition to treatment with sunitinib.

In 2006, sunitinib malate (SU11248, Sutent, Pfizer) emerged as an alternative systemic treatment for patients unable to tolerate imatinib and patients with imatinib-refractory GIST. Sunitinib is a tyrosine kinase inhibitor that targets multiple kinases, including the vascular endothelial growth factor receptors, PDGFRα, KIT, and FLT3. Sunitinib has both antiangiogenic and antiproliferative activity. In a phase 3, randomized, placebo-controlled trial, sunitinib was associated with a significant improvement in median time to progression (27.3 weeks vs. 6.4 weeks with placebo) in patients with imatinib-resistant GIST. In addition, sunitinib therapy was well tolerated; diarrhea, fatigue, and nausea were the most common adverse effects. Sunitinib has also been associated with hand-foot skin reaction, hypertension, cardiotoxicity, and hypothyroidism.

In 2006, sunitinib was approved by the FDA for treatment of patients with resistance or intolerance to imatinib.

The third FDA-approved drug recently made available for treatment of patients with imatinib- and sunitinib-resistant GIST is regorafenib. Regorafenib is a structurally unique inhibitor of multiple cancer-associated kinases, including KIT and PDGFRα, with broad-spectrum anticancer activity in preclinical and early-phase trials. Because KIT and PDGFRα remain drivers of GIST after resistance to imatinib and sunitinib, a multicenter, single-stage phase 2 trial examined regorafenib in patients with advanced GIST after failure of at least imatinib and sunitinib.

Thirty-four patients were enrolled from February to December 2010 and given regorafenib orally, 160 mg daily, on days 1 to 21 of a 28-day cycle. Clinical benefit was noted in 79% of patients with a median progression-free survival of 10 months. This trial was then followed by a phase 3, international, placebo-controlled, randomized trial of regorafenib for metastatic GIST after failure of imatinib and sunitinib. Patients with progression of disease on the placebo arm were crossed to the treatment arm of the study. Patients treated with regorafenib had a median progression-free survival of 4.8 months compared to 0.9 months on the placebo arm. Other tyrosine kinase inhibitors have demonstrated modest activity against GIST and target more than one family of protein kinases.

Among these are sorafenib, dasatinib, and nilotinib for the treatment of imatinib-resistant GIST, and these are now generally reserved for therapy after progression with regorafenib.

**Multidisciplinary Treatment**

Although imatinib has improved survival of patients with advanced GIST, most patients with advanced GIST are not cured with imatinib. Some patients develop secondary resistance to imatinib with one or more sites of disease progression after 6 months of clinical response (Fig. 36-10A [before imatinib], Fig. 36-10B [after imatinib]). The mechanisms of imatinib resistance are currently being investigated. Surgery has been shown to be beneficial for selected patients with isolated disease progression during imatinib therapy. Surgical resection of residual metastatic disease responding to imatinib-sensitive GIST has also been shown to result in progression-free survival in 70% to 96% of patients with imatinib- or sunitinib-sensitive GISTs.

The optimal timing of surgery in relation to imatinib therapy for patients with metastatic disease remains to be determined. It is not possible to compare outcomes for patients treated with kinase inhibitors alone and patients treated with kinase inhibitors plus surgical resection outside the context of randomized trials given the heterogeneity of patients and biases associated with selection of patients for surgical resection.

**Postoperative Imatinib**

Given the promising results of imatinib therapy for metastatic and locally advanced GIST, the next step was to study the efficacy of imatinib as adjuvant (postoperative) treatment in patients with surgically resectable disease, particularly those at high risk for recurrence because of large tumor size or high mitotic count. The ACOSOG first evaluated the efficacy of 1 year of postoperative imatinib in a single-arm phase 2 trial with 106 patients with high-risk GIST and compared the results with historical controls. Adjuvant treatment with imatinib for GIST patients was then examined in two key trials. In the ACOSOG randomized, double-blind, phase 3 Z9001 study, treatment with 12 months of imatinib was compared with placebo, following complete resection of a primary GIST smaller than 3 cm. Primary and secondary endpoints were recurrence-free survival.
intermediate or high risk of recurrence after resection and that at least 36 months of adjuvant imatinib be considered for patients at high risk of recurrence. Further, both the FDA and EMA updated the label, extending the duration of adjuvant therapy to at least 36 months in patients at high risk of recurrence.

Whether longer treatment durations may be of further benefit is still an open question that will be addressed by future studies. However, paralleling what is commonly done in the metastatic setting, many investigators believe that even adjuvant imatinib should become a chronic therapy. While moving toward more prolonged adjuvant treatment durations, it is all the more essential to identify the appropriate patients to treat to avoid the burden of adverse events or increased financial liability for patients who will not derive therapeutic benefit from imatinib. Risk stratification based on patients’ risk of recurrence is a key component to optimizing adjuvant treatment. The most practical stratification scheme to use for making a decision for adjuvant therapy is the modified National Institutes of Health consensus criteria. High-risk GIST patients, whose tumor harbors a sensitive genotype, should be treated by adjuvant imatinib because they have a poor prognosis. On the contrary, neither low-risk nor intermediate-risk GIST patients need adjuvant therapy, even if their tumor carries a sensitive genotype. In fact, evidence has been provided that the outcome of these patients with intermediate- or low-risk GIST is good. When using other risk stratification schemes, such as the Armed Forces Institute of Pathology table, Memorial Sloan-Kettering Cancer Center nomogram, or the heat map, there is a consensus to treat all patients having 30% or higher risk of recurrence, if their tumor carries a sensitive genotype. There is also a consensus not to treat patients having 10% or less risk of recurrence, even if their tumor carries a sensitive genotype. All patients having a risk between 10% and 30% should be evaluated on a case-by-case basis, and advantages/disadvantages of treatment should be made clear and discussed with the patient. Whenever a decision for adjuvant therapy is made, treatment duration of at least 36 months should be considered independently from the risk.

Beside the risk of recurrence, the other important factor to consider is the tumor genotype. In other words, as found in both the metastatic and adjuvant settings, GIST tumors with \textit{KIT} exon 11 and \textit{PDGFRA} non-D842V mutations are sensitive to imatinib. Patients with these mutations are suitable for adjuvant imatinib if the risk determined by stratification tools is significant. Patients with \textit{KIT} exon 9 mutations should also be treated; a higher dose (800 mg/d) may be more appropriate but remains to be studied clinically. Patients with \textit{PDGFRA} D842V–mutated tumors should not be treated with adjuvant imatinib, nor should patients with \textit{KIT} and \textit{PDGFRA} wild-type tumors associated with neurofibromatosis type 1 or the pediatric GIST/Carney-Stratakis syndromes, whatever the risk. Patients with sporadic wild-type GIST could be treated on an individual basis.

**Preoperative Imatinib**

Patients with marginally resectable GIST or at significant risk for operative morbidity should be considered for preoperative imatinib with close monitoring. Since the optimal duration of preoperative therapy is unknown, imatinib should be continued until maximal response is achieved or until there is evidence of progression. Preoperative imatinib can be stopped immediately before surgery and resumed when oral medications are restarted.
Preoperative imatinib in patients with primary GIST or resectable metastatic GIST has been evaluated in the context of two randomized phase 2 studies. The Radiologic Treatment Oncology Group (RTOG 0132/ACRIN 6665) evaluated the efficacy of preoperative imatinib (600 mg/d) for 8 to 10 weeks before surgery and 24 months after surgery in patients with primary (n = 30) or potentially resectable recurrent or metastatic (n = 22) tumors. Primarily stable disease was noted during imatinib treatment, and the 2-year progression-free survival rates were 83% for primary GIST and 77% for recurrent or metastatic GIST. In another study, 19 patients undergoing surgical resection at a single institution were randomized to preoperative imatinib (600 mg/d) for 3, 5, or 7 days followed by surgical resection and postoperative imatinib for 24 months. The response rate assessed using FDG-PET was 69%, and the median disease-free survival time following treatment with surgery and imatinib was 46 months. Similar results were observed in a prospective series of patients treated at a major institution.

All of these studies show that neoadjuvant treatment is feasible, but to maximize the benefit of preoperative therapy, the following factors need to be considered. First, the patient should in principle have a favorable mutational status; otherwise, the treatment would be in vain, allowing the tumor to continue growing. Fortunately, the majority of GISTs will respond, but it should not be forgotten that, especially in gastric location, the amount of insensitive mutations in the localized setting is less uncommon than what has been previously reported. However, we do not absolutely need to know the mutational status in advance, but we should be aware that mutation is an issue. Alternatively, if mutation status is not determined prior to treatment, we may also check response very early, either by CT, PET, or contrast-enhanced ultrasound. If a radiographic response is detected within a month, then the mutation status is likely favorable, and the treatment could be continued without necessarily pursuing mutation testing. If not, then mutation status should be investigated before continuing the treatment.

Second, the resectability of the tumor and the extent of resection necessary should be considered. Other than presentations of clearly inoperable tumors (which are usually treated with imatinib upfront) or symptomatic tumors requiring urgent intervention (hemorrhage, perforation, etc), there are few reasons today to perform extended procedures (i.e., multivisceral resections or formal organ resections) without first attempting preoperative therapy. For instance, patients with large GISTs who may require a long midline incision for resection may benefit from neoadjuvant imatinib to downstage the operation, potentially converting an open laparotomy approach to a laparoscopic one. Another typical circumstance in which neoadjuvant therapy may be beneficial is in patients with GISTs arising in the esophagus, gastroesophageal junction, duodenum, or distal rectum. Preoperative treatment may shrink the tumor and allow a more conservative local excision. In general, such patients would normally undergo postoperative adjuvant treatment as well because of the expected recurrence risk. The chance of obtaining a response is high, and the benefit for tumor shrinkage obvious, so this approach should always be discussed with patients affected by bulky and/or poorly located disease (esophagus, gastroesophageal junction, duodenum, distal rectum) as well as for those who would be candidates for an adjuvant treatment anyway.

Third, in the studies mentioned earlier, patients underwent surgery after a limited treatment duration (3 months at best). It is now well known that the preoperative treatment may result in sustained tumor shrinkage if given for a longer duration. Surgical resection may then be performed between 6 and 12 months or sooner if treatment effect plateaus. This allows for optimal tumor shrinkage, or at least shrinkage to the point where there is no further benefit to be gained by further neoadjuvant therapy without the risk of developing secondary resistance.

**DESMOIDS**

Desmoid tumors are not low-grade sarcomas but can be locally aggressive, although they do not metastasize. Approximately half of these tumors arise in the extremities; the remaining lesions are located on the trunk or in the retroperitoneum. Abdominal wall desmoids are associated with pregnancy and are thought to be the result of hormonal influence. Although usually sporadic, desmoids may occur in association with familial adenomatous polyposis, a presentation that is referred to as Gardner’s syndrome and is linked to germline mutations in the APC gene. Sporadic cases of desmoid fibromatosis are commonly linked to mutations in CTNNB1, the gene for β-catenin.

The primary therapy for desmoid tumors has long been considered surgical resection with wide local excision to achieve negative margins. However, local recurrence occurs in up to one-third of patients independently of the quality of surgical margins. Up to two-thirds of the patients operated on with positive margins do not recur. This is why there is growing evidence that the primary approach could be more conservative. Function-sparing operations should be the goal, even if a positive margin is left on a critical structure. Moreover, some authors advocate the possibility to observe patients at presentation, limiting surgery to those who progress or fail medical therapies. It has in fact been reported that by this approach, up to 50% of patients skip surgical resection.

Radiation therapy may be effective in patients with unresectable tumors or as adjuvant therapy following surgery for recurrent disease, although long-term side effects and the risk of radiation-induced sarcoma should always be considered. When used, a dose of 50 to 54 Gy is usually recommended. Systemic treatment is another option when surgery is not indicated. Hormonal therapies such as tamoxifen have been reported to be beneficial, as have nonsteroidal anti-inflammatory drugs, which are known to affect the β-catenin signaling pathways. Chemotherapy is also effective, although usually reserved for patients with tumor-associated symptoms who have not responded to other interventions. Combinations of methotrexate and vinblastine have been shown to have activity, as have single-agent pegylated liposomal doxorubicin and sorafenib.

Imatinib has also been studied with unconvincing results.

**DERMATOFIBROSARCOMA PROTUBERANS**

Dermatofibrosarcoma protuberans is a rare low-grade sarcoma arising in the dermis that rarely metastasizes but is locally aggressive. The overall annual incidence has been estimated at 4.2 cases per million individuals, and the incidence is higher among blacks than whites (6.5 vs. 3.9 per million per year). Approximately 40% of cases arise on the trunk, and most of the remaining tumors are distributed between the head and neck and the extremities. Dermatofibrosarcoma protuberans presents as a nodular, cutaneous mass that grows slowly and persistently. Satellite lesions may be found in patients with larger tumors.
Standard treatment is wide local excision, which generally results in local recurrence rates of less than 10%. Although local recurrence rates as high as 30% to 50% have been reported in population-based series, the associated 5-year survival rate is greater than 99%. Dermatofibrosarcoma protuberans arises from a specific chromosomal translocation involving chromosomes 17 and 22, in which the collagen 1 gene is fused to the gene for PDGF β-chain (PDGFB). The resultant deregulated expression of PDGFB leads to continuous activation of the PDGFR protein tyrosine kinase, which promotes tumor cell growth. The identification of this chromosomal translocation in more than 90% of cases of dermatofibrosarcoma protuberans has led to the development of targeted therapy. Inhibiting PDGFR with imatinib has been shown to induce clinical and radiologic improvement in patients with unresectable dermatofibrosarcoma protuberans. These data have resulted in the approval by the FDA of imatinib for treatment of patients with locally advanced dermatofibrosarcoma protuberans.

**PEDIATRIC SARCOMAS**

Soft tissue sarcomas in children are relatively rare, accounting for 7% to 8% of all pediatric cancers and totaling approximately 600 new cases per year. Pediatric sarcomas have traditionally been divided into two groups: rhabdomyosarcoma and nonrhabdomyosarcoma soft tissue sarcomas.

**Rhabdomyosarcoma**

Associated with skeletal muscle, rhabdomyosarcomas are the most common soft tissue tumors among children younger than 15 years and can occur at any site comprised of striated muscle. Patients with these tumors generally present with a painless enlarging mass; about 24% of tumors are located in the genitourinary system, 20% in the extremities, 20% in the head and neck, 16% in the parameningeal region, and 22% in other sites. Rhabdomyosarcoma is a small round cell tumor that demonstrates muscle differentiation upon light microscopy and immunohistochemical analysis. Two primary histologic subtypes account for 90% of cases: embryonal (70%) and alveolar (20%). Alveolar rhabdomyosarcoma is associated with cytogenetic translocation [t(2:13)(q35;q14)] in 85% to 90% of cases and [t(1:13)(p36;q14)] in 10% of cases. These translocations affect biologic activity at the levels of protein function and gene expression, thereby affecting the control of cell growth, apoptosis, differentiation, and motility and ultimately contributing to tumorigenic behavior. Whereas alveolar rhabdomyosarcomas often have translocations, most embryonal rhabdomyosarcomas have an allelic loss at chromosome 11p15.5 that is thought to inactivate a tumor suppressor gene. Both of these distinct molecular subtypes of rhabdomyosarcoma are thought to have similar alterations in downstream targets such as the p53 and Rb pathways. Further insight into these genetic alterations may lead to a better understanding of the pathogenesis of rhabdomyosarcoma and provide novel targets for therapeutic approaches.

Extent of disease is the strongest predictor of long-term outcome. Several staging systems for rhabdomyosarcoma are available. The Intergroup Rhabdomyosarcoma Study Group system is based on surgical-pathologic groupings. Multidisciplinary evaluation including pediatric oncologists, surgical subspecialists, and radiation oncologists is critical to plan the best treatment approach to maximize local tumor control while minimizing long-term treatment effects.

Complete surgical resection is the treatment of choice for rhabdomyosarcoma when function and cosmesis can be preserved. Patients who are able to undergo a complete tumor resection with negative (group I) or microscopic surgical margins (group II) are able to undergo less intensive systemic therapy and still have overall survival rates approaching 90%. At some anatomic sites, in particular the head and neck and genitourinary system, surgery is often avoided because the associated morbidity would be substantial. Recent findings suggest that chemotherapy alone can adequately control many such tumors. In the second International Society of Paediatric Oncology study of rhabdomyosarcoma (MMT84), the choice of local treatment was based on response to initial chemotherapy such that radical surgery and radiation therapy were avoided in 66% of patients. Among the patients who subsequently developed local relapse, the 5-year overall survival rate after salvage therapy was 46%.

Unlike other soft tissue sarcomas, rhabdomyosarcomas have a high propensity for lymph node metastasis, with rates up to 20% to 30% for sites such as the extremities, paratesticular nodes, and prostate. Lymph node sampling and, more recently, sentinel lymph node mapping have been used to evaluate regional node status in children with rhabdomyosarcoma. About 15% to 20% of patients with rhabdomyosarcoma have distant metastasis at presentation, most commonly (40–50% of cases) to the lungs, followed by bone marrow and bone. However, all patients with rhabdomyosarcoma are assumed to have micrometastatic disease at presentation. Therefore, multigent chemotherapy is recommended for all patients with rhabdomyosarcoma. Combination regimens including vincristine, dactinomycin, and cyclophosphamide continue to be the basis of effective curative therapy. Although various combinations including doxorubicin, ifosfamide, cisplatin, and etoposide have been shown to be active against rhabdomyosarcoma, they have not improved outcomes. Radiation therapy is given to most patients with microscopic residual disease (group II) after resection.

The prognosis for children with rhabdomyosarcomas is related to tumor site, surgical-pathologic grouping, and tumor histology. The 5-year disease-free survival rate for all patients has been reported to be 65%. Five-year disease-free survival rates for patients in groups I, II, III, and IV have been reported to be 84%, 74%, 62%, and 23%, respectively (see Table 36-3).

**Nonrhabdomyosarcoma Soft Tissue Sarcomas**

Approximately 60% of soft tissue sarcomas in children are nonrhabdomyosarcomas. These include numerous histologic subtypes, which are generally categorized into four groups: (a) fibrosarcoma, (b) Kaposi’s sarcoma, (c) other “specified” soft tissue sarcomas (e.g., synovial, angiosarcoma, hemangiopericytoma, leiomyosarcoma, liposarcoma, and extraosseous Ewing’s sarcoma), and (d) “unspecified” soft tissue sarcoma. The most common subtypes are synovial sarcoma, MPNST, and fibrosarcoma. No single histology accounts for more than 15% of all cases.

As with adult tumors, the evaluation of the soft tissue mass begins with a history and physical examination followed by imaging, which usually includes MRI. A CT scan of the chest is important for evaluation of metastatic disease. A core needle biopsy is generally required to establish a diagnosis. Surgery remains the primary treatment of nonrhabdomyosarcoma, and...
local control of large, high-grade tumors is improved with radiation therapy. The prognostic factors for children with nonrhabdomyosarcoma are similar to those for adults, and the role of chemotherapy for high-risk tumors is unclear, as for adults.

RESEARCH PERSPECTIVES

As the molecular alterations associated with various sarcoma subtypes are elucidated, many new potential targets for therapeutic intervention will be identified. A wide variety of DNA alterations have been observed in sarcomas that result in mutated genes encoding proteins ranging from transcription factors to tyrosine kinases to cytokines. The challenge in identifying therapeutic targets in sarcoma is to identify those that are specifically important to cellular function. The ideal therapeutic target has been described as a single molecule that is critical for pathogenesis, is expressed and active, is involved in a single pathway amenable to blockade (i.e., no alternative bypass pathways exist), and is critical for sarcoma cell survival.25

CONCLUSIONS

Soft tissue sarcomas are a heterogeneous family of rare tumors, accounting for approximately 1% of malignancies in adults. The etiology in the vast majority of patients is sporadic, and the management of such diverse tumors is complex. Diagnosis by light microscopy is inexact, but molecular diagnosis, although still in its infancy, holds great promise. The natural history of soft tissue sarcomas is well established. Approximately two-thirds of cases arise in the extremities, and the remaining one-third are distributed between the retroperitoneum, trunk, abdomen, and head and neck. The management algorithm for soft tissue sarcomas is complex and depends on tumor stage, site, and histology.

The most common site of metastasis is the lungs, and metastasis generally occurs within 3 years of diagnosis. Soft tissue sarcomas have unique molecular profiles that contribute to varying responses to systemic therapy. Doxorubicin-based regimens have been the mainstay of treatment for the past two decades; however, it is now clear that specific histologic subtypes have increased sensitivity to specific agents. For example, angiosarcomas are more sensitive to paclitaxel, while leiomyosarcoma is sensitive to gemcitabine and docetaxel.

Progress in understanding of soft tissue sarcoma biology is crucial for the development of additional therapeutic targets. Drug engineering will enable molecular-based therapies to become increasingly incorporated into clinical trials and, with success, into standard treatment strategies for soft tissue sarcomas in the near future.

REFERENCES

Entries highlighted in bright blue are key references.


This page intentionally left blank
INTRODUCTION

Inguinal herniorrhaphy is one of the most commonly performed operations in the United States. Based on estimates made by the National Center for Health Statistics, in 2010 nearly 515,000 inguinal hernia operations were performed in hospitals, and an additional 450,000 were performed in ambulatory surgery centers.

Approximately 75% of abdominal wall hernias occur in the groin. Of inguinal hernia repairs, 90% are performed in men, and 10% are performed in women. This is thought to be because the lifetime risk of inguinal hernia is 27% in men and 3% in women. The incidence of inguinal hernia in men has a bimodal distribution, with peaks before the first year of age and after age 40. Abramson demonstrated the age-dependence of inguinal hernias in 1978. Those age 25 to 34 years had a lifetime prevalence rate of 15%, whereas those age 75 years and over had a rate of 47% (Table 37-1). Approximately 70% of femoral hernia repairs are performed in women; however, the most common subtype of groin hernia in men and women is still the indirect inguinal hernia. Inguinal hernias are five times more common than femoral hernias.

Globally, the inguinal hernia repair has become one of the most important procedures in improving quality of life and preventing disability. In one study, an international cooperative organization performed over 1033 hernia repairs on 926 patients, and their impact was measured in disability adjusted life years (DALYs). They were able to avoid 5014 DALYs or 5.41 DALYs per patient.

History

Surgical repair of hernias has been documented as far back as in ancient Egyptian and Greek civilizations. In the past, early management of inguinal hernias often involved a conservative approach with operative management reserved only for complications. Surgery often involved routine excision of the testicle, and wounds were closed with cautery or allowed to close by secondary intention. These procedures were performed without aseptic technique, and infection and recurrence rates were high.

From the late 1700s to the early 1800s, physicians including Hesselbach, Cooper, Camper, Scarpa, Richter, and Gimbernat identified vital components of the inguinal region from cadaveric dissection. This improved understanding of the anatomy and pathophysiology of inguinal hernias. These findings, coupled with the development of aseptic technique, led surgeons such as Marcy, Kocher, and Lucas-Championnière to perform sac dissection, high ligation, and closure of the internal ring. This led to improved outcomes, but recurrence rates remained unacceptably high.

At around this time, Bassini (1844–1924) pioneered a new method that transformed inguinal hernia repair into a successful venture with minimal morbidity. The success of the Bassini repair over its predecessors ushered in an era of tissue-based repairs. The Bassini repair was then modified into the McVay and Shouldice repairs. All three of these techniques, as well as modern variations such as the Desarda operation, are currently practiced.

The next major advancement in inguinal hernia repair was in the 1980s. At this time, Lichtenstein applied a piece of mesh to the floor of the inguinal canal, allowing for a truly tension-free repair. This technique demonstrated superior outcomes compared to previous tissue-based repairs. There were several other advantages of this process. In addition to being truly tension-free, the mesh could restore the strength of the transversalis fascia, and importantly, the technique had a very short learning curve. The superior outcomes have been widely reproduced regardless of hernia size and type, and they were achievable among both expert and nonexpert hernia surgeons.

With the advent of minimally invasive surgery, inguinal hernia repair underwent its most recent transformation. Laparoscopic inguinal hernia repair offers an alternative approach, minimizes postoperative pain, and improves recovery. Since the initial description by Ger, the laparoscopic method has become more sophisticated. Refinements in approach and technique have performed without aseptic technique, and infection and recurrence rates were high.

From the late 1700s to the early 1800s, physicians including Hesselbach, Cooper, Camper, Scarpa, Richter, and Gimbernat identified vital components of the inguinal region from cadaveric dissection. This improved understanding of the anatomy and pathophysiology of inguinal hernias. These findings, coupled with the development of aseptic technique, led surgeons such as Marcy, Kocher, and Lucas-Championnière to perform sac dissection, high ligation, and closure of the internal ring. This led to improved outcomes, but recurrence rates remained unacceptably high.

At around this time, Bassini (1844–1924) pioneered a new method that transformed inguinal hernia repair into a successful venture with minimal morbidity. The success of the Bassini repair over its predecessors ushered in an era of tissue-based repairs. The Bassini repair was then modified into the McVay and Shouldice repairs. All three of these techniques, as well as modern variations such as the Desarda operation, are currently practiced.

The next major advancement in inguinal hernia repair was in the 1980s. At this time, Lichtenstein applied a piece of mesh to the floor of the inguinal canal, allowing for a truly tension-free repair. This technique demonstrated superior outcomes compared to previous tissue-based repairs. There were several other advantages of this process. In addition to being truly tension-free, the mesh could restore the strength of the transversalis fascia, and importantly, the technique had a very short learning curve. The superior outcomes have been widely reproduced regardless of hernia size and type, and they were achievable among both expert and nonexpert hernia surgeons.

With the advent of minimally invasive surgery, inguinal hernia repair underwent its most recent transformation. Laparoscopic inguinal hernia repair offers an alternative approach, minimizes postoperative pain, and improves recovery. Since the initial description by Ger, the laparoscopic method has become more sophisticated. Refinements in approach and technique have performed without aseptic technique, and infection and recurrence rates were high.

From the late 1700s to the early 1800s, physicians including Hesselbach, Cooper, Camper, Scarpa, Richter, and Gimbernat identified vital components of the inguinal region from cadaveric dissection. This improved understanding of the anatomy and pathophysiology of inguinal hernias. These findings, coupled with the development of aseptic technique, led surgeons such as Marcy, Kocher, and Lucas-Championnière to perform sac dissection, high ligation, and closure of the internal ring. This led to improved outcomes, but recurrence rates remained unacceptably high.

At around this time, Bassini (1844–1924) pioneered a new method that transformed inguinal hernia repair into a successful venture with minimal morbidity. The success of the Bassini repair over its predecessors ushered in an era of tissue-based repairs. The Bassini repair was then modified into the McVay and Shouldice repairs. All three of these techniques, as well as modern variations such as the Desarda operation, are currently practiced.

The next major advancement in inguinal hernia repair was in the 1980s. At this time, Lichtenstein applied a piece of mesh to the floor of the inguinal canal, allowing for a truly tension-free repair. This technique demonstrated superior outcomes compared to previous tissue-based repairs. There were several other advantages of this process. In addition to being truly tension-free, the mesh could restore the strength of the transversalis fascia, and importantly, the technique had a very short learning curve. The superior outcomes have been widely reproduced regardless of hernia size and type, and they were achievable among both expert and nonexpert hernia surgeons.

With the advent of minimally invasive surgery, inguinal hernia repair underwent its most recent transformation. Laparoscopic inguinal hernia repair offers an alternative approach, minimizes postoperative pain, and improves recovery. Since the initial description by Ger, the laparoscopic method has become more sophisticated. Refinements in approach and technique have performed without aseptic technique, and infection and recurrence rates were high.
Key Points

1. Conservative management of asymptomatic inguinal hernias is recommended.
2. A thorough understanding of groin anatomy is essential to successful surgical treatment of inguinal hernias.
3. Elective repair of inguinal hernias can be undertaken using a laparoscopic, robotic, or open approach.
4. Robotic-assisted hernia surgery is quickly becoming adopted by general surgeons because of its better ergonomics and visualization.
5. The use of prosthetic mesh as a reinforcement significantly improves recurrence rates, whether the repair is open or laparoscopic.
6. Recurrence, pain, and quality of life are the metrics by which hernia repair outcomes are measured.
7. Laparoscopic inguinal hernia repair results in less pain; however, mastery of this technique has a longer learning curve.

Anatomy

The inguinal canal is an approximately 4- to 6-cm long cone-shaped region situated in the anterior portion of the pelvic basin (Fig. 37-1). The canal begins on the posterior abdominal wall, where the spermatic cord passes through a hiatus in the transversalis fascia also known as the deep (internal) inguinal ring. The canal concludes medially at the superficial (external) inguinal ring, the point at which the spermatic cord crosses a defect in the external oblique aponeurosis. The boundaries of the inguinal canal are the external oblique aponeurosis anteriorly, the internal oblique muscle laterally, the transversalis fascia and transversus abdominis muscle posteriorly, the internal oblique and transversus abdominis muscle superiorly, and the inguinal (Poupart’s) ligament inferiorly. The spermatic cord traverses the inguinal canal, and it contains three arteries, three veins, two nerves, the pampiniform venous plexus, and the vas deferens. It is enveloped in three layers of spermatic fascia.

Additional important structures surrounding the inguinal canal include the iliopubic tract, the lacunar ligament, Cooper’s ligament, and the conjoined tendon (Fig. 37-2). The iliopubic tract is an aponeurotic band that begins at the anterior superior iliac spine and inserts into Cooper’s ligament from above. It forms on the deep inferior margin of the transversus abdominis and transversalis fascia. The shelving edge of the inguinal ligament is a structure that connects the iliopubic tract to the inguinal ligament. The iliopubic tract helps form the inferior margin of the internal inguinal ring as it courses medially, where it continues as the anteromedial border of the femoral canal. The lacunar ligament, or ligament of Gimbert, is the triangular fanning of the inguinal ligament as it joins the pubic tubercle. Cooper’s (pectineal) ligament is the lateral portion of the lacunar ligament that is fused to the periosteum of the pubic tubercle. The conjoined tendon is commonly described as the fusion of the inferior fibers of the internal oblique and transversus abdominis aponeurosis at the point where they insert on the pubic tubercle.

Inguinal hernias are generally classified as direct, indirect, or femoral based upon the site of herniation relative to surrounding structures. Indirect hernias protrude lateral to the inferior epigastric vessels, through the deep inguinal ring. Direct hernias protrude medial to the inferior epigastric vessels, within Hesselbach’s triangle. The borders of the triangle are the inguinal ligament inferiorly, the lateral edge of rectus sheath medially, and the inferior epigastric vessels superolaterally. Femoral hernias protrude through the small and inflexible femoral ring. They traverse the empty space between the femoral vein and the lymphatic channels. The borders of the femoral ring include the iliopubic tract and inguinal ligament anteriorly, Cooper’s ligament posteriorly, the lacunar ligament medially, and the femoral vein laterally. The Nyhus classification categorizes hernia defects by location, size, and type (Table 37-2).

Laparoscopic inguinal hernia repair requires a thorough knowledge of inguinal anatomy from a posterior perspective (Fig. 37-3). Intraperitoneal points of reference are the five peritoneal folds, bladder, inferior epigastric vessels, and psoas muscle (Fig. 37-4). Two potential spaces exist within the preperitoneum. Between the peritoneum and the posterior lamina of the transversalis fascia is Bogros’s (preperitoneal) space. This area contains preperitoneal fat and areolar tissue. The most medial aspect of the preperitoneal space, that which lies superior to the bladder, is known as the space of Retzius. The posterior perspective also allows visualization of the myopectineal orifice of Fruchaud, a relatively weak portion of the abdominal wall that is divided by the inguinal ligament (Fig. 37-5).

The vascular space is situated between the posterior and anterior laminae of the transversalis fascia, and it houses the

Table 37-1
Inguinal hernia prevalence by age

<table>
<thead>
<tr>
<th>AGE (Y)</th>
<th>25–34</th>
<th>35–44</th>
<th>45–54</th>
<th>55–64</th>
<th>65–74</th>
<th>75+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current prevalence (%)</td>
<td>12</td>
<td>15</td>
<td>20</td>
<td>26</td>
<td>29</td>
<td>34</td>
</tr>
<tr>
<td>Lifetime prevalence (%)</td>
<td>15</td>
<td>19</td>
<td>28</td>
<td>34</td>
<td>40</td>
<td>47</td>
</tr>
</tbody>
</table>

Current = repaired hernias excluded; lifetime = repaired hernias included.
in inferior epigastric vessels. The inferior epigastric artery supplies the rectus abdominis. It is derived from the external iliac artery, and it anastomoses with the superior epigastric, a continuation of the internal thoracic artery. The epigastric veins course parallel to the arteries within the rectus sheath, posterior to the rectus muscles. Inspection of the internal inguinal ring will reveal the deep location of the inferior epigastric vessels.

Nerves of interest in the inguinal region are the ilioinguinal, iliohypogastric, genitofemoral, and lateral femoral cutaneous nerves (Figs. 37-6 and 37-7). The ilioinguinal and iliohypogastric nerves arise together from the first lumbar nerve (L1). The ilioinguinal nerve emerges from the lateral border of the psoas major and passes obliquely across the quadratus lumborum. At a point just medial to the anterior
Table 37-2

Nyhus classification system

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I</td>
<td>Indirect hernia; internal abdominal ring normal; typically in infants, children, small adults</td>
</tr>
<tr>
<td>Type II</td>
<td>Indirect hernia; internal ring enlarged without impingement on the floor of the inguinal canal; does not extend to the scrotum</td>
</tr>
<tr>
<td>Type IIIA</td>
<td>Direct hernia; size is not taken into account</td>
</tr>
<tr>
<td>Type IIIB</td>
<td>Indirect hernia that has enlarged enough to encroach upon the posterior inguinal wall; indirect sliding or scrotal hernias are usually placed in this category because they are commonly associated with extension to the direct space; also includes pantaloon hernias</td>
</tr>
<tr>
<td>Type IIIC</td>
<td>Femoral hernia</td>
</tr>
<tr>
<td>Type IV</td>
<td>Recurrent hernia; modifiers A–D are sometimes added, which correspond to indirect, direct, femoral, and mixed, respectively</td>
</tr>
</tbody>
</table>

superior iliac spine, it pierces the transversus and internal oblique muscles to enter the inguinal canal and exits through the superficial inguinal ring. It supplies somatic sensation to the skin of the upper and medial thigh. In males, it also innervates the base of the penis and upper scrotum. In females, it innervates the mons pubis and labium majus. The iliohypogastric nerve arises from T12–L1. After it pierces the deep abdominal wall, it courses between the internal oblique and transversus abdominis, supplying both. It then divides into lateral and anterior cutaneous branches. A common variant is for the iliohypogastric and ilioinguinal nerves to exit around the superficial inguinal ring as a single entity. The genitofemoral nerve arises from L1 to L2, courses along the retroperitoneum, and emerges on the anterior aspect of the psoas. It then divides into genital and femoral branches. The genital branch enters the inguinal canal lateral to the inferior epigastric vessels, and it courses ventral to the iliac vessels and iliopubic tract. In males, it travels through the superficial inguinal ring and supplies the ipsilateral scrotum and cremaster muscle. In females, it supplies the ipsilateral mons pubis and labium majus. The

Figure 37-3. Anatomy of the groin region from the posterior perspective.

femoral branch courses along the femoral sheath, supplying the skin of the upper anterior thigh. The lateral femoral cutaneous nerve arises from L2 to L3, emerges lateral to the psoas muscle at the level of L4, and crosses the iliacus muscle obliquely toward the anterior superior iliac spine. It then passes inferiorly to the inguinal ligament where it divides to supply the lateral thigh (Fig. 37-8).

The preperitoneal anatomy seen in laparoscopic hernia repair led to characterization of important anatomic areas of interest, known as the triangle of doom, the triangle of pain, and the circle of death (Fig. 37-9). The triangle of doom is bordered medially by the vas deferens and laterally by the vessels of the spermatic cord. The contents of the space include the external iliac vessels, deep circumflex iliac vein, femoral nerve, and genital branch of the genitofemoral nerve. The triangle of pain is a region bordered by the iliopubic tract and gonadal vessels, and it encompasses the lateral femoral cutaneous, femoral branch of the genitofemoral and femoral nerves. The circle of death is a vascular continuation formed by the common iliac, internal iliac, obturator, inferior epigastric, and external iliac vessels.

Figure 37-5. Posterior view of the myopectineal orifice of Fruchaud. a. = artery; n. = nerve; v. = vein.

Figure 37-6. Retroperitoneal view of major inguinal nerves and their courses. m. = muscle; n. = nerve.
Pathophysiology

Inguinal hernias may be congenital or acquired. Most adult inguinal hernias are considered acquired defects in the abdominal wall. There is however, a known hereditary association that is not well understood. The most likely risk factor for inguinal hernia is weakness in the abdominal wall musculature; however, there are several other risk-factors that have been studied (Table 37-3). Congenital hernias, which make up the majority of pediatric hernias, can be considered a developmental defect rather than an acquired weakness. During the normal course of development, the testes descend from the intra-abdominal space into the scrotum in the third trimester. Their descent is guided by the gubernaculum through an evagination of the peritoneum, which protrudes through the inguinal canal and becomes the processus vaginalis. Between 36 and 40 weeks’ gestation, the processus vaginalis closes and eliminates the peritoneal opening at the internal inguinal ring. Failure of the peritoneum to close results in a patent processus vaginalis (PPV). In preterm babies, indirect inguinal hernias as a result of PPV is very high. Overall, the risk of developing a symptomatic hernia during childhood in the presence of a known PPV is relatively low.

Overall, there is limited data regarding the etiology of inguinal hernia development. Several studies have documented strenuous physical activity as a risk factor for acquired inguinal hernia. A case-controlled study of over 1400 male patients with inguinal hernia revealed that a positive family history was associated with an eightfold lifetime incidence of
Inferolateral border: iliopubic tract
Deep circumflex iliac a. & v.
Lateral border: reflected peritoneum
Lat. femoral cutaneous n.
Ant. femoral cutaneous n. or other variable branches
Femoral br. of genitofemoral n.
Superomedial border: gonadal vessels
Deep circumflex iliac a. & v.

Figure 37-9. Borders and contents of the (A) triangle of doom and (B) triangle of pain. a. = artery; Ant. = anterior; br. = branch; Lat. = lateral; n. = nerve; v. = vein. (Modified with permission from Colborn GL, Skandalakis JE: Laparoscopic cadaveric anatomy of the inguinal area, Probl Gen Surg. 1995;12(1):13-20.)

Epidemiologic studies have identified risk factors that may predispose to a hernia. Microscopic examination of skin of inguinal hernia patients demonstrated significantly decreased ratios of type I to type III collagen. Type III collagen does not contribute to wound tensile strength as significantly as type I collagen. Additional analyses of similar skin revealed disaggregated collagen tracts with decreased collagen fiber density. Collagen disorders such as Ehlers-Danlos syndrome are also associated with an increased incidence of hernia formation (Table 37-4). Recent studies have found an association between concentrations of extracellular matrix...
Table 37-3

Presumed causes of groin herniation

- Coughing
- Chronic obstructive pulmonary disease
- Obesity
- Straining
  - Constipation
  - Prostatism
- Pregnancy
- Birthweight <1500 g
- Family history of a hernia
- Valsalva’s maneuver
- Ascites
- Upright position
- Congenital connective tissue disorders
- Defective collagen synthesis
- Previous right lower quadrant incision
- Arterial aneurysms
- Cigarette smoking
- Heavy lifting
- Physical exertion

Table 37-4

Connective tissue disorders associated with groin herniation

- Osteogenesis imperfecta
- Cutis laxa (congenital elastolysis)
- Ehlers-Danlos syndrome
- Hurler-Hunter syndrome
- Marfan’s syndrome
- Congenital hip dislocation in children
- Polycystic kidney disease
- α₁-Antitrypsin deficiency
- Williams syndrome
- Androgen insensitivity syndrome
- Robinow’s syndrome
- Serpentine fibula syndrome
- Alport’s syndrome
- Tel Hashomer camptodactyly syndrome
- Leriche’s syndrome
- Testicular feminization syndrome
- Rokitsky-Mayer-Küster syndrome
- Goldenhar’s syndrome
- Morris syndrome
- Gerhardt’s syndrome
- Menkes’ syndrome
- Kawasaki disease
- Pfannenstiel syndrome
- Beckwith-Wiedemann syndrome
- Rubinstein-Taybi syndrome
- Alopecia-photophobia syndrome

DIAGNOSIS

History

Workup for inguinal hernia begins with a detailed history. The most common symptom of inguinal hernia is a groin mass that protrudes while standing, coughing, or straining. It is sometimes described as reducible while lying down. Symptoms that are extrainguinal such as a change in bowel habits or urinary symptoms are far less common but should be recognized as having the potential to be ominous. The pain is thought to be due to compression of the nerves by the sac, causing generalized pressure, localized sharp pain, or referred pain. Referred pain may involve the scrotum, testicle, or inner thigh.

Important considerations of the patient’s history include the duration and timing of symptoms. Sudden onset symptoms are more concerning. Questions should also be directed to characterize whether the hernia is reducible. Patients will often reduce the hernia by pushing the contents back into the abdomen, thereby providing temporary relief. As the defect size increases and more intra-abdominal contents fill the hernia sac, the hernia may become harder to reduce and incarcerate, prompting urgent surgical intervention.

Certain elements of the review of systems such as chronic constipation, cough, or urinary retention should prompt the surgeon to perform a thorough workup to rule out any underlying malignancy.
Physical Examination
Physical examination is essential to the diagnosis of inguinal hernia. The patient should be examined in a standing position to increase intra-abdominal pressure, with the groin and scrotum fully exposed. Inspection is performed first, with the goal of identifying an abnormal bulge along the groin or within the scrotum. If an obvious bulge is not detected, palpation is performed to confirm the presence of the hernia.

Palpation is performed by advancing the index finger through the scrotum towards the external inguinal ring (Fig. 37-11). This allows the inguinal canal to be explored. The patient is then asked to perform a Valsalva maneuver to increase intra-abdominal pressure. These maneuvers will reveal an abnormal bulge and allow the clinician to determine whether the hernia is reducible or not. Examination of the contralateral side affords the clinician the opportunity to compare the presence and extent of herniation between sides. This is especially useful in the case of a small hernia. In addition to inguinal hernia, a number of other diagnoses may be considered in the differential of a groin bulge (Table 37-5).

While very difficult to ascertain, there are certain physical examination maneuvers that can be performed to help distinguish direct vs. indirect inguinal hernias. The inguinal occlusion test entails the examiner blocking the internal inguinal ring with a finger as the patient is instructed to cough. A controlled impulse suggests an indirect hernia, while persistent herniation suggests a direct hernia. Transmission of the cough impulse to the tip of the finger implies an indirect hernia, while an impulse palpated on the dorsum of the finger implies a direct hernia. When results of physical examination are compared against operative findings, there is a probability somewhat higher than chance (i.e., 50%) of correctly diagnosing the type of hernia.21,22

External groin anatomy is difficult to assess in obese patients, making the physical diagnosis of inguinal hernia challenging. A further challenge to the physical examination is the identification of a femoral hernia. Femoral hernias should be palpable below the inguinal ligament, lateral to the pubic tubercle. In obese patients, a femoral hernia may be missed or misdiagnosed as a hernia of the inguinal canal. In contrast, a prominent inguinal fat pad in a thin patient, otherwise known as a femoral pseudohernia, may prompt an erroneous diagnosis of femoral hernia.

Imaging
In the case of an ambiguous diagnosis, radiologic investigations may be used as an adjunct to history and physical examination. Imaging in obvious cases is unnecessary. The most common radiologic modalities include ultrasonography (US), computed tomography (CT), and magnetic resonance imaging (MRI). Each technique has certain advantages over physical examination alone; however, each modality is associated with potential limitations.

US is the least invasive technique and does not impart any radiation to the patient. Anatomic structures can be more easily identified by the presence of bony landmarks; however, because there are few bones in the inguinal canal, other structures such as the inferior epigastric vessels are used to define groin anatomy. Positive intra-abdominal pressure is used to elicit the herniation of abdominal contents. Movement of these contents through the canal is essential to making the diagnosis with US.
and lack of this movement may lead to a false negative. A recent meta-analysis demonstrated that ultrasound detects inguinal hernia with a sensitivity of 86%, specificity of 77%. In thin patients, normal movement of the spermatic cord and posterior abdominal wall against the anterior abdominal wall may lead to false-positive diagnoses of hernia.

CT and MRI provide static images that are able to delineate groin anatomy, to detect groin hernias, and to exclude potentially confounding diagnoses (Fig. 37-12). Meta-analysis determined standard CT detects inguinal hernia with a sensitivity of 80%, specificity of 65%. Although direct herniography has a higher sensitivity and specificity than CT, its invasiveness and limited availability restrict its routine use. As CT imaging increases in resolution, its sensitivity in detecting inguinal hernia is expected to expand; however, this has yet to be clinically confirmed by surgical correlation.

MRI is most commonly utilized in cases where physical examination detects a groin bulge, but where ultrasonography is inconclusive. In a 1999 study of 41 patients with clinical findings of inguinal hernia, laparoscopy revealed that MRI was an effective diagnostic test with a sensitivity of 95%, specificity of 96%. The expense of MRI precludes its routine use to diagnose inguinal hernias.

**TREATMENT**

Surgical repair of hernias can be performed open, laparoscopic, or with robotic assistance. Surgical repair is the definitive treatment of inguinal hernias. The most common reason for elective repair is pain. Incarceration and strangulation are the primary indications for urgent repair. Symptomatic hernias should be operated on electively, and minimally symptomatic or asymptomatic hernias should undergo watchful waiting. Repair of minimally symptomatic inguinal hernia in patients with significant medical comorbidities surgery should be deferred and the patient medically optimized. If despite optimal management of comorbidities, the patient remains high-risk, open repair under local anesthesia can be safely performed. Although the natural history of untreated inguinal hernias is poorly defined, the rates of incarceration and strangulation are low in the asymptomatic population. As a result, nonoperative management is an appropriate consideration in minimally symptomatic patients. Prospective studies and meta-analyses have demonstrated no difference in intention-to-treat outcomes, quality of life, or cost-effectiveness between watchful waiting and elective repair among healthy inguinal hernia patients. A 2012 systematic review found that 72% of asymptomatic inguinal hernia patients developed symptoms and had surgical repair within 7.5 years of diagnosis. Nevertheless, the complication rates of immediate and delayed elective repair are equivalent. A nonoperative strategy is safe for minimally symptomatic inguinal hernia patients.

Nonoperative inguinal hernia treatment targets pain, pressure, and protrusion of abdominal contents in the symptomatic patient population. The recumbent position aids in hernia reduction via the effects of gravity and a relaxed abdominal wall. Trusses externally confine hernias to a reduced state and intermittently relieve symptoms in up to 65% of patients; however, they do not prevent complications, and they may be associated with an increased rate of incarceration. The risks of incarceration and strangulation appear to decrease over the first year, likely because gradual enlargement of the abdominal wall defect facilitates spontaneous reduction of hernia contents. The sheer volume of protruding tissue in an inguinal hernia does not necessarily signify severe morbidity.

Femoral and symptomatic inguinal hernias carry higher complication risks, and so surgical repair is performed earlier for these patients. Irrespective of symptoms, one study found the 3-month and 2-year cumulative incidences of strangulation were 2.8% and 4.5%, respectively, for inguinal hernias and 22% and 45%, respectively, for femoral hernias. Data from the Swedish Hernia Registry demonstrates that emergent operation is associated with a sevenfold increase in all-cause mortality over that of elective surgery among 107,838 groin hernia repairs. For this reason, it is recommended that femoral hernias and symptomatic inguinal hernias be electively repaired, when possible.

Incarceration occurs when hernia contents fail to reduce; however, a minimally symptomatic, chronically incarcerated hernia may also be treated nonoperatively. Taxis should be attempted for incarcerated hernias without sequelae of strangulation, and the option of surgical repair should be discussed prior to the maneuver. To perform taxis, analgesics and light sedatives are administered, and the patient is placed in the Trendelenburg position. The hernia bulge is elongated with both hands, and while slight countertraction is maintained, reduction of the contents is attempted circumferentially in a small stepwise fashion to ease their reduction into the abdomen.

The indication for emergent inguinal hernia repair is impending compromise of intestinal contents. As such, strangulation of hernia contents is a surgical emergency. Clinical signs that indicate strangulation include tenderness, fever, leukocytosis, and hemodynamic instability. The hernia bulge is usually warm, tender, and the overlying skin is often erythematos or discolored. Symptoms of bowel obstruction in patients with sliding or incarcerated inguinal hernias may also indicate strangulation. Taxis should not be performed when strangulation is suspected, as reduction of potentially gangrenous tissue into the abdomen may result in an intra-abdominal catastrophe. Preoperatively, the patient should receive fluid resuscitation, nasogastric decompression, and prophylactic intravenous antibiotics.
Prophylactic Antibiotics

The debate as to whether or not to administer preoperative prophylactic antibiotics in elective inguinal hernia repair still remains controversial as elective hernia repair is considered a clean procedure and as such are exempt from SCIP surgical prophylaxis guidelines. A Cochrane review of 17 randomized controlled trials in 2012 revealed an overall decrease in infection rates (3.1% vs. 4.5%, odds ratio [OR] 0.64, 95% confidence interval [CI] 0.50–0.82) when prophylactic antibiotics are administered in patients. In subgroup analyses, the difference was smaller in patients without mesh placement (3.5% vs. 4.9%, OR 0.71, 95% CI 0.51–1.00) than in those with mesh placement (2.4% vs. 4.2%, OR 0.56, 95% CI 0.38–0.81). However, with inguinal hernia repair, overall wound infection rates were higher than those expected for clean operations, as a result, they were unable to definitively recommend for or against antimicrobial prophylaxis.36,37 Although there is no universal guideline regarding the administration of prophylactic antibiotics for open elective hernia repair, the routine indexing of cases for quality improvement databases have resulted in the routine administration of prophylactic perioperative antibiotics in inguinal hernia repairs.

Open Approach

The most commonly performed type of hernia operation still remains the open inguinal hernia repair. These repairs can be performed tension-free with mesh or by reconstruction of the floor with tissue. Tissue repairs are less common and are primarily indicated in infected fields.

Exposure of the anterior inguinal region is common to the open approaches. An oblique or horizontal incision is performed over the groin (Fig. 37-13). The incision begins two fingerbreadths inferior and medial to the anterior superior iliac spine. It is then extended medially for approximately 6 to 8 cm. The subcutaneous tissue is dissected using electrocautery. Scarpa’s fascia is divided to expose the external oblique aponeurosis. A small incision is made in the external oblique aponeurosis parallel to the direction of the muscle fibers.
Specific Considerations

Figure 37-14. Anterior open exposure of the inguinal canal. m. = muscle; n. = nerve; v. = vein.

1. Metzenbaum scissors are introduced and spread beneath the fibers to sweep away the underlying ilioinguinal nerve. The scissors are then used to incise the aponeurosis superior to the inguinal ligament, splitting the external inguinal ring.

2. The flaps of the external oblique aponeurosis are elevated with Hemostat clamps. The internal oblique fibers are dissected bluntly from the overlying external oblique flaps. Dissection of the inferior flap reveals the shelving edge of the inguinal ligament. The iliohypogastric and ilioinguinal nerves are identified and preserved. Effort should be made to avoid removing nerves from their natural bed and disrupting the protective investing fascia. The pubic tubercle is identified, and the cord structures are dissected off of the pubis, encircled, and elevated with a Penrose drain. The cord is elevated 2 cm over the pubic symphysis in an avascular plane, and csematic fibers are preserved to avoid injuring cord structures (Fig. 37-14).

3. An indirect hernia sac will generally be found on the anteromedial surface of the spermatic cord after division of the cremasteric muscle in the direction of its fibers. The genital nerve is visualized along the inferolateral surface of the cord adjacent to the external spermatic vein. The floor of the inguinal canal is fully assessed for direct hernias. If a hernia is not visualized upon entry into the inguinal canal, the preperitoneal space should be explored for a femoral hernia. In addition to sac identification, the vas deferens and vessels of the spermatic cord must be identified to allow dissection of the sac from the cord. Blunt dissection facilitates dissection of the sac from the cord. The dissection is carried proximally toward the deep inguinal ring.

4. In cases where the viability of sac contents is in question, the sac should be incised, and hernia contents should be evaluated for signs of ischemia. The defect should be enlarged to augment blood flow to the sac contents. Viable contents may be reduced into the peritoneal cavity, while nonviable contents resected. In elective cases, the sac may be amputated at the internal inguinal ring. Care is taken to avoid injury to the preperitoneal structures, which are bluntly dissected to mobilize the upper and lower fascial flaps. At the pubic tubercle, the iliopubic tract is sutured to the lateral edge of the rectus sheath using a synthetic, nonabsorbable, monofilament suture. This continuous suture progresses laterally, approximating the edge of the inferior transversalis fascia flap to the posterior portion of the inguinal canal. The suture is tied to the tail of the original stitch. The next suture begins at the internal inguinal ring, and it continues medially, apposing the aponeuroses of the external oblique and transversus abdominis to the external oblique aponeurotic fibers. At the pubic tubercle, the suture doubles back through the same structures laterally toward the tightened internal ring.

5. Tissue Repairs. Tissue-based herniorrhaphy is a suitable alternative when prosthetic materials cannot be used safely. Indications for tissue repairs include operative field contamination, emergency surgery, and when the viability of hernia contents is uncertain.

6. Bassini Repair The Bassini repair was a historic advancement in operative technique. Its current use is limited as modern techniques reduce recurrence. The original repair includes dissection of the spermatic cord, dissection of the hernia sac with high ligation, and extensive reconstruction of the floor of the inguinal canal (Fig. 37-15). After exposing the inguinal floor, the transversalis fascia is incised from the pubic tubercle to the internal inguinal ring. Preperitoneal fat is bluntly dissected from the upper margin of the posterior side of the transversalis fascia to permit adequate tissue mobilization. A triple-layer repair is then performed. The internal oblique, transversus abdominis, and transversalis fascia are fixed to the shelving edge of the inguinal ligament and pubic periosteum with interrupted sutures. The lateral aspect of the repair reinforces the medial border of the internal inguinal ring.

7. Shouldice Repair The Shouldice repair recapitulates principles of the Bassini repair, and its distribution of tension over several tissue layers results in lower recurrence rates (Fig. 37-16). During dissection of the cord, the genital branch of the genitofemoral nerve is routinely divided, resulting in ipsilateral loss of sensation to the scrotum in men or the mons pubis and labium majus in women. With the posterior inguinal floor exposed, an incision in the transversalis fascia is made between the pubic tubercle and internal ring. Care is taken to avoid injury to preperitoneal structures, which are bluntly dissected to stabilize the upper and lower fascial flaps. At the pubic tubercle, the iliopubic tract is sutured to the lateral edge of the rectus sheath using a synthetic, nonabsorbable, monofilament suture. This continuous suture progresses laterally, approximating the edge of the inferior transversalis fascia flap to the posterior portion of the inguinal canal. The suture is tied to the tail of the original stitch. The next suture begins at the internal inguinal ring, and it continues medially, apposing the aponeuroses of the external oblique and transversus abdominis to the external oblique aponeurotic fibers. At the pubic tubercle, the suture doubles back through the same structures laterally toward the tightened internal ring.

8. McVay Repair The McVay repair addresses both inguinal and femoral ring defects. This technique is indicated for femoral hernias and in cases where the use of prosthetic material is contraindicated (Fig. 37-17). Once the spermatic cord has been isolated, an incision in the transversalis fascia permits entry into the preperitoneal space. The upper flap is mobilized by gentle blunt dissection of underlying tissue. Cooper’s ligament is bluntly dissected to expose its surface. A 2 to 4 cm relaxing incision is made in the anterior rectus sheath vertically from the...
**Figure 37-15.** Bassini repair. A. The transversalis fascia is opened. B. Reconstruction of the posterior wall by suturing the transversalis fascia (TF), the transversus abdominis muscle (TA), and the internal oblique muscle (IO) medially to the inguinal ligament (IL) laterally. EO = external oblique aponeurosis.

**Figure 37-16.** Shouldice repair. A. The iliopubic tract is sutured to the medial flap of the transversalis fascia and the internal oblique and transverse abdominis muscles. B. The second of the four suture lines, reversing toward the pubic tubercle approximating the internal oblique and transversus muscles to the inguinal ligament. Two more suture lines affix the internal oblique and transversus muscles mediadly.

The medial leaf of the external oblique aponeurosis is sutured to the inguinal ligament from the pubic tubercle to the abdominal ring using 1–0 Ethilon or Prolene interrupted sutures. The first two sutures are taken at the junction of the anterior rectus sheath and EOA. The last suture is taken so as to sufficiently close the transversus fascia to the inguinal ligament.

The pubic tubercle. This incision is essential to reduce tension on the repair; however, it may result in increased postoperative pain and higher risk of ventral abdominal herniation. Using either interrupted or continuous suture, the superior transversalis flap is then fastened to Cooper’s ligament, and the repair is continued laterally along Cooper’s ligament to occlude the femoral ring. Lateral to the femoral ring, a transition stitch is placed, affixing the transversalis fascia to the inguinal ligament. The transversalis is then sutured to the inguinal ligament laterally to the internal ring.

**Desarda Repair** The Desarda hernia repair was recently described in 2001, and it consists of a mesh-free repair utilizing a strip of external oblique aponeurosis.

An oblique skin incision is made, and dissection is carried down to the external oblique fascia. The integrity of the fascia is preserved as much as possible. The cremasteric muscle is then incised, and the spermatic cord along with the cremasteric muscle is separated from the inguinal floor. Excision of the sac is done in all cases except in small direct hernias, where it is inverted.
narrow the abdominal ring without constricting the spermatic cord (Fig. 37-18). Each suture is passed first through the inguinal ligament, then the transversalis fascia, and then the EOA. The index finger of the left hand is used to protect the femoral vessels and retract the cord structures laterally while taking lateral sutures. A splitting incision is then taken in the EOA, partially separating a strip. This splitting incision is extended medially up to the pubic symphysis and laterally 1 to 2 cm beyond the reconstructed abdominal ring.

The free border of the strip of the EOA is now sutured to the internal oblique or conjoined tendon lying close to it with 1–0 Ethilon or Prolene interrupted sutures. This is followed by closure of the superficial fascia and the skin as usual.39-41

Prosthetic Repairs. The popularization of tension-free prosthetic mesh repairs signified a paradigm shift in the surgical concept of inguinal hernia pathophysiology. Mesh-based hernioplasty is the most commonly performed general surgical procedure, owing to the technique’s efficacy and improved outcomes. The techniques of the most commonly performed prosthetic repairs are presented in this section.

Lichtenstein Tension-Free Repair The Lichtenstein technique allows for a tension-free repair of the inguinal floor by buttressing the floor with a prosthetic mesh (Fig. 37-18). Initial exposure and mobilization of cord structures is identical to other open approaches. The inguinal canal is dissected to expose the shelving edge of the inguinal ligament, the pubic tubercle, and sufficient area for mesh. The most commonly used mesh is “flat iron” shaped with a keyhole for cord egress, it is available in several sizes. It should be noted that when selecting the size, it must be large enough to extend 2 to 3 cm superior to Hesselbach’s Triangle. The medial edge of the mesh is affixed to the anterior rectus sheath such that it overlaps the pubic tubercle by 1.5 to 2 cm. This refinement to the original Lichtenstein technique minimizes medial recurrence.42

For fixation of the inferior margin of the mesh, a permanent, synthetic, monofilament suture is used taking care to avoid placing sutures directly into the periosteum of the pubic tubercle. Fixation is continued along the shelving edge of the inguinal ligament from medial to lateral, ending at the internal ring. The upper tail of the mesh is then fixed to the internal oblique aponeurosis and the medial edge to the rectus sheath using a synthetic, absorbable suture.

In the case of a femoral hernia, a triangular extension of the inferior aspect of the mesh is sutured to Cooper’s ligament medially and to the inguinal ligament laterally. The lateral tails of the mesh are tailored to fit snugly around the cord at the internal ring, but not too tight to strangulate it. The tails are then sutured to the inguinal ligament with an interrupted stitch and placed beneath the external oblique aponeurosis.

Plug and Patch Technique. A modification of the Lichtenstein repair, the Plug and Patch technique was developed by Gilbert and later popularized by Rutkow and Robbins.43 Prior to placing the prosthetic mesh patch over the inguinal floor, a three-dimensional prosthetic plug is placed in the space previously occupied by the hernia sac (Fig. 37-19). In the case of an indirect hernia, the plug is sutured to Cooper’s ligament and the inguinal ligament with interrupted sutures.44 For direct hernias, the sac is reduced, and the plug is sutured to Cooper’s ligament, the inguinal ligament, and
the internal oblique aponeurosis. While the technique has good overall outcomes, there have been some isolated case report series of complications involving the presence of the plug, including bowel obstruction and chronic pain.

**Wound Closure** Once the reconstruction of the inguinal canal is complete, the cord contents are returned to their anatomic position. The external oblique aponeurosis is then reapproximated continuously from medial to lateral using an absorbable suture. The external ring should be reconstructed in close apposition to the spermatic cord to avoid the appearance of recurrence on future examination. Scarpa’s fascia and skin are appropriately closed.

**Laparoscopic Approach**

Laparoscopic inguinal hernia repairs have become increasingly popular given the noninferiority studies, improved aesthetics, and increased surgeon experience with the procedure. Principal endoscopic methods include the transabdominal preperitoneal (TAPP) repair, the totally extraperitoneal (TEP) repair, and the less-commonly performed intraperitoneal onlay mesh (IPOM) repair.

Of note, awake patients do not tolerate abdominal insufflation well; therefore, laparoscopic repair necessitates the administration of general anesthesia and its inherent risks. Any patient with a contraindication to the use of general anesthesia should not undergo laparoscopic hernia repair. Occasionally, induction of general anesthesia may result in reduction of an incarcerated or strangulated inguinal hernia. If the surgeon suspects this might have occurred, the abdomen should be explored for non-viable tissue either via laparoscopy or upon conversion to an open laparotomy.

The indications for laparoscopic inguinal hernia repair are similar to those for open repair. Most surgeons would agree that the endoscopic approach to bilateral or recurrent inguinal hernias is superior to the open approach. Concurrent inguinal hernia repair can be considered if a hernia patient is scheduled to undergo another laparoscopic procedure without gross contamination, such as prostatectomy. International Endohernia Society (IEHS) guidelines offer a grade A recommendation that TEP and TAPP are preferred alternatives to Lichtenstein repair for recurrent hernias after open anterior repair. The possibility of bilateral repair should be discussed with all patients undergoing endoscopic inguinal hernia surgery.

The operating room configuration is identical for TAPP, TEP, and IPOM procedures. The patient is placed in the Trendelenburg position, and video screens are placed at the foot of the bed. The surgeon stands contralateral to the hernia, and the assistant stands opposite the surgeon. The patient’s arms are tucked to the sides. Figure 37-20 demonstrates a typical operating room setup for endoscopic inguinal hernia repair. The following sections outline the most commonly performed endoscopic inguinal hernia repair techniques.

**Transabdominal Preperitoneal Procedure.** The transabdominal approach confers the advantage of an intraperitoneal perspective, which is useful for bilateral hernias, large hernia defects, and scarring from previous lower abdominal surgery. The abdominal cavity is accessed using a dissecting trocar or open Hasson technique. Pneumoperitoneum to a level of 15 mmHg is achieved. Two 5-mm trocars are placed lateral and slightly inferior to the umbilical trocar, avoiding injury to the inferior epigastric vessels (Fig. 37-21). The patient is then placed in the Trendelenburg position, and the pelvis is inspected.

The bladder, median and medial umbilical ligaments, external iliac, and inferior epigastric vessels are visualized. An incision is made in the peritoneum at the medial umbilical liga ment, 3 to 4 cm superior to the hernia defect, and it is carried laterally to the anterior superior iliac spine. For bilateral inguinal hernia repair, bilateral peritoneal incisions are advisable, leaving a midline bridge of tissue to avoid injuring a potential patent urachus. The inferior edge of incised peritoneum is retracted, and the preperitoneum is dissected to expose the spermatic cord. If a direct hernia is encountered, the sac is inverted and fixed to Cooper’s ligament to prevent development of hematoma or seroma. An indirect hernia sac will usually protrude anterior to the spermatic cord. In this case, the sac is grasped and elevated superiorly from the cord and the space below is developed bluntly to allow for mesh placement. The sac is dissected from its adhesions, and the cord is skeletonized.

The mesh usually measures 10 × 15 cm to completely cover the myopectineal orifice (Fig. 37-22). It is rolled lengthwise and placed through the 12-mm trocar. It is unrolled in the preperitoneal space and secured medially to Cooper’s ligament using an endoscopic tacker. During this fixation, the surgeon palpates the end of the tacker from the abdominal surface to ensure its proper angle and to stabilize the mesh. The mesh is then pulled taut and fixed laterally to the anterior superior iliac spine. Tacks are placed above the iliopubic tract to avoid injury to the lateral cutaneous nerve of the thigh and the femoral branch of the genitofemoral nerve. The peritoneal edges are reapproximated using tacks or intracorporeal sutures as the mesh is stabilized. The peritoneum should be closed completely to avoid contact between the mesh and the intestine. The abdomen is desufflated, and the trocars are removed. The fascial defect of the 12-mm port and the skin incisions are appropriately closed.

**Totally Extraperitoneal Procedure.** The advantage of the TEP repair is the access to the preperitoneal space without intraperitoneal infiltration. Consequently, this approach minimizes the risk of injury to intra-abdominal organs and port site herniation through an iatrogenic defect in the abdominal wall. As with TAPP, TEP is indicated for repair of bilateral inguinal hernias or
for unilateral hernias when scarring makes the anterior approach challenging.

A small horizontal incision is made inferior to the umbilicus. Subcutaneous tissue is dissected to the level of the anterior rectus sheath, which is then incised lateral to the linea alba. The rectus muscle is retracted superolaterally, and a dissecting balloon is advanced through the incision toward the pubic symphysis. Under direct visualization with a 30° laparoscope, the balloon is inflated slowly to bluntly dissect the preperitoneal space (Fig. 37-23). The dissecting balloon is replaced with a 12-mm balloon trocar, and pneumoperitoneum is achieved by insufflation to 15 mmHg. A 5-mm trocar is placed suprapublically in the midline, and another is placed inferior to the insufflation port (see Fig. 37-21). The patient is placed in the Trendelenburg position, and the operation proceeds in an identical fashion to TAPP. No modifications are necessary to repair bilateral inguinal hernias with the TEP approach. Any peritoneal rents should be repaired prior to desufflation to prevent mesh from contacting intraperitoneal structures. Following mesh placement, the preperitoneal space is desufflated slowly under direct vision to ensure proper mesh positioning. Trocars are removed, and the anterior rectus sheath is closed with an interrupted suture. If there

---

**Figure 37-20.** Operating room setup for laparoscopic inguinal hernia repair.
Figure 37-21. Trocar placement for (A) transabdominal preperitoneal repair and (B) totally extraperitoneal repair.

Figure 37-22. View of mesh placement in posterior repairs. A large mesh overlaps the myopectineal orifice.

Figure 37-23. Balloon dissection of the preperitoneal space in a totally extraperitoneal inguinal hernia repair.
is violation of the peritoneum during insufflation of the dissection balloon and subsequent pneumoperitoneum, visualization can be compromised. To address this, a Veress needle or angiocatheter can be placed in the LUQ, which will allow desufflation of the peritoneum and restore visualization.

**Intraperitoneal Onlay Mesh Procedure.** In contrast to TAPP and TEP, the IPOM procedure permits the posterior approach without preperitoneal dissection. It is an attractive procedure in cases where the anterior approach is unfeasible, in recurrent hernias that are refractory to other approaches, or where extensive preperitoneal scarring would make TEP or TAPP challenging. Port placement and inguinal hernia identification are identical to TAPP. Hernia sac contents are reduced; however, the sac itself is not inverted from the preperitoneal space. Instead, mesh is placed directly over the defect and fixed in place with sutures or spiral tacks. Because these anchors are placed through the peritoneum without preperitoneal inspection, the lateral cutaneous nerve of the thigh and the genitofemoral nerve are especially prone to injury. Furthermore, intraperitoneal mesh migration is a documented phenomenon that can lead to postoperative morbidity, recurrence, and reoperation.

**Robot-Assisted Inguinal Hernia Repair**

Application of a robotic platform to hernia repair has been adapted by general surgeons across the country. The endorobot capabilities provide greatly improved manual dexterity and a relatively short learning curve. Though both total extraperitoneal repair and transabdominal preperitoneal repair can be adapted to a robotic platform, the latter has gained more traction among surgeons.

Many papers have explored the efficacy and cost-effectiveness of robot-assisted herniorrhaphy. Retrospective data have had mixed results when comparing robot-assisted surgery vs. laparoscopy. One recent study has shown longer operative time, another analysis has shown increased cost. It should be noted, however, that there is a decrease in cost with robotic surgery as the volume of procedures increases at each center, though it is still unlikely that the costs will ever converge to that of laparoscopic surgery. A retrospective, single-institution study has shown greatly reduced complication rates with robotic assisted surgery in obese patients; however, this was compared against open inguinal hernia repair (10.8% vs. 3.2%, \( P = 0.047 \)), the two groups were covariate matched for preoperative risk. Studies have also shown excellent long-term (36-month) quality of life indicators in robot-assisted TAPP, though this was a single surgeon survey. Further randomized trials will shed more light into cost issues as surgeons gain more experience with robotic application that would lead to shorter operative time and minimize additional instrument use.

Similar to laparoscopy, robot-assisted repair is ideal for recurrent inguinal hernia patients who had previous anterior repair and bilateral hernias. Contraindications to robotic hernia repair are the same as for laparoscopic repair and include coagulopathy and/or severe cardiopulmonary disease precluding induction of general anesthesia and pneumoperitoneum. Previous preperitoneal repair is a relative contraindication along with the presence of a large incarcerated inguinal hernia.

Patient evaluation should proceed similarly to workup for laparoscopic inguinal herniorrhaphy.

**Technique.** Patients are instructed to void in the preoperative area to avoid Foley catheter placement, though some surgeons advocate routine Foley catheter placement. Ideally, the operating table should have capability of synchronization with robotic arms to prevent injury to the patient during repositioning during the procedure. The patient is placed in supine position with arms tucked at both sides. Appropriate padding of extremities is important to avoid neuropraxia and trauma from robotic arm movements. Three trocars are typically used for TAPP repair. Open Hasson technique is employed for initial trocar placement at umbilicus; this can be an 8-mm trocar or alternatively a 12-mm with a telescoped 8-mm trochar. Additionally, two 8-mm trocars are placed in each side of the mid-abdomen, slightly above the level of umbilicus. After trocar placement, the robot is docked and targeted, and the patient is placed in Trendelenburg position. Typically, the surgeon will use robotic shears attached to electrocautery, Cadiere forceps, and a needle holder as the primary instruments. This combination provides optimal cost-effectiveness because the majority of the cost associated with robotic application is due to disposable instruments. Exposure starts with incising the parietal peritoneum from the medial umbilical ligament to the anterior superior iliac spine. A peritoneal flap is developed by blunt and sharp dissection with robotic shears in the preperitoneal space. Special care is taken to leave the preperitoneal fat pad containing nerves and vessels with the anterior abdominal wall. Small vessels can be coagulated with application of electrocautery with scissors. With the aid of pneumoperitoneum, the preference is first to perform a lateral dissection in the space of Bogros. Dissection continues in this plane laterally towards the anterior superior iliac spine. The generous development of a peritoneal flap will ensure successful mesh placement at the end. Then the space of Retzius is entered medially exposing the pubic symphysis. In the absence of haptic feedback, visual recognition of the pubic symphysis is crucial as this serves as an important landmark for further dissection. Inferior epigastric vessels are readily identified. Next, an inferior peritoneal flap is developed to avoid rolling of mesh during closure. Direct, indirect, and femoral spaces are carefully examined. Cadiere forceps are then used to grasp the hernia sac to provide traction. Any cord lipoma is carefully dissected free from the cord structures, and the testicular vessels, pampiniform plexus, and ductus deferens are separated from hernia sac. These structures can usually be identified at the neck of sac. Reduction is successful when the sac stays reduced after traction is released. The next step is placement of the mesh. Lightweight barbed mesh and anatomically preshaped mesh are routinely used. Mesh should be an appropriate size to cover the myopectineal orifice entirely, and the peritoneal dissection will need to be large enough to accommodate this size mesh. It is rolled and placed through one of the ports by the bedside assistant. Then it is unrolled and placed in the pelvis overlapping the pubic symphysis by several centimeters medially; this is essential as the majority of recurrences occur in this area. Utilization of tacking devices are not necessary, which helps to reduce procedural cost; however, this is surgeon preference. Finally, the peritoneal flap is placed back over the mesh layer and sutured back into place with a running locking suture that is facilitated by the increased intracorporeal dexterity of the robotic instruments. Then the fascia of the umbilical trochar site is closed with 0-absorbable suture (Fig 37-24), and the skin is closed with absorbable monofilament suture.

**Prosthesis Considerations**

The success of prosthetic repairs has generated considerable debate about the desirable physical attributes of mesh and their fixation. An ideal mesh should be easy to handle, flexible,
Figure 37-24. Steps in robotic TAPP repair. A. Image of a direct inguinal hernia. B. There is no visible hernia on the contralateral side. C. Hernia contents and sac are dissected and cleared for mesh placement. D. Unrolling and placement of mesh. E. Satisfactory placement of mesh. F. Closure of peritoneum. G. Completed repair of hernia with comparison to contralateral side.
strong, immunologically inert, contraction-resistant, infection-resistant, and inexpensive to manufacture. The following section reviews the most common types of mesh and fixatives currently available.

**Synthetic Mesh Material.** Polypropylene and polyester are the most common synthetic prosthetic materials used in hernia repair. These materials are permanent and hydrophobic, and they promote a local inflammatory response that results in cellular infiltration and scarring with slight contraction in size. Other synthetic mesh materials are under investigation with the goals of minimizing postoperative pain and preventing infection or recurrence. In selecting mesh material, considerations include mesh absorbability, thickness, weight, porosity, and strength.

Variations in the fiber diameter and fiber count of mesh materials categorize them as heavyweight or lightweight in density, though there does not seem to be a universally agreed upon set of criteria for either. Commonly used lightweight mesh materials include β-d-glucan, titanium-coated polypropylene, and polypropylene–poliglecaprone. These materials have greater elasticity and less theoretical surface area contact with surrounding tissues than their heavyweight counterparts. They are hypothesized to reduce scarring and chronic pain without compromising the strength of the repair. The use of lightweight mesh use in TEP and TAPP repairs is associated with fewer 3-month cumulative mesh-related complications. A 2012 meta-analysis of 2310 patients undergoing open or laparoscopic hernia repairs found a lower incidence of chronic pain (relative risk [RR] 0.61, CI 0.50–0.74) following use of lightweight mesh versus heavyweight mesh and no significant difference in rates of recurrence.

When available, lightweight mesh should be considered for all prosthetic repairs to minimize postoperative chronic pain.

A disadvantage of currently available commercial prostheses is their high cost. In settings where resources are limited, prosthetic repairs are performed using alternative materials. Polypropylene and polyethylene mosquito nets are inexpensive and ubiquitous in sub-Saharan Africa and India, and they have similar mechanical properties to commercially available hernioplasty meshes. Meta-analysis of 577 hernioplasties performed using sterilized mosquito nets demonstrated similar rates of short-term mesh-related complications (6.1%) and recurrence (0.17%) to those using commercial meshes. Furthermore, the disability-adjusted life years (DALYs) prevented by inguinal hernia repair signify a comparable impact to that of vaccination in sub-Saharan Africa. Expensive prostheses are not necessarily needed for hernia surgery, whether in resource-limited or in resource-abundant settings, and the anticipated benefits should be evaluated with consideration of increased costs.

**Biologic Mesh.** Although indications for the use of biologic prostheses have not been absolutely defined, they are commonly reserved for contaminated cases or when domain expansion is necessary in the face of high infection risk. This is partially on account of their high cost and high recurrence rates. There are numerous biologic materials available with differing properties, but in general, they have a lower tensile strength and subsequently higher rates of rupture than synthetic prostheses. They also have varying degrees of tensile strength and tissue biocompatibility between them. In ventral hernia repairs, xenograft material was associated with a lower rate of recurrence than allograft material. A review of biologic materials concludes cross-linked graft materials are more durable and less prone to failure than non–cross-linked grafts. Nevertheless, their diminished ability to remodel adversely affects rates of infection and incorporation. While new prosthetic materials continue to be developed, no single biologic warrants routine use. These materials will continue to evolve, and they remain an important tool for challenging cases when used judiciously.

**Fixation Technique.** Independent of prosthesis material, the method of its fixation remains disputed. Suturing, stapling, and tack ing prostheses entail tissue perforation, which may cause inflammation, neurovascular injury, and chronic pain development. Conversely, improper prosthesis fixation may result in mesh migration, repair failure, meshoma pain, and hernia recurrence. Mesh may be fixed with fibrin-derived glue, and self-gripping mesh has been developed to minimize trauma to surrounding tissues and to reduce the risk for entrapment neuropathy. For hernias repaired via a strictly preperitoneal approach, prosthesis fixation may not be necessary at all.

Fibrin glue fixation is a successful alternative to tackle fixation in hernia repair with a synthetic prosthesis. Recent studies comparing fibrin glue fixation and suture fixation in open hernia repair showed superior rates of chronic pain with both Lichtenstein and Plug and Patch techniques. Meta-analyses of endoscopic hernia repair determined the incidence of chronic postoperative pain after tacker fixation was significantly higher than after fibrin glue fixation, with one showing a relative risk of 4.64 (CI 1.9–11.7). Rates of other postoperative complications and recurrence were similar between both fixation methods. Glue fixation is a promising technical refinement, and several studies have shown long-term benefit; however, its questionable efficacy in larger hernias and cost remain considerations.

In TEP repairs, fixation of mesh may not be compulsory. A prospective randomized trial comparing fixation and no fixation in TEP repairs found a significant increase in new pain and equivalent recurrence rates in the fixation group several months after repair. A 2012 meta-analysis comparing laparoscopic tacker mesh fixation to no mesh fixation found no statistically significant differences in operative duration, pain, mesh-related complications, recurrence, or length of stay between the two methods. Studies of three-dimensional, ergonomically contoured mesh without fixation, as well as self-gripping meshes, have yielded similar results. In the preperitoneal approach, the reapproximation of surrounding tissues and physiologic intra-abdominal pressure hypothetically prevent mesh migration. Due to higher theoretical risk of mesh migration, repair without fixation is not recommended for anterior or transperitoneal approaches.

**COMPLICATIONS**

As with other clean operations, the most common complications of inguinal hernia repair include bleeding, infection, seroma, urinary retention, ileus, and injury to adjacent structures (Table 37-6). Complications specific to herniorrhaphy include hernia recurrence, chronic inguinal and pubic pain, and injury to the spermatic cord or testis. The incidence, prevention, and treatment of these complications are discussed in the ensuing section.

**Hernia Recurrence**

When a patient develops pain, bulging, or a mass at the site of an inguinal hernia repair, clinical entities such as seroma, persistent cord lipoma, and hernia recurrence should be considered. Common medical issues associated with recurrence include malnutrition, immunosuppression, diabetes, steroid use, and smoking. Technical causes of recurrence include improper mesh size, tissue
### Complications of groin hernia repairs

<table>
<thead>
<tr>
<th>Recurrence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronic groin pain</td>
</tr>
<tr>
<td>Nociceptive</td>
</tr>
<tr>
<td>Somatic</td>
</tr>
<tr>
<td>Visceral</td>
</tr>
<tr>
<td>Neuropathic</td>
</tr>
<tr>
<td>Iliohypogastric</td>
</tr>
<tr>
<td>Ilioinguinal</td>
</tr>
<tr>
<td>Genitofemoral</td>
</tr>
<tr>
<td>Lateral cutaneous</td>
</tr>
<tr>
<td>Femoral</td>
</tr>
<tr>
<td>Cord and testicular</td>
</tr>
<tr>
<td>Hematoma</td>
</tr>
<tr>
<td>Testicular atrophy</td>
</tr>
<tr>
<td>Dysejaculation</td>
</tr>
<tr>
<td>Division of vas deferens</td>
</tr>
<tr>
<td>Hydrocele</td>
</tr>
<tr>
<td>Testicular descent</td>
</tr>
<tr>
<td>Bladder injury</td>
</tr>
<tr>
<td>Wound infection</td>
</tr>
<tr>
<td>Seroma</td>
</tr>
<tr>
<td>Hematoma</td>
</tr>
<tr>
<td>Wound</td>
</tr>
<tr>
<td>Scrotal</td>
</tr>
<tr>
<td>Retroperitoneal</td>
</tr>
<tr>
<td>Osteitis pubis</td>
</tr>
<tr>
<td>Prosthetic complications</td>
</tr>
<tr>
<td>Contraction</td>
</tr>
<tr>
<td>Erosion</td>
</tr>
<tr>
<td>Infection</td>
</tr>
<tr>
<td>Rejection</td>
</tr>
<tr>
<td>Fracture</td>
</tr>
<tr>
<td>Laparoscopic</td>
</tr>
<tr>
<td>Vascular injury</td>
</tr>
<tr>
<td>Intra-abdominal</td>
</tr>
<tr>
<td>Retroperitoneal</td>
</tr>
<tr>
<td>Abdominal wall</td>
</tr>
<tr>
<td>Gas embolism</td>
</tr>
<tr>
<td>Visceral injury</td>
</tr>
<tr>
<td>Bowel perforation</td>
</tr>
<tr>
<td>Bladder perforation</td>
</tr>
<tr>
<td>Trocar site complications</td>
</tr>
<tr>
<td>Hematoma</td>
</tr>
<tr>
<td>Hernia</td>
</tr>
<tr>
<td>Wound infection</td>
</tr>
<tr>
<td>Keloid</td>
</tr>
<tr>
<td>Bowel obstruction</td>
</tr>
<tr>
<td>Trocar or peritoneal closure site hernia</td>
</tr>
<tr>
<td>Adhesions</td>
</tr>
<tr>
<td>Miscellaneous</td>
</tr>
<tr>
<td>Diaphragmatic dysfunction</td>
</tr>
<tr>
<td>Hypercapnia</td>
</tr>
<tr>
<td>General</td>
</tr>
<tr>
<td>Urinary</td>
</tr>
<tr>
<td>Paralytic ileus</td>
</tr>
<tr>
<td>Nausea and vomiting</td>
</tr>
<tr>
<td>Aspiration pneumonia</td>
</tr>
<tr>
<td>Cardiovascular and respiratory insufficiency</td>
</tr>
</tbody>
</table>

Other chronic pain syndromes include local nerve entrapment, meralgia paresthesia, and osteitis pubis. At greatest ischemia, infection, and tension in the reconstruction. A focused physical examination should be performed. As with primary hernias, US, CT, or MRI can elucidate ambiguous physical findings. When a recurrent hernia is discovered and warrants reoperation, an approach through a virgin plane facilitates its dissection and exposure. Extensive dissection of the scarred field and mesh may result in injury to cord structures, viscera, large blood vessels, and nerves. After an initial anterior approach, the posterior endoscopic approach will usually be easier and more effective than another anterior dissection. Conversely, failed preperitoneal repairs should be approached using an open anterior repair.

### Pain

Pain after inguinal hernia repair is classified into acute or chronic manifestations of three mechanisms: nociceptive (somatic), neuropathic, and visceral pain. Nociceptive pain is the most common of the three. Because it is usually a result of ligamentous or muscular trauma and inflammation, nociceptive pain is reproduced with abdominal muscle contraction. Treatment consists of rest, nonsteroidal anti-inflammatory drugs (NSAIDs), and reassurance as it resolves spontaneously in most cases. Neuropathic pain occurs as a result of direct nerve damage or entrapment. It may present early or late, and it manifests as a localized, sharp, burning, or tearing sensation. It may respond to pharmacologic therapy and to local steroid or anesthetic injections when indicated. Visceral pain refers to pain conveyed through afferent autonomic pain fibers. It is usually poorly localized and may occur during ejaculation as a result of sympathetic plexus injury.

Chronic postoperative pain remains an important measure of clinical outcome that has been reported in as many as 63% of inguinal hernia repair cases. Despite the significant anatomic variation in the three inguinal nerves, literature reviews suggest identification of all three nerves is possible in 70% to 90% of cases. Meticulous nerve identification may prevent injury that results in debilitating chronic postoperative pain syndromes. Notwithstanding, moderate-to-severe pain adversely affects physical activity, social interactions, health care utilization, employment, and productivity in 6% to 8% of patients. Pain in this subset of patients comprises a tremendous individual and societal burden.

Postherniorrhaphy inguinodynia is a debilitating chronic complication. Its incidence is independent of the method of hernia repair. Selective ilioinguinal, iliohypogastric, and genitofemoral nerve injury/myectomy, removal of mesh and fixation material, and revision of the repair are the three most common options for treatment. Nevertheless, anatomic variation and cross-innervation of the inguinal nerves in the retroperitoneum and inguinal canal make selective myectomy less reliable. When inguinodynia is refractory to pharmacologic and interventional measures, triple myectomy with removal of meshoma is routinely performed with acceptable outcomes in the majority of patients. Refractory inguinodynia with concurrent orchialgia also requires resection of the paravasal nerves.

A relatively newly described technique that has cited good outcomes is the laparoscopic triple myectomy. This involves laparoscopic approach to and division of the main trunks of the ilioinguinal and iliohypogastric nerves and additional division of the genitofemoral nerve in the lumbar plexus. Several studies with moderate numbers of patients treated showed durable reduction in pain scores.

Other chronic pain syndromes include local nerve entrapment, meralgia paresthesia, and osteitis pubis.
risk of entrapment are the ilioinguinal and iliohypogastric nerves in anterior repairs and the genitofemoral and lateral femoral cutaneous nerves in endoscopic repairs. Clinical manifestations of nerve entrapment mimic acute neuropathic pain, and they occur with a dermatomal distribution. Injury to the lateral femoral cutaneous nerve results in meralgia paresthetica, a condition characterized by persistent paresthesia of the lateral thigh. Initial treatment of nerve entrapment consists of rest, ice, NSAIDs, physical therapy, and possible local corticosteroid and anesthetic injection. This can be followed by a trial of gabapentin or its analogues. Osteitis pubis is characterized by inflammation of the pubic symphysis and usually presents as medial groin or symphyseal pain that is reproduced by thigh adduction. Avoiding the pubic periosteum when placing sutures and tacks reduces the risk of developing osteitis pubis. CT scan or MRI excludes hernia recurrence, and bone scan is confirmatory for the diagnosis. Initial treatment is identical to that of nerve entrapment; however, if pain remains intractable, orthopedic surgery consultation should be sought for possible bone resection and curettage. Irrespective of treatment, the condition often takes six months to resolve.

**Cord and Testes Injury**

Injury to spermatic cord structures may result in ischemic orchitis or testicular atrophy. Ischemic orchitis is most commonly caused by injury to the pampiniform plexus and not to the testicular artery. It usually manifests within 1 week of inguinal hernia repair as an enlarged, indurated, and painful testis, and it is almost certainly self-limited. It occurs in <1% of primary hernia repairs; however, this figure is larger for recurrent inguinal hernia repairs. US will demonstrate testicular blood flow to differentiate between ischemia and necrosis. Emergent orchiectomy is only necessary in the case of necrosis. Injury to the testicular artery itself may lead to testicular atrophy, which is manifest over a protracted period but does not always lead to testicular necrosis. This is because despite compromise of the artery, there is collateral flow from the inferior epigastric, vesical, prostatic, and scrotal arteries that supply the testes, and in the case of insufficiency, there is atrophy. Treatment for ischemic orchitis most frequently consists of reassurance, NSAIDs, and comfort measures. Intraoperatively, proximal ligation of large hernia sacs to avoid cord manipulation minimizes the risk of injury.

Injury to the vas deferens within the cord may lead to infertility. In open inguinal hernia repairs, isolating the vas deferens along with the cord structures using digital manipulation may cause injury or disruption. In endoscopic approach, grasping the vas may result in a crush injury. Transections of the vas deferens should be addressed with a urologic consult and early anastomosis, if possible. Historically, surgeons and their patients speculated that synthetic material would increase the risks of mesh rejection, carcinogenesis, and inflammation; however, as mesh became used more frequently, these concerns did not manifest. Nevertheless, one study found prosthetic mesh may exert long-term deleterious effects upon the vas deferens, causing azoospermia. Similar studies report varied results, though. A recent prospective study from the Swedish Hernia Registry discovered no difference in rates of patient-reported infertility between the general population and patients who underwent either mesh or tissue-based inguinal hernia repair. Chronic scarring may lead to vas deferens obstruction, resulting in decreased fertility rates and a dysejaculation syndrome. Pain and burning during ejaculation are usually self-limited, and more common causes, such as sexually transmitted diseases, should be excluded.

In females, the round ligament is the analog to the spermatic cord, and it maintains uterine antversion. Injury to the artery of the round ligament does not result in clinically significant morbidity.

**Laparoscopic Complications**

In general, the risks of the TEP technique mirror those of open anterior repairs, as the peritoneal space is not violated. Complications of transabdominal laparoscopy include urinai retention, paralytic ileus, visceral injuries, vascular injuries, and less commonly, bowel obstruction, hypercapnia, gas embolism, and pneumothorax. The most common complications of endoscopic inguinal hernia repair are presented in this section.

**Urinary Retention.** The most common cause of urinary retention after hernia repair is general anesthesia, which is routine in endoscopic hernia repairs. Among 880 patients undergoing inguinal hernia repair with local anesthesia only, 0.2% developed urinary retention, while the rate of urinary retention was 13% among 200 patients undergoing repair with general or spinal anesthesia. Overall, the risk of development of postoperative urinary retention is 2% to 3%. Other risk factors for postoperative urinary retention include pain, narcotic analgesia, and perioperative bladder distention. Initial treatment of urinary retention requires decompression of the bladder with short-term catheterization. Patients will generally require an overnight admission and trial of normal voiding before discharge. Failure to void normally requires reinsertion of the catheter for up to a week. Chronic requirement of a urinary catheter is rare, though older patients may require prolonged catheterization. Risk of urinary retention can be minimized by ensuring voiding prior to surgery and minimization of perioperative fluid administration.

**Ileus and Bowel Obstruction.** The laparoscopic transabdominal approach is associated with a higher incidence of ileus than other modes of repair. This complication is self-limited; however, it necessitates sustained inpatient observation, intravenous fluid maintenance, and possibly nasogastric decompression. Abdominal imaging may be helpful to confirm the diagnosis and to exclude bowel obstruction. Prolonged absence of bowel function, in conjunction with a suspicious abdominal series, should raise concern for obstruction. In this case, CT of the abdomen is helpful to distinguish anatomic sites of obstruction, inflammation, and ischemia. In TAPP repairs, obstruction occurs most commonly secondary to herniation of bowel loops through peritoneal defects or large trocar insertion sites; however, the use of smaller trocars and the preponderance of TEP repairs have reduced the frequency of this complication. True obstruction warrants reoperation.

**Visceral Injury.** Small bowel, colon, and bladder are at risk for injury in laparoscopic hernia repair. The presence of intraabdominal adhesions from previous surgeries may predispose to visceral injuries. Direct bowel injuries may also result from trocar placement. In reoperative abdominal surgery, open Hasson technique and direct visualization of trocars are recommended to reduce the likelihood of visceral injury. Bowel injury may also occur secondary to electrocautery and instrument trauma outside of the camera field. Missed bowel injuries are associated with increased mortality. If injury to the bowel is suspected, its entire length should be examined, and conversion to open repair may be necessary.
Bladder injuries are less common than visceral injuries, and they are usually associated with perioperative bladder distention or extensive dissection of perivesical adhesions. As with bladder injuries encountered in open surgery, cystotomies must be repaired in several layers with 1 to 2 weeks of Foley catheter decompression. A confirmatory cystogram may be performed before catheter removal to confirm healing of the injury.

**Vascular Injury.** The most severe vascular injuries usually occur in iliac or femoral vessels, either by misplaced sutures in anterior repairs, endoscopic tack use, or by trocar injury or direct dissection in laparoscopic repairs. In these cases, exsanguination may be swift. Conversion to an open approach may be necessary, and bleeding should be temporarily controlled with mechanical compression until vascular control is obtained.

The most commonly injured vessels in laparoscopic hernia repair include the inferior epigastrics and external iliac arteries. Although apparent upon initial approach, these vessels may be obscured during mesh positioning, and tacks or staples may injure them. Oftentimes, due to tamponade effect, injury to the inferior epigastric vessels is not apparent until the adjacent trocar is removed. If injured, the inferior epigastrics may be ligated with a percutaneous suture passer or endoscopic vessel clips.

If the tissue pressure exerted by pneumoperitoneum is greater than an injured vessel’s hydrostatic intraluminal pressure, bleeding will not manifest until pneumoperitoneum is released. The presentation of an inferior epigastric vein injury is often delayed because of this effect, and it may result in a significant rectus sheath hematoma. Accordingly, the surgeon should be aware of this intraoperative consideration.

### Hematomas and Seromas

Hematomas may present as localized collections or as diffuse bruising over the operative site. Injury to spermatic cord vessels may result in a scrotal hematoma. Although they are self-limited, characteristic dark blue discoloration of the entire scrotum may alarm patients. Intermittent warm and cold compression aids in resolution. Hematomas may also develop in the incision, retroperitoneum, rectus sheath, and peritoneal cavity. The latter three sites are more frequently associated with laparoscopic repair. Bleeding within the peritoneum or preperitoneal space may not be readily apparent on physical examination. For this reason, close monitoring of subjective complaints, vital signs, urine output, and physical parameters is necessary.

Seromas are fluid collections that most commonly develop within one week of synthetic mesh repairs. Large hernia sac remnants may fill with physiologic fluid and mimic seromas. Patients often mistake seromas for early recurrence. Treatment consists of reassurance and warm compression to accelerate resolution. To avoid secondary infection, seromas should not be aspirated unless they cause discomfort or they restrict activity for a prolonged time.

### OUTCOMES

The incidence of recurrence is the most-cited measure of postoperative outcome following inguinal hernia repair. In evaluating the various available techniques, other salient signifiers of outcome include complication rates, operative duration, hospital stay, and quality of life. The following section summarizes the evidence-based outcomes of the various approaches to inguinal hernia repair.

Among tissue repairs, the Shouldice operation is the most commonly performed technique, and it is most frequently executed at specialized centers. A 2012 meta-analysis from the Cochrane database demonstrated significantly lower rates of hernia recurrence (OR 0.62, CI 0.45–0.85) in patients undergoing Shouldice operations when compared with other open tissue-based methods.96 In experienced hands, the overall recurrence rate for the Shouldice repair is about 1%.97 Although it is an elegant procedure, its meticulous nature requires significant technical expertise to achieve favorable outcomes, and it is associated with longer operative duration and longer hospital stay. One study found the recurrence rate for Shouldice repairs decreased from 9.4% to 2.5% after surgeons performed the repair six times.99 Compared with mesh repairs, the Shouldice technique resulted in significantly higher rates of recurrence (OR 3.65, CI 1.79–7.47); however, it is the most effective tissue-based repair when mesh is unavailable or contraindicated.97

Hernia recurrence is drastically reduced as a result of the Lichtenstein tension-free repair.100 Compared with open elective tissue-based repairs, mesh repair is associated with fewer recurrences (OR 0.37, CI 0.26–0.51) and with shorter hospital stay and faster return to usual activities,101,102 In a multi-institutional series, 3019 inguinal hernias were repaired using the Lichtenstein technique, with an overall recurrence rate of 0.2%,103 Among other tension-free repairs, the Lichtenstein technique remains the most commonly performed procedure worldwide. Meta-analysis demonstrates no significant differences in outcomes between the Lichtenstein and the Plug and Patch techniques; however, intra-abdominal plug migration and erosion into contiguous structures occurs in approximately 6% of cases.101,104,105 The Stoppa technique results in longer operative duration than the Lichtenstein technique. Nevertheless, postoperative acute pain, chronic pain, and recurrence rates are similar between the two methods.106 Perhaps the most compelling advantage of the Lichtenstein technique is that nonexpert surgeons rapidly achieve similar outcomes to their expert counterparts. Guidelines issued by the European Hernia Society recommend the Lichtenstein repair for adults with either unilateral or bilateral inguinal hernias as the preferred open technique.102 Compared to open approaches, endoscopic primary inguinal hernia repair produces equivalent recurrence rates and improved recovery time, pain prevention, and return to normal activities.107 In a study of 168 patients randomized to either TEP or Lichtenstein repair, the 5-year recurrence rates were extremely low in both groups.108,109 Similarly, a study of 200 male patients randomized to either ambulatory TEP or Lichtenstein repair demonstrated no recurrences in either group after one year.110 Because endoscopic surgery requires specialized instruments and longer operative times, its cost is higher than conventional open repair; however, the potential financial benefit of shorter recovery and decreased pain may offset these costs in the long-term.

Perhaps the most salient difference between open and endoscopic techniques is the number of cases needed to develop technical proficiency. In a randomized controlled trial performed by the VA Cooperative Study, two-year recurrence rates were 10.1% in patients undergoing endoscopic repair and 4.9% in those undergoing open repair, and the outcomes of endoscopic repairs improved after each surgeon performed at least 250 cases.111 More recently, Lal and colleagues found that surgeons sustained a decrease from 9% to 2.9% in postoperative recurrences after performing 100 TEP operations.112 Other studies also suggest surgeons develop proficiency in these endoscopic techniques after performing 30
to 100 cases; however, this estimate has decreased precipitously since laparoscopic technique was first introduced.11,13,14

Although controversy persists regarding the utility of TEP versus TAPP, reviews to date find no significant differences in operative duration, length of stay, time to recovery, or short-term recurrence rate between the two approaches. In TAPP repair, the risk of intra-abdominal injury is higher than in TEP repair. This finding prompted the IEHS to recommend TAPP should only be attempted by surgeons with sufficient experience.12 A Cochrane systematic review found rates of port-site hernias and visceral injuries were higher for the TAPP technique, while TEP may be associated with a higher rate of conversion to an alternative approach; however, neither finding was sufficiently compelling to recommend one technique over the other.14

The frequency with which the aforementioned inguinal hernia repair techniques are performed reinforces the importance of broad experience. The authors recommend that surgeons become proficient in several techniques to address different manifestations of inguinal hernias. Surgeons should tailor this experience to optimize outcomes for each patient.

REFERENCES

Entries highlighted in bright blue are key references.

36. Sanchez-Manuel FJ, Lozano-Garcia J, Seco-Gil JL. Antibi-
45. Voyles CR, Hamilton BJ, Johnson WD, Kano N. Meta-


THYROID

Historical Background

Goiters (from the Latin guttur, throat), defined as an enlargement of the thyroid, have been recognized since 2700 B.C. even though the thyroid gland was not documented as such until the Renaissance period. In 1619, Hieronymus Fabricius ab Aquapendente recognized that goiters arose from the thyroid gland. The term thyroid gland (Greek thyreoïdes, shield-shaped) is, however, attributed to Thomas Wharton in his Adenographia (1656). In 1776, the thyroid was classified as a ductless gland by Albrecht von Haller and was thought to have numerous functions ranging from lubrication of the larynx to acting as a reservoir for blood to provide continuous flow to the brain, and to beautifying women’s necks. Burnt seaweed was considered to be the most effective treatment for goiters.

The first accounts of thyroid surgery for the treatment of goiters were given by Roger Frugardi in 1170. In response to failure of medical treatment, two setons were inserted at right angles into the goiter and tightened twice daily until the goiter separated. The open wound was treated with caustic powder and left to heal. However, thyroid surgery continued to be hazardous with prohibitive mortality rates (>40%) until the latter half of the 19th century, when advances in general anesthesia, antisepsis, and hemostasis enabled surgeons to perform thyroid surgery with significantly reduced mortality and morbidity rates. The most notable thyroid surgeons were Emil Theodor Kocher (1841–1917) and C.A. Theodor Billroth (1829–1894), who performed thousands of operations with increasingly successful results. However, as more patients survived thyroid operations, new problems and issues became apparent. After total thyroidectomy, patients (particularly children) became myxedematous with cretinous features. Myxedema was first effectively treated in 1891 by George Murray using a subcutaneous injection of an extract of sheep’s thyroid, and later, Edward Fox demonstrated that oral therapy was equally effective. In 1909, Kocher was awarded the Nobel Prize for medicine in recognition “for his works on the physiology, pathology, and surgery of the thyroid gland.”

Embryology

The thyroid gland arises as an outpouching of the primitive foregut around the third week of gestation. It originates at the floor of the pharyngeal anlage thicken to form the medial thyroid anlage (Fig. 38-1) that descends in the neck anterior to structures that form the hyoid bone and larynx. During its descent, the anlage remains connected to the foramen cecum via an epithelial-lined tube known as the thyroglossal duct. The epithelial cells making up the anlage give rise to the thyroid follicular cells. The paired lateral anlages originate from the fourth branchial pouch and fuse with the median anlage at approximately the fifth week of gestation. The lateral anlages are neuroectodermal in origin (ultimobranchial bodies) and provide the calcitonin producing parafollicular or C cells, which thus come to lie in the superoposterior region of the gland. Thyroid follicles are initially apparent by 8 weeks, and colloid formation begins by the 11th week of gestation.

Developmental Abnormalities

Thyroglossal Duct Cyst and Sinus. Thyroglossal duct cysts are the most commonly encountered congenital cervical anomalies. During the fifth week of gestation, the thyroglossal duct lumen starts to obliterate, and the duct disappears by the eighth week of gestation. Rarely, the thyroglossal duct may persist in whole or in part. Thyroglossal duct cysts may occur anywhere along the migratory path of the thyroid, although 80% are found in juxtaposition to the hyoid bone. They are usually asymptomatic but occasionally become infected by oral bacteria, prompting the patient to seek medical advice. Thyroglossal duct sinuses
result from infection of the cyst secondary to spontaneous or surgical drainage of the cyst and are accompanied by minor inflammation of the surrounding skin. Histologically, thyroglossal duct cysts are lined by pseudostratified ciliated columnar epithelium and squamous epithelium, with heterotopic thyroid tissue present in 20% of cases.

The diagnosis usually is established by observing a 1- to 2-cm, smooth, well-defined midline neck mass that moves upward with protrusion of the tongue. Routine thyroid imaging is not necessary, although thyroid scintigraphy and ultrasound have been performed to document the presence of normal thyroid tissue in the neck. Treatment involves the “Sistrunk operation,” which consists of en bloc cystectomy and excision of the central hyoid bone to minimize recurrence. Approximately 1% of thyroglossal duct cysts are found to contain cancer, which is usually papillary (85%). The role of total thyroidectomy in this setting is debated, but it is advised in patients with large tumors, particularly if there are additional thyroid nodules and evidence of cyst wall invasion or lymph node metastases. Squamous, Hurthle cell, and anaplastic cancers also have been reported but are rare. Medullary thyroid cancers (MTCs) are, however, not found in thyroglossal duct cysts.

**Lingual Thyroid.** A lingual thyroid represents a failure of the median thyroid anlage to descend normally and may be the only thyroid tissue present. Intervention becomes necessary for obstructive symptoms such as choking, dysphagia, airway obstruction, or hemorrhage. Many of these patients develop hypothyroidism. Medical treatment options include administration of exogenous thyroid hormone to suppress thyroid-stimulating hormone (TSH) and radioactive iodine (RAI) ablation followed by hormone replacement. Surgical excision is rarely needed but, if required, should be preceded by an evaluation of normal thyroid tissue in the neck to avoid inadvertently rendering the patient hypothyroid.

**Ectopic Thyroid.** Normal thyroid tissue may be found anywhere in the central neck compartment, including the esophagus, trachea, and anterior mediastinum. Thyroid tissue has been observed adjacent to the aortic arch, in the aortopulmonary window, within the upper pericardium, or in the interventricular septum. Often, “tongues” of thyroid tissue are seen to extend off the inferior poles of the gland and are particularly apparent in large goiters. Thyroid tissue situated lateral to the carotid sheath and jugular vein, previously termed lateral aberrant thyroid, almost always represents metastatic thyroid cancer in lymph nodes, and not remnants of the lateral anlage that had

---

**Figure 38-1.** Thyroid embryology—early development of the median thyroid anlage as a pharyngeal pouch. *(Reproduced with permission from Cady B, Rossi R: Surgery of the Thyroid and Parathyroid Glands. Philadelphia, PA: WB Saunders; 1991.)*
failed to fuse with the main thyroid, as previously suggested by Crile. Even if not readily apparent on physical examination or ultrasound imaging, the ipsilateral thyroid lobe contains a focus of papillary thyroid cancer (PTC), which may be microscopic.

**Pyramidal Lobe.** Normally the thyroglossal duct atrophies, although it may remain as a fibrous band. In about 50% of individuals, the distal end that connects to the thyroid persists as a pyramidal lobe projecting up from the isthmus, lying just to the left or right of the midline. In the normal individual, the pyramidal lobe is not palpable, but in disorders resulting in thyroid hypertrophy (e.g., Graves’ disease, diffuse nodular goiter, or lymphocytic thyroiditis), the pyramidal lobe usually is enlarged and palpable.

**Thyroid Anatomy**
The anatomic relations of the thyroid gland and surrounding structures are depicted in Fig. 38-2. The adult thyroid gland is brown in color and firm in consistency and is located posterior to the strap muscles. The normal thyroid gland weighs approximately 20 g, but gland weight varies with body weight and iodine intake. The thyroid lobes are located adjacent to the thyroid cartilage and connected in the midline by an isthmus that is located just inferior to the cricoid cartilage. A pyramidal lobe is present in about 50% of patients. The thyroid lobes extend to the midthyroid cartilage superiorly and lie adjacent to the carotid sheaths and sternocleidomastoid muscles laterally. The strap muscles (sternohyoid, sternothyroid, and superior belly of the omohyoid) are located anteriorly and are innervated by the ansa cervicalis (ansa hypoglossi). The thyroid gland is enveloped by a loosely connecting fascia that is formed from the partition of the deep cervical fascia into anterior and posterior divisions. The true capsule of the thyroid is a thin, densely adherent fibrous layer that sends out septa that invaginate into the gland, forming pseudolobules. The thyroid capsule is condensed into the posterior suspensory or Berry’s ligament near the cricoid cartilage and upper tracheal rings.

**Figure 38-2.** Anatomy of the thyroid gland and surrounding structures, viewed anteriorly (A) and in cross-section (B). a. = artery; m. = muscle; n. = nerve; v. = vein.
Blood Supply. The superior thyroid arteries arise from the ipsilateral external carotid arteries and divide into anterior and posterior branches at the apices of the thyroid lobes. The inferior thyroid arteries arise from the thyrocervical trunk shortly after their origin from the subclavian arteries. The inferior thyroid arteries travel upward in the neck posterior to the carotid sheath to enter the thyroid lobes at their midpoint. A thyroidea ima artery arises directly from the aorta or innominate in 1% to 4% of individuals to enter the isthmus or replace a missing inferior thyroid artery. The inferior thyroid artery crosses the recurrent laryngeal nerve (RLN), necessitating identification of the RLN before the arterial branches can be ligated. The venous drainage of the thyroid gland occurs via multiple small surface veins, which coalesce to form three sets of veins—the superior, middle, and inferior thyroid veins. The superior thyroid veins run with the superior thyroid arteries bilaterally. The middle vein or veins are the least consistent. The superior and middle veins drain directly into the internal jugular veins. The inferior veins often form a plexus, which drains into the brachiocephalic veins.

Nerves. The left RLN arises from the vagus nerve where it crosses the aortic arch, loops around the ligamentum arteriosum, and ascends medially in the neck within the tracheoesophageal groove. The right RLN arises from the vagus at its crossing with the right subclavian artery. The nerve usually passes posterior to the artery before ascending in the neck, its course being more oblique than the left RLN. Along their course in the neck, the RLNs may branch, and pass anterior, posterior, or interdigitate with branches of the inferior thyroid artery (Fig. 38-3). The right RLN may be nonrecurrent in 0.5% to 1% of individuals and often is associated with a vascular anomaly. Nonrecurrent left RLNs are rare but have been reported in patients with situs inversus and a right-sided aortic arch. The RLN may branch in its course in the neck, and identification of a small nerve should alert the surgeon to this possibility. Identification of the nerves or their branches often necessitates mobilization of the most lateral and posterior extent of the thyroid gland, the tubercle of Zuckerkandl, at the level of the cricoid cartilage. The last segments of the nerves often course below the tubercle and are closely approximated to the ligament of Berry. Branches of the nerve may traverse the ligament in 25% of individuals and are particularly vulnerable to injury at this junction. The RLNs terminate by entering the larynx posterior to the cricothyroid muscle.

The RLNs innervate all the intrinsic muscles of the larynx, except the cricothyroid muscles, which are innervated by the external laryngeal nerves. Injury to one RLN leads to paralysis of the ipsilateral vocal cord, which comes to lie in the paramedian or the abducted position. The paramedian position results in a normal but weak voice, whereas the abducted position leads to a hoarse voice and an ineffective cough. Bilateral RLN injury may lead to airway obstruction, necessitating emergency tracheostomy, or loss of voice. If both cords come to lie in an abducted position, air movement can occur, but the patient has an ineffective cough and is at increased risk of repeated respiratory tract infections from aspiration.

The superior laryngeal nerves also arise from the vagus nerves. After their origin at the base of the skull, these nerves...
travel along the internal carotid artery and divide into two branches at the level of the hyoid bone. The internal branch of the superior laryngeal nerve is sensory to the supraglottic larynx. Injury to this nerve is rare in thyroid surgery, but its occurrence may result in aspiration. The external branch of the superior laryngeal nerve lies on the inferior pharyngeal constrictor muscle and descends alongside the superior thyroid vessels before innervating the cricothyroid muscle. Cernea and colleagues proposed a classification system to describe the relationship of this nerve to the superior thyroid vessels (Fig. 38-4). The type 2a variant, in which the nerve crosses below the tip of the thyroid superior pole, occurs in up to 20% of individuals and places the nerve at a greater risk of injury. Therefore, the superior pole vessels should not be ligated en masse, but should be individually divided, low on the thyroid gland and dissected lateral to the cricothyroid muscle. Injury to this nerve leads to inability to tense the ipsilateral vocal cord and hence difficulty “hitting high notes,” difficulty projecting the voice, and voice fatigue during prolonged speech.

Sympathetic innervation of the thyroid gland is provided by fibers from the superior and middle cervical sympathetic ganglia. The fibers enter the gland with the blood vessels and are vasomotor in action. Parasympathetic fibers are derived from the vagus nerve and reach the gland via branches of the laryngeal nerves.

**Parathyroid Glands.** The embryology and anatomy of the parathyroid glands are discussed in detail in the “Parathyroid Gland” section of this chapter. About 85% of individuals have four parathyroid glands that can be found within 1 cm of the junction of the inferior thyroid artery and the RLN. The superior glands are usually located dorsal to the RLN, whereas the inferior glands are usually found ventral to the RLN (Fig. 38-5).

**Lymphatic System.** The thyroid gland is endowed with an extensive network of lymphatics. Intraglandular lymphatic vessels connect both thyroid lobes through the isthmus and also drain to perithyroidal structures and lymph nodes. Regional lymph nodes include pretracheal, paratracheal, perithyroidal, RLN, superior mediastinal, retropharyngeal, esophageal, and upper, middle, and lower jugular chain nodes. These lymph nodes can be classified into seven levels as depicted in Fig. 38-6. The central compartment includes nodes located in the area between the two carotid sheaths, whereas nodes lateral to the vessels are present in the lateral compartment. Thyroid cancers may metastasize to any of these regions, although metastases to submaxillary nodes (level I) are rare (<1%). There also can be “skip” metastases to nodes in the lateral ipsilateral neck without central neck nodes.

**Thyroid Histology**

Microscopically, the thyroid is divided into lobules that contain 20 to 40 follicles (Fig. 38-7). There are about $3 \times 10^6$ follicles in the adult male thyroid gland. The follicles are spherical and average 30 μm in diameter. Each follicle is lined by cuboidal epithelial cells and contains a central store of colloid secreted from the epithelial cells under the influence of the pituitary hormone TSH. The second group of thyroid secretory cells is the C cells or parafollicular cells, which contain and secrete the hormone calcitonin. They are found as individual cells or clumped in small groups in the interfollicular stroma and located in the upper poles of the thyroid lobes.

**Thyroid Physiology**

**Iodine Metabolism.** The average daily iodine requirement is 0.1 mg, which can be derived from foods such as fish, milk, and eggs or as additives in bread or salt. In the stomach and jejunum, iodine is rapidly converted to iodide and absorbed into the bloodstream, and from there it is distributed uniformly throughout the extracellular space. Iodide is actively transported into the thyroid follicular cells by an adenosine triphosphate (ATP)–dependent process. The thyroid is the storage site of >90% of the body’s iodine content and accounts for one-third of the plasma iodine loss. The remaining plasma iodine is cleared via renal excretion.

**Thyroid Hormone Synthesis, Secretion, and Transport.** The synthesis of thyroid hormone consists of several steps.

---

**Figure 38-4.** Relationship of the external branch of the superior laryngeal nerve and superior thyroid artery originally described by Cernea and colleagues. In type 1 anatomy, the nerve crosses the artery ≥1 cm above the superior aspect of the thyroid lobe. In type 2 anatomy, the nerve crosses the artery <1 cm above the thyroid pole (2a) or below (2b) it. (Reproduced with permission from Bliss RD, Gauger PG, Delbridge LW: Surgeon’s approach to the thyroid gland: surgical anatomy and the importance of technique, World J Surg. 2000 Aug;24(8):891-897.)
SPECIFIC CONSIDERATIONS

PART II

(Fig. 38-8). The first, iodide trapping, involves active (ATP-dependent) transport of iodide across the basement membrane of the thyrocyte via an intrinsic membrane protein, the sodium/iodine (Na+/I-) symporter. Thyroglobulin (Tg) is a large (660 kDa) glycoprotein, which is present in thyroid follicles and has four tyrosyl residues. The second step in thyroid hormone synthesis involves oxidation of iodide to iodine and iodination of tyrosine residues on Tg, to form monoiodotyrosines (MIT) and diiodotyrosines (DIT). Both processes are catalyzed by thyroid peroxidase (TPO). A recently identified protein, pendrin, is thought to mediate iodine efflux at the apical membrane. The third step leads to coupling of two DIT molecules to form tetra-iodothyronine or thyroxine (T₄), and one DIT molecule with one MIT molecule to form 3,5,3′-triiodothyronine (T₃) or 3,3′,5′-triiodothyronine reverse (rT₃). When stimulated by TSH, thyrocytes form pseudopodia, which encircle portions of cell membrane containing Tg, which in turn, fuse with enzyme-containing lysosomes. In the fourth step, Tg is hydrolyzed to release free iodotyrosines (T₃ and T₄) and mono- and diiodotyrosines. The latter are deiodinated in the fifth step to yield iodide, which is reused in the thyrocyte. In the euthyroid state, T₄ is produced and released entirely by the thyroid gland, whereas only 20% of the total T₃ is produced by the thyroid. Most of the T₄ is produced by peripheral deiodination (removal of 5′-iodine from the outer ring) of T₃ in the liver, muscles, kidney, and anterior pituitary, a reaction that is catalyzed by 5′-mono-deiodinase. Some T₃ is converted to rT₃, the metabolically inactive compound, by deiodination of the inner ring of T₃. In conditions such as Graves’ disease, toxic multinodular goiter, or a stimulated thyroid gland, the proportion of T₃ released from the thyroid may be dramatically elevated. Thyroid hormones are transported in serum bound to carrier proteins such as T₄-binding globulin, T₃-binding prealbumin, and albumin. Only a small fraction (0.02%) of thyroid hormone (T₃ and T₄) is free (unbound) and is the physiologically active component. T₃ is the more potent of the two thyroid hormones, although its circulating plasma level is much lower than that of T₄. T₃ is less tightly bound to protein in the plasma than T₄, and so it enters tissues more readily. T₃ is three to four times more active than T₄ per unit weight, with a half-life of about 1 day, compared to approximately 7 days for T₄.

The secretion of thyroid hormone is controlled by the hypothalamic-pituitary-thyroid axis (Fig. 38-9). The hypothalamus produces a peptide, the thyrotropin-releasing hormone (TRH), which stimulates the pituitary to release TSH or thyrotropin. TRH reaches the pituitary via the portovenous circulation. TSH, a 28-kDa glycopeptide, mediates iodide trapping, secretion, and release of thyroid hormones, in addition to increasing the cellularity and vascularity of the thyroid gland. The TSH receptor (TSH-R) belongs to a family of G-protein–coupled receptors that have seven transmembrane-spanning domains and use cyclic adenosine monophosphate in the signal-transduction pathway. TSH secretion by the anterior pituitary is also regulated via a negative feedback loop by T₄ and T₃. Because the pituitary has the ability to convert T₃ to T₄, the latter is thought to be more important in this feedback control. T₃ also inhibits the release of TRH.

The thyroid gland also is capable of autoregulation, which allows it to modify its function independent of TSH. As an adaptation to low iodide intake, the gland preferentially synthesizes
T3 rather than T4, thereby increasing the efficiency of secreted hormone. In situations of iodine excess, iodide transport, peroxide generation, and synthesis and secretion of thyroid hormones are inhibited. Excessively large doses of iodide may lead to initial increased organification, followed by suppression, a phenomenon called the Wolff-Chaikoff effect. Epinephrine and human chorionic gonadotropin hormones stimulate thyroid hormone production. Thus, elevated thyroid hormone levels are found in pregnancy and gynecologic malignancies such as hydatidiform mole. In contrast, glucocorticoids inhibit thyroid hormone production. In severely ill patients, peripheral thyroid hormones may be reduced, without a compensatory increase in TSH levels, giving rise to the euthyroid sick syndrome.

**Thyroid Hormone Function.** Free thyroid hormone enters the cell membrane by diffusion or by specific carriers and is carried to the nuclear membrane by binding to specific proteins. T4 is deiodinated to T3 and enters the nucleus via active transport, where it binds to the thyroid hormone receptor. The T3 receptor is similar to the nuclear receptors for glucocorticoids, mineralocorticoids, estrogens, vitamin D, and retinoic acid. In humans, two types of T3 receptor genes (α and β) are located on chromosomes 3 and 17. Thyroid receptor expression depends on peripheral concentrations of thyroid hormones and is tissue specific—the α form is abundant in the central nervous system, whereas the β form predominates in the liver. Each gene product has a ligand-independent, amino-terminal domain;

---

**Figure 38-6.** A and B. Lymph nodes in the neck can be divided into six regions. Upper mediastinal nodes constitute level VII. m. = muscle; n. = nerve.
The binding of thyroid hormone leads to the transcription and translation of specific hormone-responsive genes.

Thyroid hormones affect almost every system in the body. They are important for fetal brain development and skeletal maturation. T₃ increases oxygen consumption, basal metabolic rate, and heat production by stimulation of Na⁺/K⁺ ATPase in various tissues. It also has positive inotropic and chronotropic effects on the heart by increasing transcription of the Ca²⁺ ATPase in the sarcoplasmic reticulum and increasing levels of β-adrenergic receptors and concentration of G proteins. Myocardial α-receptors are decreased, and actions of catecholamines are amplified. Thyroid hormones are responsible for maintaining the normal hypoxic and hypercapnic drive in the respiratory center of the brain. They also increase gastrointestinal (GI) motility, leading to diarrhea in hyperthyroidism and constipation in hypothyroidism. Thyroid hormones also increase bone and protein turnover.

Figure 38-7. Normal thyroid histology—follicular cells surround colloid.

Figure 38-8. Thyroid follicular cell showing the major signaling pathways involved in thyroid cell growth and function and key steps in thyroid hormone synthesis. The basal membrane of the cell in contact with the circulation and its apical surface contact the thyroid follicle. Thyroid hormone synthesis is initiated by the binding of thyroid-stimulating hormone (TSH) to the TSH receptor (TSHR), a G-protein–coupled transmembrane receptor, on the basal membrane. Activation leads to an increase in cyclic adenosine monophosphate (cAMP), phosphorylation of protein kinase A (PKA), and activation of target cytosolic and nuclear proteins. The protein kinase C (PKC) pathway is stimulated at higher doses of TSH. Iodide is actively transported into the cell via the Na/I symporter (NIS) and flows down an electrical gradient to the apical membrane. There, thyroid peroxidase (TPO) oxidizes iodide and iodinated tyrosyl residues on thyroglobulin (Tg) in the presence of peroxide (H₂O₂). Mono- and diiodotyrosyl (MIT, DIT) residues are also coupled to form T₃ and T₄ by TPO. Thyroglobulin carrying T₃ and T₄ is then internalized by pinocytosis and digested in lysosomes. Thyroid hormone is released into the circulation, while MIT and DIT are deiodinated and recycled. ATP = adenosine triphosphate; CREB = cAMP response element binding protein; CREM = cAMP response element modulator; DAG = diacylglycerol; IGF-1 = insulin-like growth factor 1; IP₃ = inositol-3-phosphate; NADP⁺ = nicotinamide adenine dinucleotide phosphate, oxidized form; NADPH = nicotinamide adenine dinucleotide phosphate; PIP₂ = phosphatidylinositol; PLC = phospholipase C; T₃ = 3,5′,3′-triiodothyronine; T₄ = thyroxine. (Reproduced with permission from Kopp P: Pendred’s syndrome and genetic defects in thyroid hormone synthesis, Rev Endocr Metab Disord. 2000 Jan;1(1-2):109-121.)
and the speed of muscle contraction and relaxation. They also increase glycogenolysis, hepatic gluconeogenesis, intestinal glucose absorption, and cholesterol synthesis and degradation.

Evaluation of Patients With Thyroid Disease

Tests of Thyroid Function. A multitude of different tests are available to evaluate thyroid function. No single test is sufficient to assess thyroid function in all situations, and the results must be interpreted in the context of the patient’s clinical condition. TSH is the only test necessary in most patients with thyroid nodules that clinically appear to be euthyroid.

Serum Thyroid-Stimulating Hormone (Normal 0.5–5 μU/mL) The tests for serum TSH are based on the following principle: monoclonal TSH antibodies are bound to a solid matrix and bind serum TSH. A second monoclonal antibody binds to a separate epitope on TSH and is labeled with radioisotope, enzyme, or fluorescent tag. Therefore, the amount of serum TSH is proportional to the amount of bound secondary antibody (immunometric assay). Serum TSH levels reflect the ability of the anterior pituitary to detect free T4 levels. There is an inverse relationship between the free T4 level and the logarithm of the TSH concentration—small changes in free T4 lead to a large shift in TSH levels. The ultrasensitive TSH assay has become the most sensitive and specific test for the diagnosis of hyper- and hypothyroidism and for optimizing T4 therapy.

Total T4 (Reference Range 55–150 nmol/L) and T3 (Reference Range 1.5–3.5 nmol/L) Total T4 and T3 levels are measured by radioimmunoassay and measure both the free and bound components of the hormones. Total T4 levels reflect the output from the thyroid gland, whereas T3 levels in the nonstimulated thyroid gland are more indicative of peripheral thyroid hormone metabolism, and are, therefore, not generally suitable as a general screening test. Total T4 levels are increased not only in hyperthyroid patients, but also in those with elevated Tg levels secondary to pregnancy, estrogen/progesterone use, or congenital diseases. Similarly, total T4 levels decrease in hypothyroidism and in patients with decreased Tg levels due to anabolic steroid use and protein-losing disorders like nephrotic syndrome. Individuals with these latter disorders may be euthyroid if their free T4 levels are normal. Measurement of total T4 levels is important in clinically hyperthyroid patients with normal T3 levels, who may have T3 thyrotoxicosis. As discussed previously in “Thyroid Hormone Synthesis, Secretion, and Transport,” total T4 levels often are increased in early hypothyroidism.

Free T4 (Reference Range 12–28 pmol/L) and Free T3 (3–9 pmol/L) These radioimmunoassay-based tests are a sensitive and accurate measurement of biologically active thyroid hormone. Free T4 estimates are not performed as a routine screening tool in thyroid disease. Use of this test is confined to cases of early hyperthyroidism in which total T4 levels may be normal but free T4 levels are raised. In patients with end-organ resistance to T4 (Refetoff’s syndrome), T4 levels are increased, but TSH levels usually are normal. Free T3 is most useful in confirming the diagnosis of early hyperthyroidism, in which levels of free T3 and free T4 rise before total T4 and T3. Free T3 levels may also be measured indirectly using the T4-resin uptake test. If free T4 levels are increased, fewer hormone binding sites are available for binding radiolabeled T3 that has been added to the patient’s serum. Therefore, more T3 binds with an ion-exchange resin, and the T4-resin uptake is increased.

Thyrotropin-Releasing Hormone This test is useful to evaluate pituitary TSH secretory function and is performed by administering 500 μg of TRH intravenously and measuring TSH levels after 30 and 60 minutes. In a normal individual, TSH levels should increase at least 6 μIU/mL from the baseline. This test also was previously used to assess patients with borderline hyperthyroidism but has largely been replaced by sensitive TSH assays for this purpose.

Thyroid Antibodies Thyroid antibodies include anti-Tg, antimicrosomal, or anti-TPO and thyroid-stimulating immunoglobulin (TSI). Anti-Tg and anti-TPO antibody levels do not determine thyroid function, but rather indicate the underlying disorder, usually an autoimmune thyroiditis. About 80% of patients with Hashimoto’s thyroiditis have elevated thyroid antibody levels; however, levels may also be increased in patients with Graves’ disease, multinodular goiter, and occasionally, thyroid neoplasms.

Serum Thyroglobulin Tg is only made by normal or abnormal thyroid tissue. It normally is not released into the circulation in large amounts but increases dramatically in destructive processes of the thyroid gland, such as thyroiditis, or overactive states such as Graves’ disease and toxic multinodular goiter. The most important use for serum Tg levels is in monitoring patients with differentiated thyroid cancer for recurrence, particularly after total thyroidectomy and RAI ablation. Elevated anti-Tg antibodies can interfere with the accuracy of serum Tg levels and should always be measured when interpreting Tg levels.
**Serum Calcitonin (0–4 pg/mL Basal)** This 32-amino-acid polypeptide is secreted by the C cells and functions to lower serum calcium levels, although in humans, it has only minimal physiologic effects. It is also a sensitive marker of MTC.

**Thyroid Imaging**

**Radionuclide Imaging** Both iodine-123 (\(^{123}\text{I}\)) and iodine-131 (\(^{131}\text{I}\)) are used to image the thyroid gland. The former emits low-dose radiation, has a half-life of 12 to 14 hours, and is used to image lingual thyroids or goiters. In contrast, \(^{131}\text{I}\) has a half-life of 8 to 10 days and leads to higher-dose radiation exposure. Therefore, this isotope is used to screen and treat patients with differentiated thyroid cancers for metastatic disease. The images obtained by these studies provide information not only about the size and shape of the gland, but also the distribution of functional activity. Areas that trap less radioactivity than the surrounding gland are termed **cold** (Fig. 38-10), whereas areas that demonstrate increased activity are termed **hot**. The risk of malignancy is higher in “cold” lesions (20%) compared to “hot” or “warm” lesions (<5%). Technetium Tc 99m pertechnetate (\(^{99m}\text{Tc}\)) is taken up by the thyroid gland and is increasingly being used for thyroid evaluation. This isotope is taken up by the mitochondria, but is not organified. It also has the advantage of having a shorter half-life and minimizes radiation exposure. It is particularly sensitive for nodal metastases. More recently, \(^{18}\text{F}\)-fluorodeoxyglucose (FDG) positron emission tomography (PET) combined with computed tomography (CT) is being increasingly used to screen for metastases in patients with thyroid cancer in whom other imaging studies are negative. PET scans are not routinely used in the evaluation of thyroid nodules; however, they may show clinically occult thyroid lesions. There are several recent reports of rates of malignancy in these lesions ranging from 14% to 63%. These incidentally discovered nodules should be worked up by ultrasound and fine-needle aspiration biopsy (FNAB).

**Ultrasound** Ultrasound is an excellent noninvasive and portable imaging study of the thyroid gland with the added advantage of no radiation exposure. It is helpful in the evaluation of thyroid nodules, distinguishing solid from cystic ones, and providing information about size and multicentricity. In addition, characteristics such as echotexture, shape, borders and presence of calcifications, and vascularity can provide useful information regarding risk of malignancy. Ultrasound is also especially helpful for assessing cervical lymphadenopathy (Fig. 38-11) and to guide FNAB. An experienced ultrasonographer is necessary for the best results.

**Computed Tomography/Magnetic Resonance Imaging Scan** CT and magnetic resonance imaging (MRI) studies provide excellent imaging of the thyroid gland and adjacent nodes and are particularly useful in evaluating the extent of large, fixed, or substernal goiters (which cannot be evaluated by ultrasound) and their relationship to the airway and vascular structures. Noncontrast CT scans should be obtained for patients who are likely to require subsequent RAI therapy. If contrast is necessary, therapy needs to be delayed by several months. Combined PET-CT scans are increasingly being used for Tg-positive, RAI-negative tumors.

**Benign Thyroid Disorders**

**Hyperthyroidism.** The clinical manifestations of hyperthyroidism result from an excess of circulating thyroid hormone. Hyperthyroidism may arise from a number of conditions that are listed in Table 38-1. It is important to distinguish disorders such as Graves’ disease and toxic nodular goiters that result from increased production of thyroid hormone from those disorders that lead to a release of stored hormone from injury to the thyroid gland (thyroiditis) or from other nonthyroid gland–related conditions. The former disorders lead to an increase in RAI uptake (RAIU), whereas the latter group is characterized by low RAIU. Of these disorders, Graves’ disease, toxic multinodular goiter, and solitary toxic nodule are most relevant to the surgeon.

**Diffuse Toxic Goiter (Graves’ Disease)** Although originally described by the Welsh physician Caleb Parry in a posthumous article in 1825, this disorder is known as Graves’ disease after Robert Graves, an Irish physician who described three patients in 1835. Graves’ disease is by far the most common cause of hyperthyroidism in North America, accounting for 60% to 80% of cases. It is an autoimmune disease with a strong familial predisposition, female preponderance (5:1), and peak incidence between the ages of 40 and 60 years. Graves’ disease is characterized by thyrotoxicosis, diffuse goiter, and extrathyroidal conditions including ophthalmopathy, dermopathy (pretibial myxedema), thyroid acropachy, gynecomastia, and other manifestations.

**Etiology, Pathogenesis, and Pathology.** The exact etiology of the initiation of the autoimmune process in Graves’ disease is not known. However, conditions such as the postpartum state, iodine excess, lithium therapy, and bacterial and viral infections have been suggested as possible triggers. Genetic factors also play a role, as haplotyping studies indicate that Graves’ disease is associated with certain human leukocyte antigen (HLA) haplotypes, including HLA-B8, HLA-DR3, and HLA-DR4*0501 in Caucasian patients, whereas HLA-DRB1*0701 is protective.
against it. Polymorphisms of the cytotoxic T-lymphocyte antigen 4 (CTLA-4) gene also have been associated with Graves’ disease development. CD40 has also been recognized as a Graves’ susceptibility gene. It has an important role in B-cell function and its upregulation leads to a lower threshold for B-cell activation. It can also lead to enhanced IL-6 secretion and activation of T-cells in thyrocytes leading to a local inflammatory response. Other susceptibility genes include PTPN22 (encodes the lymphoid tyrosine phosphatase) and CD25, which encodes for the interleukin-2 receptor α-chain (IL-2Rα). Once initiated, the inflammatory process causes sensitized T-helper lymphocytes to stimulate B lymphocytes, which produce antibodies directed against the thyroid hormone receptor. TSIs or antibodies that stimulate the TSH-R, as well as TSH-binding inhibiting immunoglobulins or antibodies, have been described. The thyroid-stimulating antibodies stimulate the thyrocytes to grow and synthesize excess thyroid hormone, which is a hallmark of Graves’ disease. Graves’ disease also is associated with other autoimmune conditions such as type 1 diabetes mellitus, Addison’s disease, pernicious anemia, and myasthenia gravis.

Macroscopically, the thyroid gland in patients with Graves’ disease is diffusely and smoothly enlarged, with a concomitant increase in vascularity. Microscopically, the gland is hyperplastic, and the epithelium is columnar with minimal colloid present. The nuclei exhibit mitosis, and papillary projections of hyperplastic epithelium are common. There may be aggregates of lymphoid tissue, and vascularity is markedly increased.

**Clinical Features.** The clinical manifestations of Graves’ disease can be divided into those related to hyperthyroidism and those specific to Graves’ disease. Hyperthyroid symptoms include heat intolerance, increased sweating and thirst, and weight loss despite adequate caloric intake. Symptoms of increased adrenergic stimulation include palpitations, nervousness, fatigue, emotional lability, hyperkinesis, and tremors. The most common GI symptoms include increased frequency of bowel movements and diarrhea. Female patients often develop amenorrhea, decreased fertility, and an increased incidence of miscarriages. Children experience rapid growth with early bone maturation, whereas older patients may present with cardiovascular complications such as atrial fibrillation and congestive heart failure.

On physical examination, weight loss and facial flushing may be evident. The skin is warm and moist, and African American patients often note darkening of their skin. Tachycardia or atrial fibrillation is present, with cutaneous vasodilation leading to a widening of the pulse pressure and a rapid falloff in the transmitted pulse wave (collapsing pulse). A fine tremor, muscle wasting, and proximal muscle group weakness with hyperactive tendon reflexes often are present.

Approximately 50% of patients with Graves’ disease also develop clinically evident ophthalmopathy, and dermopathy occurs in 1% to 2% of patients. It is characterized by deposition of glycosaminoglycans, leading to thickened skin in the pretibial region and dorsum of the foot (Fig. 38-12). Eye symptoms include lid lag (von Graefe’s sign), spasm of the upper eyelid revealing the sclera above the corneoscleral limbus (Dalyrple’s sign), and a prominent stare, due to catecholamine excess. True infiltrative eye disease results in periorbital edema, conjunctival swelling and congestion (chemosis), proptosis, limitation of upward and lateral gaze (from involvement of the inferior and medial rectus muscles, respectively), keratitis, and even blindness due to optic nerve involvement. The etiology of Graves’ ophthalmopathy is not completely known; however,

<table>
<thead>
<tr>
<th><strong>Table 38-1</strong></th>
<th><strong>Differential diagnosis of hyperthyroidism</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INCREASED HORMONE SYNTHESIS (INCREASED RAIU)</strong></td>
<td><strong>RELEASE OF PREFORMED HORMONE (DECREASED RAIU)</strong></td>
</tr>
<tr>
<td>Graves’ disease (diffuse toxic goiter)</td>
<td>Thyroiditis—acute phase of Hashimoto’s thyroiditis, subacute thyroiditis</td>
</tr>
<tr>
<td>Toxic multinodular goiter</td>
<td>Factitious (iatrogenic) thyrotoxicosis</td>
</tr>
<tr>
<td>Toxic adenoma</td>
<td>“Hamburger thyrotoxicosis”</td>
</tr>
<tr>
<td>Drug induced—amiodarone, iodine</td>
<td></td>
</tr>
<tr>
<td>Thyroid cancer</td>
<td></td>
</tr>
<tr>
<td>Struma ovarii</td>
<td></td>
</tr>
<tr>
<td>Hydatidiform mole</td>
<td></td>
</tr>
<tr>
<td>TSH-secreting pituitary adenoma</td>
<td></td>
</tr>
</tbody>
</table>

RAIU = radioactive iodine uptake; TSH = thyroid-stimulating hormone.
orbital fibroblasts and muscles are thought to share a common antigen, the TSH-R. Ophthalmopathy is thought to result from inflammation caused by cytokines released from sensitized killer T lymphocytes and cytotoxic antibodies. Gynecomastia is common in young men. Rare bony involvement leads to subperiosteal bone formation and swelling in the metacarpals (thyroid acropachy). Onycholysis, or separation of fingernails from their beds, is a commonly observed finding. On physical examination, the thyroid usually is diffusely and symmetrically enlarged, as evidenced by an enlarged pyramidal lobe. There may be an overlying bruit or thrill over the thyroid gland and a loud venous hum in the supraclavicular space.

**Diagnostic Tests.** The diagnosis of hyperthyroidism is made by a suppressed TSH with or without an elevated free $T_4$ or $T_3$ level. If eye signs are present, other tests are generally not needed. However, in the absence of eye findings, an $^{123}$I uptake and scan should be performed. An elevated uptake, with a diffusely enlarged gland, confirms the diagnosis of Graves’ disease and helps to differentiate it from other causes of hyperthyroidism. Technetium scintigraphy (using pertechnetate, which is trapped by the thyroid, but not organified) can also be used to determine etiology. While technetium scans result in low range of normal uptake and high background activity, total-body radiation exposure is less than that of $^{123}$I scans. If free $T_4$ levels are normal, free $T_3$ levels should be determined, as they often are elevated in early Graves’ disease or toxic nodules ($T_4$ toxicosis). Anti-$T_g$ and anti-TPO antibodies are elevated in up to 75% of patients but are not specific. Elevated TSH-R or thyroid-stimulating antibodies (TSAb) are diagnostic of Graves’ disease and are increased in about 90% of patients. CT or MRI scans of the orbits are useful in evaluating Graves’ ophthalmopathy.

**Treatment.** Graves’ disease may be treated by any of three treatment modalities: antithyroid drugs, thyroid ablation with radioactive $^{131}$I, and thyroidectomy. The choice of treatment depends on several factors, as discussed in the following sections.

**Antithyroid Drugs** Antithyroid medications generally are administered in preparation for RAI ablation or surgery. The drugs commonly used are propylthiouracil (PTU, 100 to 300 mg three times daily) and methimazole (10 to 30 mg three times daily, then once daily). Methimazole has a longer half-life and can be dosed once daily. Both drugs reduce thyroid hormone production by inhibiting the organic binding of iodine and the coupling of iodothyrosines (mediated by TPO). In addition, PTU also inhibits the peripheral conversion of $T_4$ to $T_3$, making it useful for the treatment of thyroid storm. Both drugs can cross the placenta, inhibiting fetal thyroid function, and are excreted in breast milk, although PTU has a lower risk of transplacental transfer. Methimazole also has been associated with congenital aplasia; therefore, PTU is preferred in pregnant and breastfeeding women. Side effects of treatment include reversible granulocytopenia, skin rashes, fever, peripheral neuritis, polyarteritis, vasculitis, hepatitis, and, rarely, agranulocytosis and aplastic anemia. Patients should be monitored for these possible complications and should always be warned to stop PTU or methimazole immediately and seek medical advice should they develop a sore throat or fever. Treatment of agranulocytosis involves admission to the hospital, discontinuation of the drug, and broad-spectrum antibiotic therapy. Surgery should be postponed until the granulocyte count reaches 1000 cells/mm$^3$.

The dose of antithyroid medication is titrated as needed in accordance with TSH and $T_4$ levels. Most patients have improved symptoms in 2 weeks and become euthyroid in about 6 weeks. Some physicians use the block-replace regimen, by adding $T_4$ (0.05 to 0.10 mg) to prevent hypothyroidism and suppress TSH secretion, because some, but not all, studies suggest that this reduces recurrence rates. The length of therapy is debated. Treatment with antithyroid medications is associated
with a high relapse rate when these drugs are discontinued, with 40% to 80% of patients developing recurrent disease after a 1- to 2-year course. Patients with small glands are less likely to recur, so that treatment for curative intent is reserved for patients with (a) small, nontoxic goiters less than 40 g; (b) mildly elevated thyroid hormone levels; (c) negative or low titers of thyroid hormone receptor antibodies; and (d) rapid decrease in gland size with antithyroid medications. The catecholamine response of thyrotoxicosis can be alleviated by administering β-blocking agents. β-Blockade should be considered in all patients with symptomatic thyrotoxicosis and is recommended for elderly patients, those with coexistent cardiac disease, and patients with resting heart rates >90 bpm. These drugs have the added effect of decreasing the peripheral conversion of T4 to T3. Propranolol is the most commonly prescribed medication in doses of about 20 to 40 mg four times daily. Higher doses are sometimes required due to increased clearance of the medication. Caution should be exercised in patients with asthma. Calcium channel blockers are useful for rate control in patients in whom β-blockers are contraindicated.

Radioactive Iodine Therapy (131I) RAI forms the mainstay of Graves’ disease treatment in North America. The major advantages of this treatment are the avoidance of a surgical procedure and its concomitant risks, reduced overall treatment costs, and ease of treatment. Antithyroid drugs are given until the patient is euthyroid and then discontinued to maximize drug uptake. The 131I dose is calculated after a preliminary scan and usually consists of 8 to 12 mCi administered orally. After standard treatment with RAI, most patients become euthyroid within 2 months. However, only about 50% of patients treated with RAI are euthyroid 6 months after treatment, and the remaining are still hyperthyroid or already hypothyroid. After 1 year, about 2.5% of patients develop hypothyroidism each year. RAI also has been documented to lead to progression of Graves’ ophthalmopathy (33% after RAI compared to 16% after surgery), and ophthalmopathy is more common in smokers. Although there is no evidence of long-term problems with infertility, and overall cancer incidence rates are unchanged, there is a small increased risk of nodular goiter, thyroid cancer, and hyperparathyroidism (HPT) in patients who have been treated with RAI. Patients treated with RAI have an unexplained increase in their overall and cardiovascular mortality rates when compared to the general population.

RAI therapy is therefore most often used in older patients with small or moderate-sized goiters, those who have relapsed after medical or surgical therapy, and those in whom antithyroid drugs or surgery are contraindicated. Absolute contraindications to RAI include women who are pregnant (or planning pregnancy within 6 months of treatment) or breastfeeding. Relative contraindications include young patients (i.e., especially children and adolescents), those with thyroid nodules, and those with ophthalmopathy. Lack of access to a high-volume thyroid surgeon is also a consideration. The higher the initial dose of 131I, the earlier the onset and the higher the incidence of hypothyroidism.

Surgical Treatment In North America, surgery is recommended when RAI is contraindicated as in patients who (a) have confirmed cancer or suspicious thyroid nodules, (b) are young, (c) desire to conceive soon (<6 months) after treatment, (d) have had severe reactions to antithyroid medications, (e) have large goiters (>80 g) causing compressive symptoms, and (f) are reluctant to undergo RAI therapy. Relative indications for thyrotoxicosis usually is managed by radioiodine treatment.

Toxic Multinodular Goiter Toxic multinodular goiters usually occur in older individuals, who often have a prior history of a nontoxic multinodular goiter. Over several years, enough thyroid nodules become autonomous to cause hyperthyroidism. The presentation often is insidious in that hyperthyroidism may only become apparent when patients are placed on low doses of thyroid hormone suppression for the goiter. Some patients have T3 toxicity, whereas others may present only with atrial fibrillation or congestive heart failure. Hyperthyroidism also can be precipitated by iodide-containing drugs such as contrast media and the antiarrhythmic agent amiodarone (Jod-Basedow...
hyperthyroidism). Symptoms and signs of hyperthyroidism are similar to Graves’ disease, but extrathyroidal manifestations are absent.

**Diagnostic Studies.** Blood tests are similar to Graves’ disease with a suppressed TSH level and elevated free T<sub>3</sub> or T<sub>4</sub> levels. RAI uptake also is increased, showing multiple nodules with increased uptake and suppression of the remaining gland.

**Treatment.** Hyperthyroidism must be adequately controlled. Both RAI and surgical resection may be used for treatment. When surgery is performed, near-total or total thyroidectomy is recommended to avoid recurrence and the consequent increased complication rates with repeat surgery. Care must be taken in identifying the RLN, which may be found laterally on the thyroid (rather than posterior) or stretched anteriorly over a nodule. RAI therapy is reserved for elderly patients who represent very poor operative risks, provided there is no airway compression from the goiter and thyroid cancer is not a concern. However, because uptake is less than in Graves’ disease, larger doses of RAI often are needed to treat the hyperthyroidism. Furthermore, RAI-induced thyroiditis has the potential to cause swelling and acute airway compromise and leaves the goiter intact, with the possibility of recurrent hyperthyroidism.

**Toxic Adenoma** Hyperthyroidism from a single hyperfunctioning nodule typically occurs in younger patients who note recent growth of a long-standing nodule along with the symptoms of hyperthyroidism. Toxic adenomas are characterized by somatic mutations in the TSH-R gene, although G-protein–stimulating gene (gsp) mutations may occur also. Most hyperfunctioning or autonomous thyroid nodules have attained a size of at least 3 cm before hyperthyroidism occurs. Physical examination usually reveals a solitary thyroid nodule without palpable thyroid tissue on the contralateral side. RAI scanning shows a “hot” nodule with suppression of the rest of the thyroid gland. These nodules are rarely malignant. Smaller nodules may be managed with antithyroid medications and RAI. Larger nodules can require higher doses, which can lead to hypothyroidism. Surgery (lobectomy and isthmusectomy) is preferred to treat young patients and those with larger nodules. Percutaneous ethanol injection (PEI) has been reported to have reasonable success rates but has not been directly compared with surgery.

**Thyroid Storm** Thyroid storm is a condition of hyperthyroidism accompanied by fever, central nervous system agitation or depression, and cardiovascular and GI dysfunction, including hepatic failure. The condition may be precipitated by abrupt cessation of antithyroid medications, infection, thyroid or non-thyroid surgery, and trauma in patients with untreated thyrotoxicosis. Occasionally, thyroid storm may result from amiodarone administration or exposure to iodinated contrast agents or following RAI therapy. This condition was previously associated with high mortality rates but can be appropriately managed in an intensive care unit setting. β-Blockers are given to reduce peripheral T<sub>3</sub> to T<sub>4</sub> conversion and decrease the hyperthyroid symptoms. Oxygen supplementation and hemodynamic support should be instituted. Nonaspirin compounds can be used to treat pyrexia, and Lugol’s iodine or sodium ipodate (intravenously) should be administered to decrease iodine uptake and thyroid hormone secretion. PTU therapy blocks the formation of new thyroid hormone and reduces peripheral conversion of T<sub>4</sub> to T<sub>3</sub>. Corticosteroids often are helpful to prevent adrenal exhaustion and block hepatic thyroid hormone conversion.

<table>
<thead>
<tr>
<th>Table 38–2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Causes of hypothyroidism</strong></td>
</tr>
<tr>
<td>PRIMARY (INCREASED TSH LEVELS)</td>
</tr>
<tr>
<td>Hashimoto’s thyroiditis</td>
</tr>
<tr>
<td>RAI therapy for Graves’ disease</td>
</tr>
<tr>
<td>Postthyroidectomy</td>
</tr>
<tr>
<td>Subacute thyroiditis</td>
</tr>
<tr>
<td>Rare: iodine deficiency, dyshormogenesis</td>
</tr>
</tbody>
</table>

**Hypothyroidism.** Deficiency in circulating levels of thyroid hormone leads to hypothyroidism and, in neonates, to cretinism, which is characterized by neurologic impairment and mental retardation. Hypothyroidism also may occur in Pendred’s syndrome (associated with deafness) and Turner’s syndrome. Conditions that cause hypothyroidism are listed in Table 38–2.

**Clinical Features** Failure of thyroid gland development or function in utero leads to cretinism and characteristic facies similar to those of children with Down syndrome and dwarfism. Failure to thrive and severe mental retardation often are present. Immediate testing and treatment with thyroid hormone at birth can lessen the neurologic and intellectual deficits. Hypothyroidism developing in childhood or adolescence results in delayed development and may also lead to abdominal distention, umbilical hernia, and rectal prolapse. In adults, symptoms in general are nonspecific, including tiredness, weight gain, cold intolerance, constipation, and menorrhagia. Patients with severe hypothyroidism or myxedema develop characteristic facial features due to the deposition of glycosaminoglycans in the subcutaneous tissues, leading to facial and peri orbital puffiness. The skin becomes rough and dry and often develops a yellowish hue from reduced conversion of carotene to vitamin A. Hair becomes dry and brittle, and severe hair loss may occur. There is also a characteristic loss of the outer two thirds of the eyebrows. An enlarged tongue may impair speech, which is already slowed, in keeping with the impairment of mental processes. Patients may also have nonspecific abdominal pain accompanied by distention and constipation. Libido and fertility are impaired in both sexes. Cardiovascular changes in hypothyroidism include bradycardia, cardiomegaly, pericardial effusion, reduced cardiac output, and pulmonary effusions. When hypothyroidism occurs as a result of pituitary failure, other features of hypopituitarism, such as pale, waxy skin; loss of body hair; and atrophic genitalia, may be present.
Laboratory Findings Hypothyroidism is characterized by low circulating levels of T₄ and T₃. Raised TSH levels are found in primary thyroid failure, whereas secondary hypothyroidism is characterized by low TSH levels that do not increase following TRH stimulation. Thyroid autoantibodies are highest in patients with autoimmune disease (Hashimoto’s thyroiditis, Graves’ disease) and may also be elevated in patients with nodular goiter and thyroid neoplasms. An electrocardiogram demonstrates decreased voltage with flattening or inversion of T waves.

Treatment T₄ is the treatment of choice and is administered in dosages varying from 50 to 200 μg per day, depending on the patient’s size and condition. Starting doses of 100 μg of T₄ daily are well tolerated; however, elderly patients and those with coexisting heart disease and profound hypothyroidism should be started on a considerably lower dose such as 25 to 50 μg daily because of associated hypercholesterolemia and atherosclerosis. The dose can be slowly increased over weeks to months to attain a euthyroid state. A baseline electrocardiogram should always be obtained in patients with severe hypothyroidism before treatment. T₄ dosage is titrated against clinical response and TSH levels, which should return to normal. The management of patients with subclinical hypothyroidism (normal T₄, slightly raised TSH) is controversial. Some evidence suggests that patients with subclinical hypothyroidism and increased antithyroid antibody levels should be treated because they will subsequently develop hypothyroidism. Patients who present with myxedema coma require initial emergency treatment with large doses of IV T₄ (300 to 400 μg), with careful monitoring in an intensive care unit setting.

Thyroiditis. Thyroiditis usually is classified into acute, subacute, and chronic forms, each associated with a distinct clinical presentation and histology.

Acute (Suppurative) Thyroiditis The thyroid gland is inherently resistant to infection due to its extensive blood and lymphatic supply, high iodide content, and fibrous capsule. However, infectious agents can seed it (a) via the hematogenous or lymphatic route, (b) via direct spread from persistent pyriform sinus fistulae or thyroglossal duct cysts, (c) as a result of penetrating trauma to the thyroid gland, or (d) due to immunosuppression. Streptococcus and anaerobes account for about 70% of cases; however, other species also have been cultured. Acute suppurative thyroiditis is more common in children and often is preceded by an upper respiratory tract infection or otitis media. It is characterized by severe neck pain radiating to the jaws or ear, fever, chills, odynophagia, and dysphonia. Complications such as systemic sepsis, tracheal or esophageal rupture, jugular vein thrombosis, laryngeal chondritis, and perichondritis or sympathetic trunk paralysis may also occur.

The diagnosis is established by leukocytosis on blood tests and FNAB for Gram’s stain, culture, and cytology. CT scans may help to delineate the extent of infection and identify abscesses. A persistent pyriform sinus fistula should always be suspected in children with recurrent acute thyroiditis. The sensitivity of identification of fistulae in the acute setting is lowest for barium esophagography (50%) and best for direct endoscopy (100%), with CT scans being intermediate (80%). Both barium esophagogram and CT scans have improved sensitivity once the acute inflammation has resolved (100% and 83%, respectively), with CT being better at defining the accurate anatomic pathway and its relationship to the thyroid gland. Treatment consists of parenteral antibiotics and drainage of abscesses. Thyroidectomy may be needed for persistent abscesses or failure of open drainage. Patients with pyriform sinus fistulae require complete resection of the sinus tract, including the area of the thyroid where the tract terminates, to prevent recurrence. Transnasal flexible fiberoptic laryngoscopy is being increasingly used to identify the internal opening of the pyriform sinus tract and may also allow electrocauterization of the tract, and success rates similar to open surgery have been reported.

Subacute Thyroiditis Subacute thyroiditis can occur in the painful or painless forms. Although the exact etiology is not known, painful thyroiditis is thought to be viral in origin or result from a postviral inflammatory response. Genetic predisposition may also play a role, as manifested by its strong association with the HLA-B35 haplotype. One model of pathogenesis suggests that viral or thyroid antigens, when presented by macrophages in the context of HLA-B35, stimulate cytotoxic T lymphocytes and damage thyroid follicular cells.

Painful thyroiditis most commonly occurs in 30- to 40-year-old women and is characterized by the sudden or gradual onset of neck pain, which may radiate toward the mandible or ear. History of a preceding upper respiratory tract infection often can be elicited. The gland is enlarged, exquisitely tender, and firm. The disorder classically progresses through four stages. An initial hyperthyroid phase, due to release of thyroid hormone, is followed by a second, euthyroid phase. The third phase, hypothyroidism, occurs in about 20% to 30% of patients and is followed by resolution and return to the euthyroid state in >90% of patients. A few patients develop recurrent disease.

In the early stages of the disease, TSH is decreased, and Tg, T₄, and T₃ levels are elevated due to the release of preformed thyroid hormone from destroyed follicles. The erythrocyte sedimentation rate is typically >100 mm/h. RA IU also is decreased (<2% at 24 hours), even in euthyroid patients, due to the release of thyroid hormones from destruction of the thyroid parenchyma. Painful thyroiditis is self-limited, and therefore, treatment is primarily symptomatic. Aspirin and other nonsteroidal anti-inflammatory drugs are used for pain relief, but steroids may be indicated in more severe cases. Short-term thyroid replacement may be needed and may shorten the duration of symptoms. Thyroidectomy is reserved for the rare patient who has a prolonged course not responsive to medical measures or for recurrent disease.

Painless thyroiditis is considered to be autoimmune in origin and may occur sporadically or in the postpartum period; the latter typically occurs at about 6 weeks after delivery in women with high TPO antibody titers in early pregnancy. This timing is thought to coincide with a decrease in the normal immune tolerance of pregnancy and consequent rebound elevation of antibody titers.

Painless thyroiditis also is more common in women and usually occurs between 30 and 60 years of age. Physical examination demonstrates a normal sized or minimally enlarged, slightly firm, nontender gland. Laboratory tests and RA IU are similar to those in painful thyroiditis, except for a normal erythrocyte sedimentation rate. The clinical course also parallels painful thyroiditis. Patients with symptoms may require β-blockers and thyroid hormone replacement. Thyroidectomy or RAI ablation is only indicated for the rare patient with recurrent, disabling episodes of thyroiditis.
Chronic Thyroiditis

Lymphocytic (Hashimoto’s) Thyroiditis. Lymphocytic thyroiditis was first described by Hashimoto in 1912 as struma lymphomatosa—a transformation of thyroid tissue to lymphoid tissue. It is the most common inflammatory disorder of the thyroid and the leading cause of hypothyroidism today.

Etiology, Pathogenesis, and Pathology Hashimoto’s thyroiditis is an autoimmune process that is thought to be initiated by the activation of CD4+ T (helper) lymphocytes with specificity for thyroid antigens. Once activated, T cells can recruit cytotoxic CD8+ T cells to the thyroid. Hypothyroidism results not only from the destruction of thyrocytes by cytotoxic T cells but also by autoantibodies, which lead to complement fixation and killing by natural killer cells or block the TSH-R. Antibodies are directed against three main antigens—Tg (60%), TPO (95%), and TSH-R (60%)—and, less commonly, the sodium/iodine symporter (25%). Apoptosis (programmed cell death) also has been implicated in the pathogenesis of Hashimoto’s thyroiditis. Chronic thyroiditis also has been associated with increased intake of iodine and administration of medications such as interferon-α, lithium, and amiodarone. Support for an inherited predisposition includes an increased incidence of thyroid autoantibodies in first-degree relatives of patients with Hashimoto’s thyroiditis compared to controls and the occurrence of the autoantibodies and hypothyroidism in patients with specific chromosomal abnormalities such as Turner’s syndrome and Down syndrome. Associations with HLA-B8, DR3, and DR5 haplotypes of the major histocompatibility complex also have been described. Alterations in CTLA4 have also been shown to increase the risk of developing Hashimoto’s thyroiditis. Other associated genes include various cytokine genes, GITR (glucocorticoid-induced tumor necrosis factor-receptor) and STAT3; however, these need further confirmatory studies.

On gross examination, the thyroid gland is usually mildly enlarged throughout and has a pale, gray-tan cut surface that is granular, nodular, and firm. On microscopic examination, the gland is diffusely infiltrated by small lymphocytes and plasma cells and occasionally shows well-developed germinal centers. Thyroid follicles are smaller than normal with reduced amounts of colloid and increased interstitial connective tissue. The follicles are lined by Hürthle or Askanazy cells, which are characterized by abundant eosinophilic, granular cytoplasm.

Clinical Presentation Hashimoto’s thyroiditis is also more common in women (male-to-female ratio is 1:10 to 20) between the ages of 30 and 50 years old. The most common presentation is that of a minimally or moderately enlarged firm granular gland discovered on routine physical examination or the awareness of a painless anterior neck mass, although 20% of patients present with hypothyroidism, and 5% present with hyperthyroidism (Hashitoxicosis). In classic goitrous Hashimoto’s thyroiditis, physical examination reveals a diffusely enlarged, firm gland, which also is lobulated. An enlarged pyramidal lobe often is palpable.

Diagnostic Studies When Hashimoto’s thyroiditis is suspected clinically, an elevated TSH and the presence of thyroid autoantibodies usually confirm the diagnosis. FNAB with ultrasound guidance is indicated in patients who present with a solitary suspicious nodule or a rapidly enlarging goiter. Thyroid lymphoma is a rare but well-recognized, ominous complication of chronic autoimmune thyroiditis and has a prevalence 80 times higher than expected frequency in this population than in a control population without thyroiditis. Studies of clonal similarity indicate that lymphoma may, in fact, evolve from Hashimoto’s thyroiditis.

Treatment Thyroid hormone replacement therapy is indicated in overtly hypothyroid patients, with a goal of maintaining normal TSH levels. The management of patients with subclinical hypothyroidism (normal T4 and elevated TSH) is controversial. A systematic review of cohort studies showed that in age- and sex-adjusted analyses, subclinical hypothyroidism is associated with a hazard ratio (HR) for coronary heart disease events of 1.89 (95% confidence interval [CI], 1.28 to 2.80; P < .001) and coronary heart disease mortality of 1.58 (95% CI, 1.10 to 2.27; P = .005) for a TSH level of 10 to 19.9 μIU/mL. The data for TSH levels of 5 to 10 μIU/mL was less convincing. An evaluation of the 12 randomized controlled trials in this area only showed a trend toward improvement of some lipid parameters, and none of the included trials evaluated overall mortality or cardiac morbidity. For this reason, levothyroxine is recommended for all patients with TSH levels >10 μIU/mL and patients with levels of 5 to 10 μIU/mL in the presence of a goiter or anti-TPO antibodies. Treatment is also advised especially for middle-aged patients with cardiovascular risk factors such as hyperlipidemia or hypertension and in pregnant patients. Surgery may occasionally be indicated for suspicion of malignancy or for goiters causing compressive symptoms or cosmetic deformity.

Riedel’s Thyroiditis Riedel’s thyroiditis is a rare variant of thyroiditis also known as Riedel’s struma or invasive fibrous thyroiditis that is characterized by the replacement of all or part of the thyroid parenchyma by fibrous tissue, which also invades into adjacent tissues. The etiology of this disorder is controversial, and it has been reported to occur in patients with other autoimmune diseases. This association, coupled with the presence of lymphoid infiltration and response to steroid therapy, suggests a primary autoimmune etiology. Riedel’s thyroiditis is also associated with other focal sclerosing syndromes including mediastinal, retroperitoneal, periportal, and retro-orbital fibrosis and sclerosing cholangitis, suggesting that it may, in fact, be a primary fibrotic disorder. It is now considered a manifestation of IgG4-related systemic disease characterized by elevated serum IgG4 levels and a lymphoplasmycatic infiltrate with an abundance of IgG4 bearing plasma cells. The disease occurs predominantly in women between the ages of 30 and 60 years old. It typically presents as a painless, hard anterior neck mass, which progresses over weeks to years to produce symptoms of compression, including dysphagia, dyspnea, choking, and hoarseness. Patients may present with symptoms of hypothyroidism and hypoparathyroidism as the gland is replaced by fibrous tissue. Physical examination reveals a hard, “woody” thyroid gland with fixation to surrounding tissues. The diagnosis needs to be confirmed by open thyroid biopsy because the firm and fibrous nature of the gland renders FNAB inadequate.

Surgery is the mainstay of the treatment. The chief goal of operation is to decompress the trachea by wedge excision of the thyroid isthmus and to make a tissue diagnosis. More extensive resections are not advised due to the infiltrative nature of the fibrotic process that obscures usual landmarks and structures. Hypothyroid patients are treated with thyroid hormone replacement. Some patients who remain symptomatic have been reported to experience dramatic improvement after treatment with corticosteroids and tamoxifen. Mycophenolate mofetil and more recently rituximab has also been used to attenuate the
inflammatory process and led to dramatic symptom improvements in some patients.15

**Goiter.** Any enlargement of the thyroid gland is referred to as a goiter. The causes of nontoxic goiters are listed in Table 38-3. Goiters may be diffuse, uninodular, or multinodular. Most nontoxic goiters are thought to result from TSH stimulation secondary to inadequate thyroid hormone synthesis and other paracrine growth factors.16 Elevated TSH levels induce diffuse thyroid hyperplasia, followed by focal hyperplasia, resulting in nodules that may or may not concentrate iodine, colloid nodules, or microfollicular nodules. The TSH-dependent nodules progress to become autonomous. Familial goiters resulting from inherited deficiencies in enzymes necessary for thyroid hormone synthesis may be complete or partial. The term endemic goiter refers to the occurrence of a goiter in a significant proportion of individuals in a particular geographic region. In the past, dietary iodine deficiency was the most common cause of endemic goiter. This condition has largely disappeared in North America due to routine use of iodized salt and iodination of fertilizers, animal feeds, and preservatives. However, in areas of iodine deficiency, such as Central Asia, South America, and Indonesia, up to 90% of the population have goiters. Other dietary goitrogens that may participate in endemic goiter formation include kelp, cassava, and cabbage. In many sporadic goiters, no obvious cause can be identified.

**Clinical Features** Most patients with nontoxic goiters are asymptomatic, although patients often complain of a pressure sensation in the neck. As the goiters become very large, compressive symptoms such as dyspnea and dysphagia ensue. Patients also describe having to clear their throats frequently (catarrh). Dysphonia from RLN injury is rare, except when malignancy is present. Obstruction of venous return at the thoracic inlet from a substernal goiter results in a positive Pemberton’s sign—facial flushing and dilatation of cervical veins upon raising the arms above the head (Fig. 38-13A). Sudden enlargement of nodules or cysts due to hemorrhage may cause acute pain. Physical examination may reveal a soft, diffusely enlarged gland (simple goiter) or nodules of various size and consistency in case of a multinodular goiter. Deviation or compression of the trachea may be apparent.

**Diagnostic Tests** Patients usually are euthyroid with normal TSH and low-normal or normal free T4 levels. If some nodules develop autonomy, patients have suppressed TSH levels or become hyperthyroid. RAI uptake often shows patchy uptake with areas of hot and cold nodules. FNAB is recommended in patients who have a dominant nodule or one that is painful or enlarging, as carcinomas have been reported in 5% to 10% of multinodular goiters. CT scans are helpful to evaluate the extent of retrosternal extension and airway compression (Fig. 38-13B).

**Treatment** Most euthyroid patients with small, diffuse goiters do not require treatment. Some physicians give patients with large goiters exogenous thyroid hormone to reduce the TSH stimulation of gland growth; this treatment may result in decrease and/or stabilization of goiter size and is most effective for small diffuse goiters. Endemic goiters are treated by iodine administration. Surgical resection is reserved for goiters that (a) continue to increase despite T4 suppression, (b) cause obstructive symptoms, (c) have substernal extension (considered a relative indication by some groups), (d) have malignancy suspected or proven by FNAB, and (e) are cosmetically unacceptable. Near-total or total thyroidectomy is the treatment of choice, and patients require lifelong T4 therapy.

**Solitary Thyroid Nodule** Solitary thyroid nodules are present in approximately 4% of individuals in the United States, whereas thyroid cancer has a much lower incidence of 40 new cases per 1 million. Therefore, it is of utmost importance to determine which patients with solitary thyroid nodule would benefit from surgery.

**History.** Details regarding the nodule, such as time of onset, change in size, and associated symptoms such as pain, dysphagia, dyspnea, or choking, should be elicited. Pain is an unusual symptom and, when present, should raise suspicion for intrathyroidal hemorrhage in a benign nodule, thyroiditis, or malignancy. Patients with MTC may complain of a dull, aching sensation. A history of hoarseness is worrisome, as it may be secondary to malignant involvement of the RLNs. Most importantly, patients should be questioned regarding risk factors for malignancy, such as exposure to ionizing radiation and family history of thyroid and other malignancies associated with thyroid cancer.

**External-Beam Radiation** Low-dose therapeutic radiation has been used to treat conditions such as tinea capitis (6.5 cGy), thymic enlargement (100 to 400 cGy), enlarged tonsils and adenoids (750 cGy), acne vulgaris (200 to 1500 cGy), and other conditions such as hemangioma and scrofula. Radiation (approximately 4000 cGy) is also an integral part of the management of patients with Hodgkin’s disease. It is now known that a history of exposure to low-dose ionizing radiation to the thyroid gland places the patient at increased risk for developing thyroid cancer. The risk increases linearly from 6.5 to 2000 cGy, beyond which the incidence declines as the radiation causes destruction of the thyroid tissue. The risk is maximum 20 to 30 years after exposure, but these patients require lifelong monitoring. During the nuclear fallout from Chernobyl in 1986, 131I release was accompanied by a marked increase in the incidence of both benign and malignant thyroid lesions noted within 4 years of exposure, particularly in children.17 Most thyroid carcinomas following radiation exposure are papillary, and some of these cancers with a solid type of histology and presence of RET/PTC translocations appear to be more aggressive. In general, there is a 40% chance that patients presenting with a thyroid nodule and a history of radiation have thyroid cancer. Of those patients who have thyroid cancer, the cancer is located in the dominant nodule in 60% of patients, but in the remaining 40% of patients, the cancer is in another nodule in the thyroid gland.

<table>
<thead>
<tr>
<th>CLASSIFICATION</th>
<th>SPECIFIC ETIOLOGY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endemic</td>
<td>Iodine deficiency, dietary goitrogens (cassava, cabbage)</td>
</tr>
<tr>
<td>Medications</td>
<td>Iodide, amiodarone, lithium</td>
</tr>
<tr>
<td>Thyroiditis</td>
<td>Subacute, chronic (Hashimoto’s)</td>
</tr>
<tr>
<td>Familial</td>
<td>Impaired hormone synthesis from enzyme defects</td>
</tr>
<tr>
<td>Neoplasm</td>
<td>Adenoma, carcinoma</td>
</tr>
<tr>
<td>Resistance to thyroid hormone</td>
<td>—</td>
</tr>
</tbody>
</table>
Figure 38-13. A. Retrosternal extension of a large goiter may result in impeded flow in the superior vena cava, leading to dilated veins over the chest wall. This may become more prominent when patients raise their arms above the head—Pemberton’s sign. B. Computed tomography scan demonstrating retrosternal extension and consequent tracheal deviation and compression from a large goiter.
A family history of thyroid cancer is a risk factor for the development of both medullary and nonmedullary thyroid cancer. Familial MTCs occur in isolation or in association with other tumors as part of multiple endocrine neoplasia type 2 (MEN2) syndromes. Nonmedullary thyroid cancers can occur in association with other known familial cancer syndromes such as Cowden’s syndrome, Werner’s syndrome (adult progeroid syndrome), familial adenomatous polyposis, and DICER 1 (Table 38-4). Nonmedullary thyroid cancers can also occur independently of these syndromes as the predominant tumors in the families, and in fact nonsyndromic FNMTC accounts for 95% of cases. The definition of familial nonmedullary thyroid cancer (FNMTMC) is variable across the literature; however, in most studies, it is defined by the presence of two or more first-degree relatives with follicular cell–derived cancers.

**Table 38-4**

<table>
<thead>
<tr>
<th>SYNDROME</th>
<th>GENE</th>
<th>MANIFESTATION</th>
<th>THYROID TUMOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cowden’s syndrome</td>
<td>PTEN</td>
<td>Intestinal hamartomas, benign and malignant breast tumors</td>
<td>FTC, rarely PTC and Hürthle cell tumors</td>
</tr>
<tr>
<td>FAP</td>
<td>APC</td>
<td>Colon polyps and cancer, duodenal neoplasias, desmoids</td>
<td>PTC cribriform growth pattern</td>
</tr>
<tr>
<td>Werner’s syndrome</td>
<td>WRN</td>
<td>Adult progeroid syndrome</td>
<td>PTC, FTC, anaplastic cancer</td>
</tr>
<tr>
<td>Carney complex type 1</td>
<td>PRKAR1α</td>
<td>Cutaneous and cardiac myxomas, breast and adrenal tumors</td>
<td>PTC, FTC</td>
</tr>
<tr>
<td>McCune-Albright syndrome</td>
<td>GNASI</td>
<td>Polystotic fibrous dysplasia, endocrine abnormalities, café-au-lait spots</td>
<td>PTC clear cell</td>
</tr>
<tr>
<td>DICER 1 syndrome</td>
<td></td>
<td>Pleuropulmonary blastoma, cystic nephroma, ovarian sex cord-stromal tumors</td>
<td>Multinodular goiter, thyroid cancer</td>
</tr>
</tbody>
</table>

FAP = familial adenomatous polyposis; FTC = follicular thyroid cancer; PTC = papillary thyroid cancer.

**Family History** A family history of thyroid cancer is a risk factor for the development of both medullary and nonmedullary thyroid cancer. Familial MTCs occur in isolation or in association with other tumors as part of multiple endocrine neoplasia type 2 (MEN2) syndromes. Nonmedullary thyroid cancers can occur in association with other known familial cancer syndromes such as Cowden’s syndrome, Werner’s syndrome (adult progeroid syndrome), familial adenomatous polyposis, and DICER 1 (Table 38-4). Nonmedullary thyroid cancers can also occur independently of these syndromes as the predominant tumors in the families, and in fact nonsyndromic FNMTC accounts for 95% of cases. The definition of familial nonmedullary thyroid cancer (FNMTMC) is variable across the literature; however, in most studies, it is defined by the presence of two or more first-degree relatives with follicular cell–derived cancers.

FNMTMC is now recognized as a distinct clinical entity associated with a high incidence of multifocal tumors and benign thyroid nodules. Some studies report that these patients have higher locoregional recurrence rates and consequently shorter disease-free survival. Several candidate chromosomal loci that predispose to these tumors have been identified, including MNG1 (14q32), thyroid carcinoma with ophthilia (TCO, on 19p13.2), PTPC/papillary renal neoplasia (PRN, on 1q21), NMTC1 (2q21), and FTEN (8p23.1-p22). Susceptibility genes include SRGAP1 (12q14), TITF-1/NKX2.1 (14q13), FOXE1 (9q22), and the telomere-telomerase complex.18

**Physical Examination.** The thyroid gland is best palpated from behind the patient and with the neck in mild extension. The cricoid cartilage is an important landmark, as the isthmus is situated just below it. Nodules that are hard, gritty, or fixed to surrounding structures such as the trachea or strap muscles are more likely to be malignant. The cervical chain of lymph nodes should be assessed as well as the nodes in the posterior triangle.

**Diagnostic Investigations.** An algorithm for the workup of a solitary thyroid nodule is shown in Fig. 38-14.

**Fine-Needle Aspiration Biopsy** FNAB has become the single most important test in the evaluation of thyroid masses and can be performed with or without ultrasound guidance. Ultrasound guidance is recommended for nodules that are difficult to palpate, for cystic or solid-cystic nodules that recur after the initial aspiration, and for multinodular goiters. A 23-gauge needle is inserted into the thyroid mass, and several passes are made while aspirating the syringe. After releasing the suction on the syringe, the needle is withdrawn and the cells are immediately placed on prelabeled dry glass slides; some are immersed in a 70% alcohol solution while others are air dried. A sample of the aspirate is also placed in a 90% alcohol solution for cytospin or cell pellet. The slides are stained by Papanicolaou’s or Wright’s stains and examined under the microscope. If a bloody aspirate is obtained, the patient should be repositioned in a more upright position and the biopsy repeated with a finer (25- to 30-gauge) needle.

After FNAB, the majority of nodules can be classified into several categories that determine further management. To address the issue of variability in the terminology of fine-needle aspiration (FNA), the National Cancer Institute (NCI) hosted the “NCI Thyroid Fine Needle Aspiration State of the Science Conference,” which then defined the Bethesda criteria for thyroid FNA.19 Accordingly, optimum cytology specimens should have at least six follicles each containing at least 10 to 15 cells from at least two aspirates.

The FNA is classified as “nondiagnostic or unsatisfactory” in 2% to 20% of cases and typically results from a virtually acellular specimen, cyst fluid, or the presence of blood or clotting artifact. The risk of malignancy in this setting ranges from 1% to 4%, and reaspiration under ultrasound guidance is recommended. A “benign” result is obtained in 60% to 70% of thyroid FNAs. The most common lesion in this setting is a follicular nodule (includes adenomatoid nodule, colloid nodule, and follicular adenoma). Other diagnoses include lymphocytic (Hashimoto’s) thyroiditis and granulomatous thyroiditis. False-negative results are reported in up to 3% of cases, and follow-up is recommended. A result of “atypia of unknown significance (AUS) or follicular lesion of unknown significance (FLUS)” is obtained in 3% to 6% of biopsies. The risk of malignancy in this scenario is difficult to determine; however, it is thought to be in the range of 5% to 15%. Clinical correlation and a repeat FNA are recommended for AUS lesions (which often results in...
a more definitive interpretation), although clinical observation or surgery may be appropriate because of worrisome clinical or ultrasound findings. The category of “follicular neoplasm” is intended to identify nodules that might be follicular carcinomas. The term suspicious for a follicular neoplasm is preferred by some laboratories for this category because up to 35% of cases turn out not to be neoplasms but hyperplastic proliferations of follicular cells, most commonly those of multinodular goiter. Lobectomy is the preferred treatment for this result, and approximately 15% to 35% of lesions placed in this category prove to be malignant. Hürthle cell neoplasms are also included in this category. Most papillary and other carcinomas can be diagnosed by FNA, but the features are subtle at times, such as in follicular variant of papillary carcinomas. If the diagnosis is uncertain, the lesions are classified as “suspicious for malignancy.” Lobectomy or near-total thyroidectomy is recommended because 60% to 75% turn out to be malignant. This category also includes lesions suspicious for medullary carcinoma and lymphoma, and ancillary testing such as immunohistochemical analysis and flow cytometry may be helpful. The risk of malignancy in lesions classified as “malignant” by FNA is 97% to 99%, and near-total/total thyroidectomy is recommended.

**Laboratory Studies** Most patients with thyroid nodules are euthyroid. Determining the blood TSH level is helpful. If a patient with a nodule is found to be hyperthyroid, the risk of malignancy is approximately 1%. Serum Tg levels cannot differentiate benign from malignant thyroid nodules unless the levels are extremely high, in which case metastatic thyroid cancer should be suspected. Tg levels are, however, useful in following patients who have undergone total thyroidectomy for thyroid cancer and also for serial evaluation of patients undergoing nonoperative management of thyroid nodules. Serum calcitonin levels should be obtained in patients with MTC or a family history of thyroid cancer; FNAB = fine-needle aspiration biopsy; AUS = atypia of unknown significance; FLUS = follicular lesion of unknown significance; FN = follicular neoplasm.

**Imaging** Ultrasound is helpful for detecting nonpalpable thyroid nodules, differentiating solid from cystic nodules, and identifying adjacent lymphadenopathy. Ultrasound evaluation can identify features of a nodule that increase the a priori risk of malignancy, such as fine stippled calcification and enlarged regional nodes; however, a tissue diagnosis is strongly recommended before thyroidectomy. Ultrasound also provides a noninvasive and inexpensive method of following the size of suspected benign nodules diagnosed by FNAB and for identifying enlarged lymph nodes. Ultrasound elastography is used to evaluate tissue stiffness noninvasively. This technique takes advantage of the fact that malignant nodules tend to be harder than benign nodules and thus deform less compared with the surrounding normal thyroid parenchyma. Larger studies are warranted before elastography and newer techniques such as contrast-enhanced ultrasound can be routinely included in the evaluation of thyroid nodules. CT and MRI are unnecessary in the routine evaluation of thyroid tumors except for large, fixed, or substernal lesions. Scanning the thyroid with 123I or 99mTc is rarely necessary, and thyroid scanning currently is recommended in the assessment of thyroid nodules only in patients...
who have follicular thyroid nodules on FNAB and a suppressed TSH. PET scanning does not play a major role in the primary evaluation of thyroid nodules.

**Management.** Malignant tumors are treated by thyroidectomy, as discussed earlier and later in this chapter in “Surgical Treatment under Malignant Thyroid Disease.” Simple thyroid cysts resolve with aspiration in about 75% of cases, although some require a second or third aspiration. If the cyst persists after three attempts at aspiration, unilateral thyroid lobectomy is recommended. Lobectomy also is recommended for cysts >4 cm in diameter or complex cysts with solid and cystic components, as the latter have a higher incidence of malignancy (15%). When FNAB is used in complex nodules, the solid portion should be sampled. If a colloid nodule is diagnosed by FNAB, patients should still be observed with serial ultrasound and Tg measurements. If the nodule enlarges, repeat FNAB often is indicated. Although controversial, levothyroxine in doses sufficient to maintain a serum TSH level between 0.1 and 1.0 μU/mL may also be administered. In areas with a high prevalence of iodine deficiency, this can decrease nodule size and potentially prevent the growth of new nodules. In iodine-sufficient populations, the data are less impressive. Randomized controlled trial analyses have shown that less than 25% of benign nodules shrink more than 50% with TSH suppression in iodine-replete populations.

Thyroidectomy should be performed if a nodule enlarges on TSH suppression, causes compressive symptoms, or for cosmetic reasons. An exception to this general rule is the patient who has had previous irradiation of the thyroid gland or has a family history of thyroid cancer. In these patients, total or near-total thyroidectomy is recommended because of the high incidence of thyroid cancer and decreased reliability of FNAB in this setting.

**Malignant Thyroid Disease**

In the United States, thyroid cancer accounts for <1% of all malignancies (2% of women and 0.5% of men) and is the most rapidly increasing cancer in women. Thyroid cancer is responsible for six deaths per million persons annually. Most patients present with a palpable swelling in the neck, which initiates assessment through a combination of history, physical examination, and FNAB.

**Molecular Genetics of Thyroid Tumorigenesis.** Several oncogenes and tumor suppressor genes are involved in thyroid tumorigenesis,22 as depicted in Table 38-5. The RET proto-oncogene (Fig. 38-15) plays a significant role in the pathogenesis of thyroid cancers. It is located on chromosome 10 and encodes a receptor tyrosine kinase, which binds several growth factors such as glial-derived neurotrophic factor and neurturin. The RET protein is expressed in tissues derived from the embryonic nervous and excretory systems. Therefore,

---

**Table 38-5**

<table>
<thead>
<tr>
<th>GENE</th>
<th>FUNCTION</th>
<th>TUMOR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Oncogenes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RET</td>
<td>Membrane receptor with tyrosine kinase activity</td>
<td>Sporadic and familial MTC, PTC (RET/PTC rearrangements)</td>
</tr>
<tr>
<td>MET</td>
<td>Same</td>
<td>Overexpressed in PTC</td>
</tr>
<tr>
<td>TRK1</td>
<td>Same</td>
<td>Activated in some PTC</td>
</tr>
<tr>
<td>TSH-R</td>
<td>Linked to heterotrimeric G protein</td>
<td>Hyperfunctioning adenoma</td>
</tr>
<tr>
<td>Gsα (gsp)</td>
<td>Signal transduction molecule (GTP binding)</td>
<td>Hyperfunctioning adenoma, follicular adenoma</td>
</tr>
<tr>
<td>Ras</td>
<td>Signal transduction protein</td>
<td>Follicular adenoma and carcinoma, PTC</td>
</tr>
<tr>
<td>PAX8/PPARγ1</td>
<td>Oncoprotein</td>
<td>Follicular adenoma, follicular carcinoma</td>
</tr>
<tr>
<td>B-Raf (BRAF)</td>
<td>Signal transduction</td>
<td>PTC, tall cell and poorly differentiated, anaplastic</td>
</tr>
<tr>
<td>CTNNB1 (β-catenin)</td>
<td>Signal transduction</td>
<td>Upregulated in poorly differentiated and anaplastic cancers</td>
</tr>
<tr>
<td>TERT promoter</td>
<td>Chromosome integrity</td>
<td>Mutated in thyroid cancers including PTC, FTC and anaplastic cancers</td>
</tr>
<tr>
<td><strong>Tumor suppressors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p53</td>
<td>Cell cycle regulator, arrests cells in G1, induces apoptosis</td>
<td>Dedifferentiated PTC, FTC, anaplastic cancers</td>
</tr>
<tr>
<td>p16</td>
<td>Cell cycle regulator, inhibits cyclin-dependent kinase</td>
<td>Thyroid cancer cell lines</td>
</tr>
<tr>
<td>PTEN</td>
<td>Protein tyrosine phosphatase</td>
<td>Follicular adenoma and carcinoma</td>
</tr>
<tr>
<td><strong>Other genetic alterations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>microRNA</td>
<td>Small, noncoding RNA</td>
<td>Specific types upregulated in papillary and some follicular carcinomas</td>
</tr>
</tbody>
</table>

FTC = follicular thyroid cancer; GTP = guanosine triphosphate; MTC = medullary thyroid cancer; PTC = papillary thyroid cancer.
RET disruption can lead to developmental abnormalities in organs derived from these systems, such as the enteric nervous system (Hirschsprung’s disease) and kidney. Germline mutations in the RET proto-oncogene are known to predispose to MEN2A, MEN2B, and familial MTCs, and somatic mutations have been demonstrated in tumors derived from the neural crest, such as MTCs (30%) and pheochromocytomas. The tyrosine kinase domain of RET can fuse with other genes by rearrangement. These fusion products also function as oncogenes and have been implicated in the pathogenesis of PTCs. At least 15 RET/PTC rearrangements have been described and appear to be early events in tumorigenesis. Young age and radiation exposure seem to be independent risk factors for the development of RET/PTC rearrangements. Up to 70% of papillary cancers in children exposed to the radiation fallout from the 1986 Chernobyl disaster carry RET/PTC rearrangements, the most common being RET/PTC1 and RET/PTC3. These rearrangements confer constitutive activation of the receptor tyrosine kinases. RET/PTC3 is associated with a solid type of PTC that appears to present at a higher stage and to be more aggressive. It has now been established that RET/PTC3 signaling involves the mitogen-activated protein kinase (MAPK) pathway via other signaling molecules such as Ras, Raf, and MEK. In normal cells, physiological activation of Raf kinases occurs via direct interaction with guanosine triphosphate (GTP)-bound Ras, a membrane-bound small G protein. Activated Raf, a serine-threonine kinase, in turn phosphorylates MEK, another serine-threonine kinase. This leads to phosphorylation of ERK/MAPK, which phosphorylates regulatory molecules in the nucleus, thereby altering gene expression. Aberrant activation of the MAPK pathway leads to tumorigenesis. Aside from RET/PTC alterations, mutations in the Ras genes can also activate the MAPK pathway. Mutated RAS oncogenes have been identified in up to 20% to 40% of thyroid follicular adenomas and carcinomas, multinodular goiters, and papillary and anaplastic carcinomas. There are three Raf kinases, A-Raf, B-Raf (BRAF), and C-Raf. Mutations in BRAF also have been implicated in aberrant MAPK pathway activation and tumorigenesis. Of the various identified BRAF mutations, T1799A (V600E amino acid substitution) is the most common and occurs frequently in thyroid cancers. Interestingly, BRAF mutations occur in papillary and anaplastic tumors (average prevalence of 44% and 22%, respectively) but not in follicular thyroid cancers, suggesting a role in the pathogenesis of these malignancies. Studies also show that BRAF mutations are associated with more aggressive clinicopathologic features, including larger tumor size, invasion, and lymphadenopathy, and may have a role as prognostic markers.

The p53 gene is a tumor suppressor gene encoding a transcriptional regulator, which causes cell cycle arrest allowing repair of damaged DNA, thus helping to maintain genomic integrity. Mutations of p53 are rare in PTCs but common in undifferentiated thyroid cancers and thyroid cancer cell lines. Other cell cycle regulators and tumor suppressors such as p15 and p16 are mutated more commonly in thyroid cancer cell lines than in primary tumors. An oncogene resulting from the fusion of the DNA binding domain of the thyroid-transcription factor PAX8 gene to the peroxisome proliferator-activated receptor gamma 1 (PPARγ1) has been noted to play an important role
in the development of follicular neoplasms, including follicular cancers. Mutations in the telomerase reverse transcriptase catalytic subunit (TERT) promoter unit have also been recently been identified in well-differentiated thyroid cancers and appear to be related to poor prognosis. Thyroid cancer stem cells have also been identified; however, their role in thyroid carcinogenesis remains to be determined. Mutations in the kinases PIK3CA and AKT1 are rare in thyroid cancers and tend to occur as late events in tumorigenesis.

Specific Tumor Types

Papillary Carcinoma Papillary carcinoma accounts for 80% of all thyroid malignancies in iodine-sufficient areas and is the predominant thyroid cancer in children and individuals exposed to external radiation. Papillary carcinoma occurs more often in women, with a 2:1 female-to-male ratio, and the mean age at presentation is 30 to 40 years. Most patients are euthyroid and present with a slow-growing painless mass in the neck. Dysphagia, dyspnea, and dysphonnia usually are associated with locally advanced invasive disease. Lymph node metastases are common, especially in children and young adults, and may be the presenting complaint. “Lateral aberrant thyroid” almost always denotes a cervical lymph node that has been invaded by metastatic cancer. Suspicion of thyroid cancer often originates through physical examination of patients and a review of their history. Diagnosis is established by FNAB of the thyroid mass or lymph node. Once thyroid cancer is diagnosed on FNAB, a complete neck ultrasound is strongly recommended to evaluate the contralateral lobe and for lymph node metastases in the central and lateral neck compartments. Distant metastases are uncommon at initial presentation, but may ultimately develop in up to 20% of patients. The most common sites are lungs, pleura, liver, and brain.

Pathology. On gross examination, PTCs generally are hard and whitish and remain flat on sectioning with a blade, in contrast to normal tissue or benign nodular lesions that tend to bulge. Macroscopic calcification, necrosis, or cystic change may be apparent. Histologically, papillary carcinomas may exhibit papillary projections (Fig. 38-16A), a mixed pattern of papillary and follicular structures, or a pure follicular pattern (follicular variant). The diagnosis is established by characteristic nuclear cellular features. Cells are cuboidal with pale, abundant cytoplasm, crowded nuclei that may demonstrate “grooving,” and intranuclear cytoplasmic inclusions (leading to the designation of Orphan Annie nuclei [Fig. 38-16B]), which allow diagnosis by FNAB. Psammoma bodies, which are microscopic, calcified deposits representing clumps of sloughed cells, also may be present. Mixed papillary-follicular tumors and follicular variant of papillary thyroid carcinoma (FVPTC) are classified as papillary carcinomas because they behave biologically as papillary carcinomas. Two main subtypes of FVPTC are recognized: encapsulated and nonencapsulated (infiltrative). The former is challenging to diagnose. Since the tumors have no invasion, the diagnosis relies on the finding of characteristic nuclei, which can be subjective. In addition, several studies have shown that the encapsulated tumors have an indolent behavior and are genetically distinct from their infiltrative counterparts. As such, these tumors are now designated noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP).

Multifocality is common in papillary carcinoma and may be present in up to 85% of cases on microscopic examination. Multifocality is associated with an increased risk of cervical nodal metastases, and these tumors may rarely invade adjacent structures such as the trachea, esophagus, and RLNs. Other variants of papillary carcinoma include tall cell, insular, columnar, diffuse sclerosing, clear cell, trabecular, and poorly differentiated types. These variants account for about 1% of all papillary carcinomas and are generally associated with a worse prognosis.

Minimal or occult/microcarcinoma refers to tumors of 1 cm or less in size with no evidence of local invasiveness through the thyroid capsule or angioinvasion, and that are not associated with lymph node metastases. They are nonpalpable and usually are incidental findings at operative, histologic, or autopsy examination. Studies have demonstrated occult PTC to be present in 2% to 36% of thyroid glands removed at autopsy. These tumors are also being identified more frequently due to the widespread use of ultrasound. These occult tumors are generally associated with a better prognosis than larger tumors, but they may be more aggressive than previously appreciated. About 25% of patients with these tumors have associated occult lymph node metastases.

Prognostic Indicators. In general, patients with PTC have an excellent prognosis with a >95% 10-year survival rate. Several prognostic indicators have been incorporated into various
staging systems, which enable patients to be stratified into low-risk and high-risk groups. Unfortunately, all of these classification systems rely on data that are not available preoperatively.

In 1987, Hay and colleagues at the Mayo Clinic proposed the AGES scoring system, which incorporates Age, histologic Grade, Extrathyroidal invasion, and metastases and tumor Size to predict the risk of dying from papillary cancer. Low-risk patients are young, with well-differentiated tumors, no metastases, and small primary lesions, whereas high-risk patients are older, with poorly differentiated tumors, local invasion, distant metastases, and large primary lesions. The MACIS scale is a postoperative system modified from the AGES scale. This scale incorporates distant Metastases, Age at presentation (<40 or >40 years old), Completeness of original surgical resection, extrathyroidal Invasion, and Size of original lesion (in centimeters) and classifies patients into four risk groups based on their scores. Cady proposed the AMES system to classify differentiated thyroid tumors into low- and high-risk groups using Age (men <40 years old, women <50 years old), Metastases, Extrathyroidal spread, and Size of tumors (< or >5 cm). A simplified system by DeGroot and associates uses four groups—class I (intradathroidal), class II (cervical nodal metastases), class III (extrathyroidal invasion), and class IV (distant metastases)—to determine prognosis. Another classification system is the TNM system (Tumor, Nodal status, Metastases; Table 38-6), which used by most medical centers in North America and has been recently updated. In this new version, minimal extrathyroidal extension is no longer considered T3a disease. Thyroglobulin doubling time (using levels obtained when TSH is <0.1 mIU/L) has also been demonstrated to be an independent prognostic marker for metastatic disease and recurrence.

Several molecular and genetic markers such as tumor DNA aneuploidy, decreased cyclic adenosine monophosphate response to TSH, increased epidermal growth factor binding, presence of N-ras and gsp mutations, overexpression of c-myc, and presence of p53 mutations also have been associated with a worse prognosis. The presence of BRAF V600E mutation, as previously mentioned, is associated with aggressive tumor characteristics, including extrathyroidal extension, older age at presentation, and lymph node and distant metastases. This mutation also appears to be an independent predictor of both tumor recurrence (even for early-stage disease) and tumor-related mortality. Some studies propose that BRAF mutation status on FNAB can be used to tailor initial management including more extensive initial surgical excision, high-dose postoperative RAI therapy, increased TSH suppression, and closer follow-up. The correlation of RET/PTC rearrangements and Ras mutations with prognosis is less clear. TERT promoter mutations have been associated with poor disease-specific and disease-free survival.

Surgical Treatment. Most authors agree that patients with high-risk tumors (judged by any of the classification systems discussed earlier in “Prognostic Indicators”) or bilateral tumors should undergo total or near-total thyroidectomy. The optimal surgical strategy in the majority of patients with low-risk (small, unilateral) cancers was controversial for many years, with the focus of the debate centering around outcome data and risks associated with extent of thyroidectomy in this group of patients. Proponents of total thyroidectomy indicate that it enables the use of RAI to effectively detect and treat residual thyroid tissue or metastatic disease and makes serum Tg level a more sensitive marker of recurrent or persistent disease. It is also known that a significant proportion (33% to 50%) of patients who develop a recurrence die from their disease, and even though the data are retrospective, long-term, follow-up studies suggest that recurrence rates are lowered and that survival is improved in patients undergoing near-total or total thyroidectomy. In addition, diminished survival is noted in patients with low-risk disease (mortality rates of 5% at 10 to 20 years), and it is not possible to accurately risk stratify patients preoperatively. In the last 10 years, a large study of >50,000 patients with papillary cancer demonstrated that, in multivariate analyses, total thyroidectomy led to a significantly improved recurrence and survival for tumors >1 cm. Furthermore, the authors also showed that patients with tumors 1 to 2 cm in diameter who were treated with lobectomy had a 24% higher risk of recurrence and a 49% higher risk of thyroid cancer mortality. Based on this information, the American Thyroid Association 2009 guidelines for the evidence-based management of thyroid cancers recommended a near-total or total thyroidectomy for primary cancers >1 cm unless there are contraindications to the surgery. However, additional studies since then have demonstrated no survival differences based on initial surgical procedure when adjusting for complexity/risk and comorbid diseases. This finding, coupled with a trend for increased use of ultrasound and Tg measurements to assess for recurrences and the declining use of RAI ablation, led to revised guidelines in 2015. Accordingly, either near-total(total thyroidectomy or lobectomy constitute appropriate initial treatment for tumors >1 cm and <4 cm without extrathyroidal extension or lymph node involvement (cN0). Of note, the guidelines do state that the treatment team may elect near-total/total thyroidectomy to facilitate RAI therapy, enhance follow-up based on disease features, or if the patient expresses a preference for complete thyroid excision.

There has also been in change in the management of papillary microcarcinomas (<1 cm) since at least two trials from Japan have shown that active surveillance (defined as observation without immediate surgery) can be a viable and safe first line of treatment for these very-low-risk tumors without extrathyroidal extension or lymph node metastases. Tumors that progress during monitoring are treated by surgery. If surgery is chosen as initial treatment for these patients, a thyroid lobectomy is considered sufficient.

When PTC is diagnosed by FNAB, the definitive operation can be done without confirming the diagnosis by frozen section during the operation. Patients with a nodule that is suspicious for papillary cancer should be treated by thyroid lobectomy, isthmectomy, and removal of any pyramidal lobe or adjacent lymph nodes. If intraoperative frozen-section examination of a lymph node or primary tumor confirms carcinoma, completion total or near-total thyroidectomy is performed. If a definitive diagnosis cannot be made or the surgeon is concerned about the viability of the parathyroid glands or the status of the RLN, the operation is terminated. When final histology confirms carcinoma, completion thyroidectomy is performed if deemed necessary based on risk-stratification. During thyroidectomy, enlarged or obviously involved central neck nodes should be removed (therapeutic central-compartment, level VI), along with nodes with known lateral neck metastases. Some investigators recommend routine bilateral central neck dissection due to the high incidence of microscopic metastases and data showing improved rates of recurrence and survival (compared to historic controls).
## Table 38-6

### TNM classification of thyroid tumors PAPILLARY or FOLLICULAR TUMORS

<table>
<thead>
<tr>
<th>DIAGNOSIS AGE</th>
<th>TUMOR</th>
<th>NODE</th>
<th>METASTASIS</th>
<th>STAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;55 years</td>
<td>Any T</td>
<td>Any N</td>
<td>M0</td>
<td>I</td>
</tr>
<tr>
<td>&lt;55 years</td>
<td>Any T</td>
<td>Any N</td>
<td>M1</td>
<td>II</td>
</tr>
<tr>
<td>≥55 years</td>
<td>T1</td>
<td>N0/NX</td>
<td>M0</td>
<td>I</td>
</tr>
<tr>
<td>≥55 years</td>
<td>T1</td>
<td>N1</td>
<td>M0</td>
<td>II</td>
</tr>
<tr>
<td>≥55 years</td>
<td>T2</td>
<td>N0/NX</td>
<td>M0</td>
<td>I</td>
</tr>
<tr>
<td>≥55 years</td>
<td>T2</td>
<td>N1</td>
<td>M0</td>
<td>II</td>
</tr>
<tr>
<td>≥55 years</td>
<td>T3a/T3b</td>
<td>Any N</td>
<td>M0</td>
<td>II</td>
</tr>
<tr>
<td>≥55 years</td>
<td>T4a</td>
<td>Any N</td>
<td>M0</td>
<td>II</td>
</tr>
<tr>
<td>≥55 years</td>
<td>T4b</td>
<td>Any N</td>
<td>M0</td>
<td>IVA</td>
</tr>
<tr>
<td>≥55 years</td>
<td>Any T</td>
<td>Any N</td>
<td>M1</td>
<td>IVC</td>
</tr>
</tbody>
</table>

### MEDULLARY THYROID CANCER

<table>
<thead>
<tr>
<th>TUMOR</th>
<th>NODE</th>
<th>METASTASIS</th>
<th>STAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>N0</td>
<td>M0</td>
<td>I</td>
</tr>
<tr>
<td>T2</td>
<td>N0</td>
<td>M0</td>
<td>II</td>
</tr>
<tr>
<td>T3</td>
<td>N0</td>
<td>M0</td>
<td>II</td>
</tr>
<tr>
<td>T1–3</td>
<td>N1a</td>
<td>M0</td>
<td>III</td>
</tr>
<tr>
<td>T4a</td>
<td>Any N</td>
<td>M0</td>
<td>IVA</td>
</tr>
<tr>
<td>T1–3</td>
<td>N1b</td>
<td>M0</td>
<td>IVA</td>
</tr>
<tr>
<td>T4b</td>
<td>Any N</td>
<td>M0</td>
<td>IVB</td>
</tr>
<tr>
<td>Any T</td>
<td>Any N</td>
<td>M1</td>
<td>IVC</td>
</tr>
</tbody>
</table>

### ANAPLASTIC CANCER

<table>
<thead>
<tr>
<th>TUMOR</th>
<th>NODE</th>
<th>METASTASIS</th>
<th>STAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1–T3a</td>
<td>N0/NX</td>
<td>M0</td>
<td>IVA</td>
</tr>
<tr>
<td>T1–T3a</td>
<td>N1</td>
<td>M0</td>
<td>IVB</td>
</tr>
<tr>
<td>T3b</td>
<td>Any N</td>
<td>M0</td>
<td>IVB</td>
</tr>
<tr>
<td>T4</td>
<td>Any N</td>
<td>M0</td>
<td>IVB</td>
</tr>
<tr>
<td>Any T</td>
<td>Any N</td>
<td>M1</td>
<td>IVC</td>
</tr>
</tbody>
</table>

### DEFINITIONS

#### Primary tumor (T)
- **TX** = Primary tumor cannot be assessed
- **T0** = No evidence of primary tumor
- **T1** = Tumor ≤2 cm in greatest dimension limited to the thyroid
  - **T1a** = Tumor ≤1 cm in greatest dimension limited to the thyroid
  - **T1b** = Tumor >1 cm but ≤2 cm in greatest dimension limited to the thyroid
- **T2** = Tumor >2 cm but ≤4 cm in greatest dimension limited to the thyroid
- **T3** = Tumor >4 cm limited to the thyroid, or gross extrathyroidal extension invading only strap muscles
  - **T3a** = Tumor >4 cm limited to the thyroid
  - **T3b** = Gross extrathyroidal extension invading only strap muscles (sternohyoid, sternothyroid, thyrohyoid, or omohyoid muscles) from a tumor of any size
- **T4** = Includes gross extrathyroidal extension beyond the strap muscles
  - **T4a** = Gross extrathyroidal extension invading subcutaneous soft tissues, larynx, trachea, esophagus, or recurrent laryngeal nerve from a tumor of any size
  - **T4b** = Gross extrathyroidal extension invading prevertebral fascia or encasing the carotid artery or mediastinal vessels from a tumor of any size

#### Regional lymph nodes (N)
- **NX** = Regional lymph nodes cannot be assessed
- **N0** = No evidence of locoregional lymph node metastasis
  - **N0a** = One or more cytologically or histologically confirmed benign lymph nodes
  - **N0b** = No radiologic or clinical evidence of locoregional lymph node metastasis
- **N1** = Metastasis to regional nodes
  - **N1a** = Metastasis to level VI or VII (pretracheal, paratracheal, or prelaryngeal/Delphian, or upper mediastinal) lymph nodes. This can be unilateral or bilateral disease.
  - **N1b** = Metastasis to unilateral, bilateral, or contralateral lateral neck lymph nodes (levels I, II, III, IV, or V) or retropharyngeal lymph nodes

#### Distant Metastasis (M)
- **M0** = No distant metastasis
- **M1** = Distant metastasis

However, these risks need to be balanced with the increased risk of hypoparathyroidism with routine central neck dissection and the fact that some studies do not show any difference in recurrence rates or rates of low or undetectable Tg levels. The updated 2015 ATA guidelines for thyroid cancer management suggest that prophylactic (ipsilateral or bilateral) dissection may be performed in patients with advanced (T3 or T4) papillary thyroid carcinoma, or if the lateral neck nodes are involved with tumor (N1b), or if the information will help in treatment planning.40 The American Head and Neck Society urges the involvement of multidisciplinary teams in the decision-making process for prophylactic CND and indicates that it can be considered in patients with high risk of recurrence (to include older or young age, multifocal disease and extrathyroidal extension in addition to the aforementioned factors).42 Further prospective studies are needed before definitive recommendations can be made in this regard.

Biopsy-proven lymph node metastases detected clinically or by imaging in the lateral neck in patients with papillary carcinoma are managed with modified radical or functional neck dissection,40 as described later in this chapter in “Thyroid Surgery.” Dissection of the posterior triangle and suprathyroid dissection usually are not necessary unless there is extensive metastatic disease in levels 2, 3, and 4, but should be performed when appropriate. Prophylactic lateral neck node dissection is not necessary in patients with PTC because these cancers do not appear to metastasize systemically from lymph nodes, and micrometastases often can be ablated with RAI therapy.

**Follicular Carcinoma** Follicular carcinomas account for 10% of thyroid cancers and occur more commonly in iodine-deficient areas. The overall incidence of this tumor is declining in the United States, probably due to iodine supplementation and improved histologic classification. Women have a higher incidence of follicular cancer, with a female-to-male ratio of 3:1, and a mean age at presentation of 50 years old. Follicular cancers usually present as solitary thyroid nodules, occasionally with a history of rapid size increase, and long-standing goiter. Pain is uncommon, unless hemorrhage into the nodule has occurred. Unlike papillary cancers, cervical lymphadenopathy is uncommon at initial presentation (about 5%), although distant metastases may be present. In <1% of cases, follicular cancers may be hyperfunctioning, leading patients to present with signs and symptoms of thyrotoxicosis. FNAB is unable to distinguish benign follicular lesions from follicular carcinomas. Therefore, preoperative clinical diagnosis of cancer is difficult unless distant metastases are present. Large follicular tumors (>4 cm) in older men are more likely to be malignant.

Due to the limitations inherent in the FNAB diagnosis, a number of studies have focused on identifying molecular markers to distinguish benign from malignant follicular lesions. Many of these genetic changes can be identified using tissue obtained during FNAB. While no single marker has met the ideal characteristics of being simple to use, reproducible, and cost-effective, several combinations of markers appear to be useful in differentiating benign from malignant follicular lesions. A commonly used panel of seven genes used to “rule in” malignancy detects mutations in BRAF, Ras, RET/PTC, and PAX/PPARγ and has been associated with a sensitivity of 57% to 75%, specificity of 97% to 100%, PPV of 87% to 100%, and NPV of 79% to 86% in the case of nodules consistent with Follicular/Hürthle cell neoplasms or suspicious of the same.43 In contrast, the Afirma Gene Expression Classifier (GEC) uses a “rule out” strategy to identify benign nodules. This method uses material from additional FNA passes (in an RNA-preserving solution) to analyze a 167 gene panel, and the results obtained are reported as benign or suspicious.44 It is reported to have a lower PPV of 37% but a better NPV of 94% for Follicular/Hürthle cell neoplasms. Next generation sequencing techniques have been used to enhance malignancy detection.

![Graph showing cumulative death rates](image-url)
by including additional mutations and gene arrangements. The advanced version of this assay (ThyroSeq V2) had a sensitivity of 90%, specificity of 93%, PPV of 83%, and NPV of 96% in a study of Follicular/Hürthle cell neoplasm/suspicious of the same nodules, making it useful as both a “rule in” and “rule out” test. Of note, these assays have also been evaluated in “AUS/FLUS” and “suspicious for malignancy” nodules with varying results, i.e., the performance characteristics of these tests have been noted to change depending upon the prevalence of malignancy in the tested population (pretest probability). At this time, the ATA guidelines do not advise molecular testing in the work-up of “suspicious for malignancy” nodules. Molecular testing may be used to supplement cytology results for malignancy risk assessment in “AUS/FLUS” or “follicular/Hürthle cell neoplasm/suspicious of the same” nodules depending on feasibility and informed patient preference. Expression arrays also have been used to investigate the role of microRNAs, which are a new class of small, noncoding RNAs that have been implicated in carcinogenesis. The specific microRNAs miR-197 and miR-346 are upregulated in follicular thyroid cancers and have the potential to be used as diagnostic markers. Additional studies have also demonstrated the feasibility of studying a panel of microRNAs in a small number of FNA samples. ThyGenX/ThyraMIR uses a mutation panel supplemented with 10 miRNA markers, whereas Rosetta GX Reveal is exclusively based on miRNA markers; however, both require further validation.

Pathology. Follicular carcinomas usually are solitary lesions, and the majority are surrounded by a capsule. Histologically, follicles are present, but the lumen may be devoid of colloid. Architectural patterns depend on the degree of differentiation demonstrated by the tumor. Malignancy is defined by the presence of capsular and vascular invasion (Fig. 38-18). In general, minimally invasive tumors appear grossly encapsulated and have microscopic invasion through the tumor capsule without extension into the parenchyma and/or invasion into small- to medium-sized vessels (venous caliber) in or immediately outside the capsule, but not within the tumor. On the other hand, widely invasive tumors demonstrate evidence of large vessel invasion and/or broad areas of tumor invasion through the capsule. They may, in fact, be unencapsulated. It is important to note that there is a wide variation of opinion among clinicians and pathologists with respect to the above definitions. Tumor infiltration and invasion, as well as tumor thrombi within the middle thyroid or jugular veins, may be apparent at operation.

Surgical Treatment and Prognosis. Patients diagnosed by FNAB as having a follicular lesion should undergo thyroid lobectomy because at least 70% to 80% of these patients will have benign adenomas. Total thyroidectomy is recommended by some surgeons in older patients with follicular lesions >4 cm because of the higher risk of cancer in this setting (50%) and certainly should be performed in patients with atypia on FNA, a family history of thyroid cancer, or a history of radiation exposure. Intraoperative frozen-section examination usually is not helpful, but it should be performed when there is evidence of capsular or vascular invasion or when adjacent lymphadenopathy is present. Total thyroidectomy should be performed when thyroid cancer is diagnosed. There is debate among experts about whether patients with minimally invasive follicular cancers should undergo completion thyroidectomy because the prognosis is so good in these patients. A diagnosis of frankly invasive carcinoma or follicular carcinoma with angioinvasion, with or without capsular invasion, necessitates completion of total thyroidectomy primarily so that 131I can be used to detect and ablate metastatic disease. Propylactic nodal dissection is not needed because nodal involvement is infrequent, but in the unusual patient with nodal metastases, therapeutic neck dissection is recommended. Propylactic central neck dissection may be considered in patients with large tumors. The cumulative mortality from follicular thyroid cancer is approximately 15% at 10 years and 30% at 20 years. Poor long-term prognosis is predicted by age over 50 years old at presentation, tumor size >4 cm, higher tumor grade, marked vascular invasion, extrathyroidal invasion, and distant metastases at the time of diagnosis.

Hürthle Cell Carcinoma. Hürthle cell carcinomas account for approximately 3% of all thyroid malignancies and, under the World Health Organization classification, are considered to be a subtype of follicular thyroid cancer. Hürthle cell cancers also are characterized by vascular or capsular invasion and, therefore, cannot be diagnosed by FNAB. Tumors contain sheets of eosinophilic cells packed with mitochondria, which are derived from the oxyphilic cells of the thyroid gland. Hürthle cell tumors differ from follicular carcinomas in that they are more often multifocal and bilateral (about 30%), usually do not take up RAI (about 5%), are more likely to metastasize to local nodes (25%) and distant sites, and are associated with a higher mortality rate (about 20% at 10 years). Hence, they are considered to be a separate class of tumors by some groups.

Management is similar to that of follicular neoplasms, with lobectomy and isthmusectomy being sufficient surgical treatment for unilateral Hürthle cell adenomas. When Hürthle cell neoplasms are found to be invasive on definitive paraffin-section histology, then total thyroidectomy should be performed. These patients should also undergo routine central neck node removal, similar to patients with MTC, and modified radical neck dissection when lateral neck nodes are palpable or identified by ultrasonography. Although RAI scanning and ablation usually are ineffective, they probably should be considered to ablate any residual normal thyroid tissue and occasionally ablate tumors because there is no other good therapy.

Postoperative Management of Differentiated Thyroid Cancer

Radioiodine Therapy. The issue of whether RAI therapy offers any benefit to patients with differentiated thyroid cancer

Figure 38-18. Hematoxylin-eosin–stained section from follicular thyroid carcinoma showing capsular invasion.
remains controversial in the absence of prospective, random- ized controlled trials. Long-term cohort studies by Mazzaferri and associates and DeGroot and colleagues demonstrate that postoperative RAI therapy reduces recurrence (Fig. 38-19) and provides a small improvement in survival, even in low-risk patients. Screening with RAI is more sensitive than chest X-ray or CT scanning for detecting metastases; however, it is less sensitive than Tg measurements for detecting metastatic disease in most differentiated thyroid cancers except Hürthle cell tumors. Screening and treatment are facilitated by the removal of all normal thyroid tissue, which effectively competes for iodine uptake. Metastatic differentiated thyroid carcinoma can be detected and treated by \( ^{131}I \) in about 75% of patients. Multiple studies show that RAI effectively treats >70% of lung micrometastases that are detected by RAI scan in the presence of a normal chest X-ray, whereas the success rates drop to <10% with pulmonary macrometastases. Early detection therefore is very important to improve prognosis.

Several features place patients at increased risk for local recurrences or metastases. The 2015 ATA guidelines use various features to risk-stratify tumors. Low-risk papillary thyroid cancer includes those without local tumor invasion, all macroscopic tumor resected, absence of aggressive histology (e.g., tall cell, columnar cell carcinoma), no known distant metastases (clinical or on RAI scan if done), no vascular invasion, clinical N0 or ≤5 pathologic N1 micrometastases (<0.2 cm in largest dimension), intrathyroidal, encapsulated follicular variant of papillary thyroid cancer, intrathyroidal, well differentiated follicular thyroid cancer with capsular invasion and no or minimal (<4 foci) vascular invasion and intrathyroidal papillary microcarcinoma (unifocal or multifocal, including \( BRAF ^{V_{600}E} \) mutated). Intermediate-risk tumors include those showing microscopic invasion of tumor into the perithyroidal soft tissues or RAI-avid metastatic foci in the neck on the first posttreatment whole-body RAI scan. This group also includes tumors with aggressive histology (e.g., tall cell, columnar cell carcinoma), papillary thyroid cancer with vascular invasion, clinical N1 or >5 pathologic N1 with all involved lymph nodes <3 cm in largest dimension and multifocal papillary microcarcinoma with extra-thyroidal extension (ETE) and \( BRAF ^{V_{600}E} \) mutated (if known). High-risk tumors include those demonstrating macroscopic invasion of tumor into the perithyroidal soft tissues (gross ETE), incomplete tumor resection, and presence of distant metastases (or postoperative serum thyroglobulin suggestive of distant metastate) or pathologic N1 with any metastatic lymph node ≥3 cm in largest dimension. Follicular thyroid cancers with extensive vascular invasion (>4 foci of vascular invasion) also fall into this category. It is important to note that this risk assessment represents a continuum with recurrence rates from 1% to 2% for low-risk cancers to >50% for high-risk cancers.

The current ATA guidelines recommend RAI therapy after surgical treatment for all patients with high-risk disease, i.e., those with gross ETE and M1 disease. RAI therapy is not recommended for patients with papillary microcarcinomas, either uni- or multifocal. RAI remnant ablation is not routinely recommended after thyroidectomy for ATA low-risk DTC patients. However, it may be considered in patients with aggressive histology or vascular invasion. Consideration of RAI is recommended for patients with intermediate-risk disease and “generally favored” for patients with microscopic ETE due to the risk of recurrent disease, large (>2–3 cm) or clinically evident lymph nodes (central, mediastinal, and lateral neck) or presence of extranodal extension. Advancing age may also favor RAI use. However, RAI is not needed for patients with a few (<5) microscopic nodal metastases in the central compartment in the absence of other adverse features as there is insufficient evidence for its utility in this setting. It is generally favored for patients with lateral neck disease. There is currently no established role for molecular testing in determining RAI therapy.

Remnant ablation can be performed with either thyroid hormone withdrawal or recombinant TSH (rTSH) stimulation. This is based on randomized studies showing that both techniques are equally effective in preparing patients for ablation, with the latter being associated with an improved quality of life.49,50 In patients with ATA high-risk disease (including distant metastases), there is insufficient data to recommend thyro- gen-mediated ablation, and hormone withdrawal is preferred. Furthermore, if patients have comorbidities that can be exacerbated by severe hypothyroidism (cardiac or psychiatric conditions), consideration should be given to thyrogen-mediated RAI. If hormone withdrawal is used, \( T_3 \) therapy should be discontinued for approximately 6 weeks before scanning with \( ^{131}I \). Patients should receive \( T_3 \) during this time period to decrease the period of hypothyroidism. \( T_3 \) has a shorter half-life than \( T_4 \) (1 day vs. 1 week) and needs to be discontinued for 2 weeks to allow TSH levels to rise before treatment. Levels >30 mU/L are considered optimal, based on noncontrolled studies. A low-iodine diet also is recommended during this 2-week period. The usual protocol involved administering a screening dose of 1 to 3 mCi and measuring uptake 24 hours later. After a total thyro- idectomy, this value should be <1%. A “hot” spot in the neck after initial screening usually represents residual normal tissue in the thyroid bed. Some investigators recommend omitting the scanning dose altogether to minimize thyrocyte “stunning” and subsequent requirement for higher treatment doses. Others recommend scanning only if the size of the remnant cannot be determined by the operative report or ultrasound, or if the results would alter the decision to treat or the dose to be administered. Current guidelines recommend using either \( ^{123}I \) or low-activity \( ^{131}I \) (1- to 3-mCi dose) and delivering a therapeutic dose within 72 hours.

The recommended dose of RAI is 30 mCi if remnant abla- tion is performed after total thyroidectomy for ATA low-risk thyroid or intermediate-risk cancer with lower risk features (i.e., low-volume central neck nodal metastases with no other known gross residual disease or any other adverse features). If RAI is given for adjuvant treatment to treat suspected microscopic disease (in the absence of metastatic disease), doses ranging from 30 to 150 mCi are recommended, and there is no solid evidence to show that higher doses reduce the recurrence rates for T3 and N1 disease in this setting.

If patients have an elevated Tg level, but negative RAI scan on follow-up, some physicians recommend treating once with 100 mCi of \( ^{131}I \) and repeating the imaging 1 to 2 weeks later. Approximately one-third of these patients demonstrate uptake on posttreatment imaging, and Tg levels usually decrease in these patients, documenting therapeutic benefit. The maximum dose of radiiodine that can be administered at one time without performing dosimetry is approximately 200 mCi with a cumulative dose of 1000 to 1500 mCi. Up to 500 mCi can be given with proper pretreatment dosimetry. Recent studies also
Figure 38-19. Tumor recurrence at a median of 16.7 years after thyroid surgery. The numerator is the number of patients with recurrence, and the denominator is the number of patients in each time interval. The $P$ values are derived from log-rank statistical analysis of 40-year life-table data. Figure shows that all recurrences (A) and distant metastases (B) were reduced in patients who received radioactive iodine (RAI) in addition to thyroxine (T4) therapy. (Reproduced with permission from Mazzaferri E, Kloos R: Current approaches to primary therapy for papillary and follicular thyroid cancer, Clin Endocrinol Metab. 2001 Apr; 86(4):1447-1363.)
show an increase in the number of second cancers in patients treated with RAI. The early and delayed complications of RAI therapy are listed in Table 38-7.

**Thyroid Hormone** $T_4$ is necessary as replacement therapy in patients after total or near-total thyroidectomy, and also has the additional effect of suppressing TSH and reducing the growth stimulus for any possible residual thyroid cancer cells. TSH suppression reduces tumor recurrence rates. Current guidelines advise maintaining initial TSH levels <0.1 mU/mL in patients with high-risk thyroid cancer and in the range of 0.1 to 0.5 mU/mL in patients with intermediate-risk disease. For low-risk patients (with or without remnant ablation) with undetectable serum Tg levels, TSH levels can be maintained at the lower end of the reference range (0.5–2 mU/L). If these patients have low measureable Tg levels, it is recommended that TSH be maintained at or slightly below lower limit of normal (0.1 to 0.5 mU/L) while continuing surveillance for recurrence. In low-risk patients treated with lobectomy alone, it is advised to keep TSH in the mid to lower reference range (0.5–2 mU/L), and hormone therapy may be needed to maintain these levels. Further TSH suppression levels are determined by response to therapy. The risk of tumor recurrence must be balanced with the side effects associated with prolonged TSH suppression, including osteopenia and cardiac problems, particularly in older patients.

**Follow-Up of Patients With Differentiated Thyroid Cancer**

**Thyroglobulin Measurement** Tg and anti-Tg antibody levels should be measured initially at 6 to 12 month intervals and more frequently in patients with high-risk tumors. Further measurements are guided by response to therapy. Patients are considered to have an excellent response to treatment if suppressed Tg is <0.2 ng/mL and stimulated Tg is <1 ng/mL with negative imaging. In these patients, Tg levels can be followed every 12 to 24 months while on thyroid hormone as their risk of recurrence is low (1–4%). Patients with structurally or biochemically incomplete (negative imaging but suppressed Tg ≥1 ng/mL or stimulated Tg ≥10 ng/mL or rising anti-Tg levels) or indeterminate responses (nonspecific imaging findings, suppressed Tg detectable but <1 ng/mL, and stimulated Tg detectable but <10 ng/mL or stable or declining anti-Tg levels) require additional investigations. Tg measurements in FNAB aspirates have also been shown to be useful in the detection of nodal metastatic disease.

**Imaging** After the first posttreatment scan, low- and some intermediate-risk patients with negative TSH-stimulated Tg and cervical ultrasound do not require routine diagnostic whole-body radiiodine scans. However, diagnostic whole-body scans 6 to 12 months after remnant ablation may be of value in the follow-up of patients with high- or intermediate-risk patients with higher risk features. Other scenarios for follow-up scans include patients with abnormal uptake outside the thyroid bed on posttherapy scan, those with poorly informative postablotion scans (e.g., due to high thyroid bed uptake), and patients with Tg antibodies.

Cervical ultrasound be performed to evaluate the thyroid bed and central and lateral cervical nodal compartments at 6 and 12 months after thyroidectomy and then annually for at least 3 to 5 years, depending on the patient’s risk for recurrent disease and Tg status. Sonographically suspicious nodes ≥8 to 10 mm on the smallest diameter measurement should be biopsied for cytology as well as Tg measurement in the aspirate washout. Smaller nodes can be followed and biopsied if there is continued growth. FDG-PET and PET-CT scans have been shown to be useful to localize recurrent or persistent thyroid cancer in patients who have Tg-positive, RAI scan–negative disease. FDG-PET can also be useful for the initial staging of patients with poorly differentiated thyroid carcinomas or Hürthle cell tumors, particularly in patients with other evidence of disease on imaging or Tg levels. In addition, they may be used as a prognostic tool in patients with metastatic disease and to evaluate the response to treatment in patients with metastatic or locally advanced disease.

**Additional Treatment Modalities**

**Radiotherapy, Thermal Ablation, and Chemotherapy**

External-beam radiotherapy is occasionally required to control unresectable, locally invasive, or recurrent disease and to treat metastases in support bones to decrease the risk of fractures. It also is of value for the treatment and control of pain from bony metastases when there is minimal or no RAIU. Stereotactic brain radiotherapy and intensity-modulated radiation therapy have both been used successfully for metastatic lesions. Percutaneous thermal ablation by increasing (radiofrequency ablation) or decreasing temperature (cryoablation) in the lesion to induce irreversible cellular damage has shown promise for lung, bone, and liver lesions. Single-drug and multidrug chemotherapy has been used with little success in disseminated thyroid cancer, and there is no role for routine chemotherapy. Doxorubicin (Adriamycin) and paclitaxel (Taxol) were previously the most frequently used agents. The former acts as a radiation sensitizer and should be considered in patients undergoing external-beam radiation.

**Novel Therapies** These therapies are directed at the molecular pathways known to be involved in thyroid cancers. Sorafenib
and lenvatinib are U.S. Food and Drug Administration (FDA) and European Medical Agency (EMA)-approved for use in patients with advanced differentiated thyroid cancer that is non-responsive to RAI after evaluation in phase 3 placebo-controlled double blinded trials (the DECISION study and the SELECT study, respectively).55,56 Both drugs are multikinase inhibitors and target RET kinase and the vascular endothelial growth factor (VEGF)-receptor; however, lenvatinib also inhibits the fibroblast growth factor and the platelet-derived growth factor receptor. Sorafenib demonstrated progression-free survival (PFS) improvement by 5 months with about 12% partial response rates, whereas lenvatinib prolonged median PFS by 15.7 months compared with placebo with response rates of 65%, including some complete responses. Vandetanib is mainly a RET-kinase inhibitor, but it also affects the VEGF-receptor and epidermal growth factor receptor. It has been evaluated in a phase 2 trial and also demonstrated an improved PFS.57 However, none of the agents show improvements in overall survival. Moreover, they are associated with significant side effects (diarrhea, fatigue, hypertension, hepatotoxicity, bleeding, and thrombosis) that affect patient quality of life. As such, they are considered only in patients with metastatic, rapidly progressive, symptomatic disease that is unable to respond to other local treatment approaches and generally in the context of clinical trials. Oncogenic kinase inhibitors that selectively inhibit the mutant V600E BRAF kinase (dabrafenib) has also shown promise in treating a subset of patients with advanced differentiated thyroid cancer.58

**Medullary Carcinoma** MTC accounts for about 5% of thyroid malignancies and arises from the parafollicular or C cells of the thyroid, which, in turn, are derived from the ultimobranchial bodies. These cells are concentrated superolaterally in the thyroid lobes, and this is where MTC usually develops. C cells secrete calcitonin, a 32-amino-acid polypeptide that functions to lower serum calcium levels, although its effects in humans are minimal. Most MTCs occur sporadically. However, approximately 25% occur within the spectrum of several inherited syndromes such as familial MTC, MEN2A, and MEN2B. All these variants are known to result secondary to germline mutations in the RET proto-oncogene. The syndromes also are characterized by genotype-phenotype correlations, with specific mutations leading to particular clinical manifestations. The salient clinical and genetic features of these syndromes are outlined in Table 38-8. Some clinical features of MEN2B patients are shown in Fig. 38-20.

**Table 38-8**

<table>
<thead>
<tr>
<th>SYNDROME</th>
<th>MANIFESTATIONS</th>
<th>RET MUTATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEN2A</td>
<td>MTC, pheochromocytoma, primary hyperparathyroidism, lichen planus amyloidosis</td>
<td>Exon 10—codons 609, 611, 618, 620 Exon 11—codon 634 (more commonly associated with pheochromocytoma and primary hyperparathyroidism)</td>
</tr>
<tr>
<td>MEN2B</td>
<td>MTC, pheochromocytoma, Marfanoid habitus, mucocutaneous ganglioneuromatosis</td>
<td>Exon 16—codon 918</td>
</tr>
<tr>
<td>Familial MTC</td>
<td>MTC</td>
<td>Codons 609, 611, 618, 620, and 634 Codons 768, 790, 791, or 804 (rare)</td>
</tr>
<tr>
<td>MEN2A and Hirschsprung’s disease</td>
<td>MTC, pheochromocytoma, primary hyperparathyroidism, Hirschsprung’s disease</td>
<td>Codons 609, 618, 620</td>
</tr>
</tbody>
</table>

MEN2 = multiple endocrine neoplasia type 2; MTC = medullary thyroid cancer.

![Figure 38-20](image-url). Features of MEN2B: thickened lips (A) and mucosal neuromas (A and B).
Patients with MTC often present with a neck mass that may be associated with palpable cervical lymphadenopathy (15% to 20%). Pain or aching is more common in patients with these tumors, and local invasion may produce symptoms of dysphagia, dyspnea, or dysphonia. Distant blood-borne metastases to the liver, bone (frequently osteoblastic), and lung occur later in the disease. The female-to-male ratio is 1.5:1. Most patients present between 50 and 60 years old, although patients with familial disease present at a younger age. Medullary thyroid tumors secrete not only calcitonin and carcinoembryonic antigen (CEA), but also other peptides such as calcitonin gene–related peptide, histaminases, prostaglandins E₂, and F₂ alpha, and serotonin. Patients with extensive metastatic disease frequently develop diarrhea, which may result from increased intestinal motility and impaired intestinal water and electrolyte flux. About 2% to 4% of patients develop Cushing’s syndrome as a result of ectopic production of adrenocorticotropic hormone (ACTH).

Pathology. MTCs typically are unilateral (80%) in patients with sporadic disease and multicentric in familial cases, with bilateral tumors occurring in up to 90% of familial patients. Familial cases also are associated with C-cell hyperplasia, which is considered a premalignant lesion. Microscopically, tumors are composed of sheets of infiltrating neoplastic cells separated by collagen and amyloid. Marked heterogeneity is present; cells may be polygonal or spindle shaped. The presence of amyloid is a diagnostic finding, but immunohistochemistry for calcitonin is more commonly used as a diagnostic tumor marker. These tumors also stain positively for CEA and calcitonin gene–related peptide.

Diagnosis. The diagnosis of MTC is established by history, physical examination, raised serum calcitonin, or CEA levels, and FNAB cytology of the thyroid mass. Attention to family history is important because about 25% of patients with MTC have familial disease. Because it is not possible to distinguish sporadic from familial disease at initial presentation, all new patients with MTC should be screened for RET point mutations, pheochromocytoma, and HPT. Screening of patients with familial MTC for RET point mutations has largely replaced using provocative testing with pentagastrin or calcium–stimulated calcitonin levels to make the diagnosis. Calcitonin and CEA are used to identify patients with persistent or recurrent MTC. Calcitonin is a more sensitive tumor marker, but CEA is a better predictor of prognosis.

Treatment. The ATA published revised guidelines for the management of medullary cancers in 2015.⁵⁹ A neck ultrasound is recommended to evaluate the central and lateral neck compartments and the superior mediastinum. Serum calcitonin, CEA, calcium levels should also be measured, and RET proto-oncogene mutation testing should be performed. Pheochromocytomas need to be excluded. If patients are found to have a pheochromocytoma, this must be operated on first. Primary hyperparathyroidism, if present, is treated at the time of thyroidectomy. These tumors are generally (>50%) bilateral. Total thyroidectomy is the treatment of choice for patients with MTC because of the high incidence of multicentricity, the more aggressive course, and the fact that ¹³¹I therapy usually is not effective.

Central compartment nodes frequently are involved early in the disease process, so that a bilateral prophylactic central neck node dissection should be routinely performed. In patients with palpable or imaging-detected cervical nodes, symptoms and signs of distant disease or calcitonin levels >500 pg/mL, additional imaging to include a neck and chest CT and a triple-phase liver CT or contrast-enhanced MRI and an axial MRI/bone scan is recommended to assess for metastatic disease. In patients with no distant disease but nodal involvement, an ipsilateral or bilateral lateral neck dissection (levels IIa, III, IV, and V) is performed. Less aggressive neck surgery should be considered to preserve speech and swallowing while maintaining locoregional control in patients with limited metastatic disease. The role of prophylactic lateral neck dissection is controversial and may be considered based on calcitonin levels. Some groups favor this procedure if central neck lymph nodes are involved or if the primary tumor is ≥1.5 cm.

In the case of locally recurrent or widely metastatic disease, tumor debulking is advised not only to ameliorate symptoms of pain, flushing, and diarrhea, but also to decrease risk of death from recurrent central neck or mediastinal disease. External-beam radiotherapy is controversial but can be considered for patients with resected T4 disease and for patients with unresectable residual or recurrent tumor and symptomatic bony metastases. Liver metastases tend to be multiple and are typically not amenable to resection, percutaneous ethanol ablation, or radiofrequency ablation. However, chemoembolization may be helpful in this setting. There is no effective chemotherapy regimen.

Various targeted therapies directed against the RET kinase have been investigated for the treatment of MTC.⁵⁸ Many of these also inhibit VEGF receptor due to their close structural similarities. Sorafenib, sunitinib, lenvatinib, and cabozantinib are some such multikinase inhibitors, whereas axitinib and pazopanib act only on VEGFR. Vandetanib inhibits both targets and is also an EGF receptor inhibitor, and cabozantinib targets c-MET in addition to RET and VEGF receptor. Both drugs are currently approved by the FDA and EMA for the treatment of advanced and progressive MTC based on data that they prolong progression-free survival, in addition to reducing secretion of calcitonin and CEA.⁶⁰,⁶¹ They are recommended as first-line systemic therapy in symptomatic patients with advanced MTC. An anti-CEA monoclonal antibody (labetuzumab) also has shown antitumor response in a small group of patients. Patients with recurrent/metastatic disease should be enrolled in well-designed clinical trials.

In patients who have hypercalcemia and an increased PTH at the time of thyroidectomy, only obviously enlarged parathyroid glands should be removed. The other parathyroid glands should be preserved and marked in patients with normocalcemia, as only about 20% of patients with MEN2A develop HPT. When a normal parathyroid cannot be maintained on a vascular pedicle, it should be removed, biopsied to confirm that it is a parathyroid, and then autotransplanted to the forearm of the nondominant arm, particularly in patients with MEN2A. Reimplantation into the sternocleidomastoid muscle is also acceptable for patients with known MEN2B and familial MTC.

Prophylactic total thyroidectomy is indicated in RET mutation carriers once the mutation is confirmed. The ATA guidelines stratify mutation into various risk levels to offer recommendations regarding age at which a prophylactic...
thyroidectomy should be performed and to predict phenotypes, including pheochromocytomas. In general, in patients with less aggressive mutations (designated ATA moderate-risk), thyroideotomy may be delayed >5 years, especially if there is a normal annual serum calcitonin, neck ultrasound, less aggressive family history, or family preference. Children with MEN2A and mutations at codon 634 (designated high-risk) are advised to undergo thyroideotomy at <5 years of age, and those with MEN2B-related mutations (designated highest-risk) should undergo the procedure before age 1. Central neck dissection can be avoided in children who are RET-positive and calcitonin-negative with a normal ultrasound examination. When the calcitonin is increased or the ultrasound suggests a thyroid cancer, a prophylactic central neck dissection is indicated.

Postoperative Follow-Up and Prognosis. Patients are followed by annual measurements of calcitonin and CEA levels, in addition to history and physical examination. Other modalities used to localize recurrent disease include ultrasound, CT, MRI, and more recently, FDG-PET/CT scans. Prognosis is related to disease stage. The 10-year survival rate is approximately 80% but decreases to 45% in patients with lymph node involvement. Survival also is significantly influenced by disease type. It is best in patients with non-MEN familial MTC, followed by those with MEN2A, and then those with sporadic disease. Prognosis is the worst (survival of 35% at 10 years) in patients with MEN2B. Performing prophylactic surgery in RET oncogene mutation carriers not only improves survival rates but also renders most patients calcitonin free.

Anaplastic Carcinoma Anaplastic carcinoma accounts for approximately 1% of all thyroid malignancies in the United States. Women are more commonly affected, and the majority of tumors present in the seventh and eighth decade of life. The typical patient has a long-standing neck mass, which rapidly enlarges and may be painful. Associated symptoms such as dysphonia, dysphagia, and dyspnea are common. The tumor is large and may be fixed to surrounding structures or may be ulcerated with areas of necrosis (Fig. 38-21). Lymph nodes usually are palpable at presentation. Evidence of metastatic spread also may be present. Diagnosis is confirmed by FNAB revealing characteristic giant and multinucleated cells. Differential diagnoses on FNA can include lymphomas, medullary carcinomas, direct extension from a laryngeal carcinoma, or other metastatic disease. When spindle cell elements are present, primary and metastatic sarcomas need to be considered as well. Immunohistochemical markers can aid with excluding other carcinomas or melanoma. When spindle cell elements are present, a prophylactic central neck dissection is indicated.

Pathology. On gross inspection, anaplastic tumors are firm and whitish in appearance. Microscopically, sheets of cells with marked heterogeneity are seen. The three main histologic growth patterns are spindle cell, squamoid, and pleomorphic giant cell. Tumors may show a predominance of one pattern or a mixture of various patterns. Foci of more differentiated thyroid tumors, either follicular or papillary, may be seen, suggesting that anaplastic tumors arise from more well-differentiated tumors.

Treatment and Prognosis. This tumor is one of the most aggressive thyroid malignancies, with few patients surviving 6 months beyond diagnosis. All forms of treatment have been disappointing. The ATA has published guidelines for the management of patients with anaplastic cancer. Imaging ( ultrasound, CT, MRI, or PET-CT) should be obtained to assess resectability. All patients should have preoperative laryngoscopy to assess the status of the vocal cords. A total or near-total thyroideotomy with therapeutic lymph node dissection is recommended for patients with an intrathyroidal mass (although lobectomy may also be appropriate, particularly if there is concern for vocal cord paralysis). If extrathyroidal extension is present, an en bloc resection should be considered if all gross disease can be removed (R1). Tracheostomy should be avoided as long as possible unless there is impending airway loss. Adjuvant radiation which should be offered to patients with a good performance status and no metastatic disease who desire aggressive management. Cytotoxic chemotherapy (with some combination of a taxane, anthracycline, and platinum) is typically given concurrently and has been associated with prolonged survival, although these agents are also being used in a neoadjuvant fashion, particularly in patients with unresectable disease.

Lymphoma Lymphomas account for <1% of thyroid malignancies, and most are of the non-Hodgkin’s B-cell type. Although the disease can arise as part of a generalized lymphomatous condition, most thyroid lymphomas develop in patients with chronic lymphocytic thyroiditis. Chronic antigenic lymphocyte stimulation has been suggested to result in lymphocyte transformation. Patients usually present with symptoms similar to those of patients with anaplastic carcinoma, although the rapidly enlarging neck mass often is painless. Patients may present with acute respiratory distress. Ultrasound can be useful for early diagnosis, and lymphoma appears as a well-defined hypoechoic mass. The diagnosis usually is suggested by FNAB, but FNAB can be nondiagnostic, particularly in the setting of low-grade lymphomas. Therefore, needle core or open biopsy may be necessary for definitive diagnosis. Staging studies should be obtained expeditiously to assess the extent of extrathyroidal spread.
Treatment and Prognosis. Patients with thyroid lymphoma respond rapidly to chemotherapy (CHOP—cyclophosphamide, doxorubicin, vincristine, and prednisone), which also has been associated with improved survival. Combined treatment with radiotherapy and chemotherapy often is recommended. Thyroidectomy and nodal resection are used to alleviate symptoms of airway obstruction in patients who do not respond quickly to the above regimens or who have completed the regimen before diagnosis. Prognosis depends on the histologic grade of the tumor and whether the lymphoma is confined to the thyroid gland or is disseminated. The overall 5-year survival rate is about 50%; patients with extrathyroidal disease have markedly lower survival rates.

Metastatic Carcinoma The thyroid gland is a rare site of metastases from other cancers, including kidney, breast, lung, and melanoma. Clinical examination and a review of the patient’s history often suggest the source of the metastatic disease, and FNAB usually provides definitive diagnosis. Resection of the thyroid, usually lobectomy, may be helpful in many patients, depending on the status of their primary tumor.

Thyroid Surgery

Conduct of Thyroidectomy Patients with any recent or remote history of altered phonation or prior neck or upper chest surgery that places the recurrent laryngeal or vagus nerves at risk should undergo vocal cord assessment by direct or indirect laryngoscopy before thyroidectomy. Laryngeal examination is also advised in patients with known posterior extension of thyroid cancer and extensive central nodal metastases. The patient is positioned supine, with a sandbag between the scapulae. The head is placed on a donut cushion, and the neck is extended to provide maximal exposure. A Kocher transverse collar incision, typically 3 to 5 cm in length, is placed in or parallel to a natural skin crease 1 cm below the cricoid cartilage (Fig. 38-22A), although longer incisions may be needed. The subcutaneous tissues and platysma are incised sharply, and subplatysmal flaps are raised superiorly to the level of the thyroid cartilage and inferiorly to the suprasternal notch (Fig. 38-22B). The strap muscles are divided in the midline along the entire length of the mobilized flaps, and the thyroid gland is exposed. On the side to be approached first, the sternothyroid muscles are separated from the underlying sternothyroid muscle by blunt dissection until the internal jugular vein and ansa cervicalis nerve are identified. The strap muscles rarely need to be divided to gain exposure to the thyroid gland. If this maneuver is necessary, the muscles should be divided high to preserve their innervation by branches of the ansa cervicalis. If there is evidence of direct tumor invasion into the strap muscles, the portion of involved muscle should be resected en bloc with the thyroid gland. The sternothyroid muscle is then dissected off the underlying thyroid by a combination of sharp and blunt dissection, thus exposing the middle thyroid veins. The thyroid lobe is retracted medially and anteriorly, and the lateral tissues are swept posterolaterally using a peanut sponge. The middle thyroid veins are ligated and divided (Fig. 38-22C). Attention is then turned to the midline where Delphian nodes and the pyramidal lobe are identified. The fascia just cephalad and caudal to the isthmus is divided. The superior thyroid pole is identified by retracting the thyroid first inferiorly and medially, and then the upper pole of the thyroid is mobilized caudally and laterally. The dissection plane is kept as close to the thyroid as possible, and the superior pole vessels are individually identified, skeletonized, ligated, and divided low on the thyroid gland to avoid injury to the external branch of the superior laryngeal nerve (Fig. 38-22D). Once these vessels are divided, the tissues posterior and lateral to the superior pole can be swept from the gland in a posteromedial direction, to reduce the risk of damaging vessels supplying the upper parathyroid.

The RLNs should then be identified, and the ATA 2015 guidelines strongly recommend visual identification in all cases. The course of the right RLN is more oblique than the left RLN. The nerves can be most consistently identified at the level of the cricoid cartilage. The parathyroids usually can be identified within 1 cm of the crossing of the inferior thyroid artery and the RLN, although they also may be ectopic in location. The lower pole of the thyroid gland should be mobilized by gently sweeping all tissues dorsally. The inferior thyroid vessels are dissected, skeletonized, ligated, and divided as close to the surface of the thyroid gland as possible to minimize devascularization of the parathyroids (extracapsular dissection) or injury to the RLN. The RLN is most vulnerable to injury in the vicinity of the ligament of Berry. The nerve often passes through this structure along with small crossing arterial and venous branches (Fig. 38-22E). Any bleeding in this area should be controlled with gentle pressure before carefully identifying the vessel and ligating it. Use of the electrocautery should be avoided in proximity to the RLN. Once the ligament is divided, the thyroid can be separated from the underlying trachea by sharp dissection. The pyramidal lobe, if present, must be dissected in a cephalad direction to above the level of the notch in the thyroid cartilage or higher in continuity with the thyroid gland. If a lobectomy is to be performed, the isthmus is divided flush with the trachea on the contralateral side and suture ligated. The procedure is repeated on the opposite side for a total thyroidectomy.

Parathyroid glands located anteriorly on the surface of the thyroid that cannot be dissected from the thyroid with a good blood supply or that have been inadvertently removed during the thyroidectomy should be resected, confirmed as parathyroid tissue by frozen section, divided into 1-mm fragments, and reimplemented into individual pockets in the sternocleidomastoid muscle. The sites should be marked with silk sutures and a clip. Various novel techniques using indocyanine fluorescence angiography and near-infrared autofluorescence have shown utility in the identification and viability assessment of parathyroid glands; however, they are not routinely used at the present time. If a subtotal thyroidectomy is to be performed, once the superior pole vessels are divided and the thyroid lobe mobilized anteriorly, the thyroid lobe is cross-clamped with a Mayo clamp, leaving approximately 4 g of the posterior portion of the thyroid. The thyroid remnant is suture ligated, taking care to avoid injury to the RLN. Routine drain placement rarely is necessary. After adequate hemostasis is obtained, the strap muscles are reaproximated in the midline. The platysma is approximated in a similar fashion. The skin can be closed with subcuticular sutures or clips.

Nerve Monitoring Intraoperative RLN and external laryngeal nerve monitoring techniques are being increasingly used during thyroid and parathyroid surgery. Both continuous monitoring using endotracheal tube electrodes and intermittent monitoring by periodic stimulation and laryngeal palpation are used. Many published studies have established the feasibility of nerve monitoring; however, none were able to show that the technique equivocally reduces nerve injury (particularly by experienced
Figure 38-22. Conduct of thyroidectomy. A. Correct placement of thyroidectomy incision. B. Raising subplatysmal flaps. C. Dissection of middle thyroid vein. D. Dissection of the superior pole vessels, which should be individually ligated. E. Dissection at the ligament of Berry. Note small artery and vein within the ligament and the recurrent laryngeal nerve coursing laterally. F. Endoscopic thyroidectomy via axillary incisions. m. = muscle; n. = nerve; v. = vein.
Adopted. The current ATA guidelines recommend that intraoperative neural stimulation may be used to facilitate nerve identification of the RLN and confirm its function, especially prior to proceeding with contralateral thyroidectomy.

**Minimally Invasive Approaches** Several approaches to minimally invasive thyroidectomy have been described. Mini-incision procedures use a small, 3-cm incision with no flap creation and minimal dissection to deliver the thyroid into the wound and then perform the pretracheal and paratracheal dissection. Video assistance can be used to improve the visualization.
via the small incision. Totally endoscopic approaches also have been described, via the supraventricular, anterior chest, axillary, and breast approach. The axillary, anterior chest, and breast approaches eliminate the skin incision in the neck but are more invasive. The endoscopic approaches can also be performed with the assistance of robotic techniques. More recently, there have been studies of transoral robotic-assisted thyroidectomy in which the thyroid is approached through the oral cavity. These methods are feasible, but clear benefits over the “traditional” open approach via small neck incisions have not been established.

Typically, endoscopic thyroidectomies are performed under general anesthesia. For the axillary approach, a 30-mm skin incision is made in the axilla, and 12-mm and 5-mm trocars are inserted through this incision (Fig. 38-22F). An additional 5-mm trocar is inserted adjacent to the incision. For the anterior chest approach, a 12-mm skin incision is made in the skin of the anterior chest approximately 3 to 5 cm below the border of the ipsilateral clavicle. Two additional 5-mm trocars are inserted by endoscopic guidance below the ipsilateral clavicle, and carbon dioxide (CO₂) is then insufflated up to a pressure of 4 mmHg to facilitate creation of a working space. The anterior border of the sternocleidomastoid muscle is then separated from the sternohyoid muscle to expose the sternothyroid muscle. The thyroid gland is exposed by splitting the sternothyroid muscle. The lower pole is retracted upward and dissected from the adipose tissue to identify the RLN. As the RLN is exposed, Berry’s ligament is exposed and incised with a 5-mm clip or laparoscopic coagulating shears. The upper pole of the thyroid gland is separated from the cricothyroid muscle, and the external branch of the superior laryngeal nerve can be identified during this maneuver. The upper pole of the thyroid gland then is dissected free.

**Surgical Removal of Intrathoracic Goiter** A goiter is considered mediastinal if at least 50% of the thyroid tissue is located intrathoracically. Mediastinal goiters can be primary or secondary. Primary mediastinal goiters constitute approximately 1% of all mediastinal goiters and arise from accessory (ectopic) thyroid tissue located in the chest. These goiters are supplied by intrathoracic blood vessels and do not have any connection to thyroid tissue in the neck. The vast majority of mediastinal goiters are, however, secondary mediastinal goiters that arise from downward extension of cervical thyroid tissue along the fascial planes of the neck and derive their blood supply from the superior and inferior thyroid arteries. Virtually all intrathoracic goiters can be removed via a cervical incision. Patients who have (a) invasive thyroid cancers, (b) had previous thyroid operations and may have developed parasitic mediastinal vessels, or (c) primary mediastinal goiters with no thyroid tissue in the neck may require a median sternotomy for removal. The chest, however, should be prepared in most cases in the event it is necessary to perform a median sternotomy to control mediastinal bleeding or completely remove an unsuspected invasive cancer. The goiter is approached via a neck incision. The superior pole vessels and the middle thyroid veins are identified and ligated first. Early division of the isthmus helps with subsequent mobilization of the substernal goiter from beneath the sternum. Placement of large 1-0 or 2-0 sutures deep into the goiter, when necessary, helps deliver it. For patients in whom thyroid cancer is suspected or demonstrated in an intrathoracic gland, attempts should be made to avoid rupture of the thyroid capsule. When sternotomy is indicated, the sternum usually should be divided to the level of the third intercostal space and then laterally on one side at the space between the third and fourth ribs (Fig. 38-23).

**Central and Lateral Neck Dissection for Nodal Metastases** Central compartment (medial to the carotid sheath) lymph nodes frequently are involved in patients with papillary, medullary, and Hürthle cell carcinomas and should be removed at the time of thyroidectomy, preserving the RLNs and parathyroid glands. Central neck dissection is particularly important in patients with medullary and Hürthle cell carcinoma because of the high frequency of microscopic tumor spread and because these tumors cannot be ablated with ¹³¹I. An ipsilateral modified radical neck dissection is indicated in the presence of palpable cervical lymph nodes or prophylactically in some patients with medullary carcinoma.

A modified radical (functional) neck dissection can be performed via the cervical incision used for thyroidectomy, which can be extended laterally (Fig. 38-24A) to the anterior margin of the trapezius muscle. The procedure involves removal of all fibro-fatty tissue along the internal jugular vein (levels II, III, and IV) and the posterior triangle (level V). In contrast to a radical neck dissection, the internal jugular vein, the spinal accessory nerve, the cervical sensory nerves, and the sternocleidomastoid muscle are preserved unless they are adherent to or invaded by tumor. The procedure begins by opening the plane between the strap muscles medially and the sternocleidomastoid muscle laterally. The anterior belly of the omohyoid muscle is retracted laterally, and the dissection is carried posteriorly until the carotid sheath is reached. The internal jugular vein is retracted medially with a vein retractor and the fibro-fatty tissue and lymph nodes are dissected away from it by a combination of sharp and blunt dissection. The lateral dissection is carried along the posterior border of the sternocleidomastoid muscle, removing the tissue from the posterior triangle. The deep dissection plane is the anterior scalenus muscle, the phrenic nerve, the brachial plexus, and the medial scalenus muscle. The phrenic nerve is preserved on the scalenus anterior muscle, as are the cervical sensory nerves in most patients (Fig. 38-24B). Dissection along the spinal accessory nerve superiorly is most important because this is a frequent site of metastatic disease.
Complications of Thyroid Surgery  Nerves, parathyroids, and surrounding structures are all at risk of injury during thyroidectomy. Injury to the RLN may occur by severance, ligation, or traction, but should occur in <1% of patients undergoing thyroidectomy by experienced surgeons. The RLN is most vulnerable to injury during the last 2 to 3 cm of its course, but also can be damaged if the surgeon is not alert to the possibility of nerve branches and the presence of a nonrecurrent nerve, particularly on the right side. If the injury is recognized intraoperatively, most surgeons advocate primary reapproximation of the perineurium using nonabsorbable sutures. Approximately 20% of patients are at risk of injury to the external branches of the

Figure 38-24. Conduct of neck dissection. A. Incisions for modified radical neck dissection. B. Anatomic relations of structures identified during a modified radical neck dissection. a. = artery; m. = muscle; n. = nerve.
superior laryngeal nerve, especially if superior pole vessels are ligated en masse. The cervical sympathetic trunk is at risk of injury in invasive thyroid cancers and retroesophageal goiters and may result in Horner’s syndrome. Transient hypocalcemia (from surgical injury or inadvertent removal of parathyroid tissue) has been reported in up to 50% of cases, but permanent hypoparathyroidism occurs <2% of the time. Postoperative hypocalcemia is more likely in patients who undergo concomitant thyroidectomy and central and lateral neck dissection and in patients with Graves’ disease. Postoperative hematomas or bleeding may also complicate thyroidectomies and rarely necessitate emergency reoperation to evacuate the hematoma. Bilateral vocal cord dysfunction with airway compromises requires immediate reintubation and tracheostomy. Seromas may need aspiration to relieve patient discomfort. Wound cellulitis and infection and injury to surrounding structures, such as the carotid artery, jugular vein, and esophagus, are infrequent.

**PARATHYROID**

**Historical Background**

In 1849, the curator of the London Zoological Gardens, Sir Richard Owen, provided the first accurate description of the normal parathyroid gland after autopsy examination of an Indian rhinoceros. However, human parathyroids were not grossly and microscopically described until 1879 by Ivar Sandström, a medical student in Uppsala, Sweden. He suggested that these glands be named the *glandulae parathyroideae*, although their function was not known.

The association of HPT and the bone disease osteitis fibrosa cystica (described by von Recklinghausen) was recognized in 1903. Calcium measurement became possible in 1909, and the association between serum calcium levels and the parathyroid glands was established. The first successful parathyroidectomy was performed in 1925 by Felix Mandl on a 38-year-old man who had severe bone pain secondary to advanced osteitis fibrosa cystica. The patient’s condition dramatically improved after the operation, and he lived for another 7 years before dying of recurrent HPT or renal failure. In 1926, the first parathyroid operation was performed at Massachusetts General Hospital. Edward Churchill, assisted by an intern named Oliver Cope, operated on the famous sea captain Charles Martell for severe primary HPT (PHPT). It was not until his seventh operation, which included total thyroidectomy, that an ectopic adenoma was found subterminally. Unfortunately, Captain Martell died 6 weeks later, likely due to laryngeal spasm and complications of renal stones and ureteral obstruction. The first successful parathyroidectomy for HPT in the United States was performed on a 56-year-old woman in 1928 by Isaac Y. Olch at the Barnes Hospital in St. Louis, Missouri. At operation, a parathyroid adenoma was found attached to the left lower lobe of the thyroid gland. Postoperatively, the patient developed tetany, requiring lifelong supplemental calcium.

**Embryology**

In humans, the superior parathyroid glands are derived from the fourth branchial pouch, which also gives rise to the thyroid gland. The third branchial pouches give rise to the inferior parathyroid glands and the thymus (Fig. 38-25). The parathyroids remain closely associated with their respective branchial pouch derivatives. The position of normal superior parathyroid glands is more consistent, with 80% of these glands being found near the posterior aspect of the upper and middle thyroid lobes, at the level of the cricoid cartilage. Approximately 1% of normal upper glands may be found in the parasiophageal or retroesophageal space. Enlarged superior glands may descend in the tracheoesophageal groove and come to lie caudal to the inferior glands. Truly ectopic superior parathyroid glands are rare, but they may be found in the middle or posterior mediastinum or in the aortopulmonary window. As the embryo matures, the thymus and inferior parathyroids migrate together caudally in the neck. The most common location for inferior glands is within a distance of 1 cm from a point centered where the inferior thyroid artery and RLN cross. Approximately 15% of inferior glands...
are found in the thymus. The position of the inferior glands, however, tends to be more variable due to their longer migratory path. Undescended inferior glands may be found near the skull base, angle of the mandible, or superior to the upper parathyroid glands along with an undescended thymus. The frequency of intrathyroidal glands is about 2%.

Anatomy and Histology
Most patients have four parathyroid glands. The superior glands usually are dorsal to the RLN at the level of the cricoid cartilage, whereas the inferior parathyroid glands are located ventral to the nerve. Normal parathyroid glands are gray and semitransparent in newborns but appear golden yellow to light brown in adults. Parathyroid color depends on cellularity, fat content, and vascularity. Moreover, they often are embedded in and sometimes difficult to discern from surrounding fat. Normal parathyroid glands are located in loose tissue or fat and are ovoid. They measure up to 7 mm in size and weigh approximately 40 to 50 mg each. Parathyroid glands usually derive their blood supply from branches of the inferior thyroid artery, although branches from the superior thyroid artery supply at least 20% of upper glands. Branches from the thyroidea ima, and vessels to the trachea, esophagus, larynx, and mediastinum may also be found. The parathyroid glands drain ipsilaterally by the superior, middle, and inferior thyroid veins.

Akerström and colleagues, in an autopsy series of 503 cadavers, found four parathyroid glands in 84% of cases. Supernumerary glands were present in 13% of patients, most commonly in the thymus. Only 3% of patients had less than four glands. Similar results were obtained in other dissection studies of 428 human subjects by Gilmour who reported a 6.7% incidence of supernumerary glands.

Histologically, parathyroid glands are composed of chief cells and oxyphil cells arranged in trabeculae, within a stroma composed primarily of adipose cells (Fig. 38-26). The parathyroid glands of infants and children are composed mainly of chief cells, which produce parathyroid hormone (PTH). Acidophilic, mitochondria-rich oxyphil cells are derived from chief cells, can be seen around puberty, and increase in numbers in adulthood. A third group of cells, known as water-clear cells, are also derived from chief cells, are present in small numbers, and are rich in glycogen. Although most oxyphil and water-clear cells retain the ability to secrete PTH, their functional significance is not known.

Parathyroid Physiology and Calcium Homeostasis
Calcium is the most abundant cation in human beings and has several crucial functions. Extracellular calcium levels are 10,000-fold higher than intracellular levels, and both are tightly controlled. Extracellular calcium is important for excitation-contraction coupling in muscle tissues, synaptic transmission in the nervous system, coagulation, and secretion of other hormones. Intracellular calcium is an important second messenger regulating cell division, motility, membrane trafficking, and secretion. Calcium is absorbed from the small intestine in its inorganic form. Calcium fluxes in the steady state are depicted in Fig. 38-27.

Extracellular calcium (900 mg) accounts for only 1% of the body’s calcium stores, the majority of which is sequestered in the skeletal system. Approximately 50% of the serum calcium is in the ionized form, which is the active component. The remainder is bound to albumin (40%) and organic anions such as phosphate and citrate (10%). The total serum calcium levels range from 8.5 to 10.5 mg/dL (2.1 to 2.6 mmol/L), and ionized calcium levels range from 4.4 to 5.2 mg/dL (1.1 to 1.3 mmol/L). Both concentrations are tightly regulated. The total serum calcium level must always be considered in its relationship to plasma protein levels, especially serum albumin. For each gram per deciliter of alteration of serum albumin above or below 4.0 mg/dL, there is a 0.8 mg/dL increase or decrease in protein-bound calcium and, thus, in total serum calcium levels. Total and, particularly, ionized calcium levels are influenced by various hormone systems.

Parathyroid Hormone. The parathyroid cells rely on a G-protein–coupled membrane receptor, designated the calcium-sensing receptor (CASR), to regulate PTH secretion by sensing extracellular calcium levels (Fig. 38-28). PTH secretion also is stimulated by low levels of 1,25-dihydroxy vitamin D, calcitriol (1-hydroxycholecalciferol), and hypomagnesemia. The PTH gene is located on chromosome 11. PTH is synthesized in the parathyroid gland as a precursor hormone proPTH, which is cleaved first to proPTH and then to the final 84-amino-acid PTH. Secreted PTH has a half-life of 2 to 4 minutes. In the liver, PTH is metabolized into the active N-terminal component and the relatively inactive C-terminal fragment. The C-terminal component is excreted by the kidneys and accumulates in chronic renal failure.

PTH functions to regulate calcium levels via its actions on three target organs, the bone, kidney, and gut. PTH increases the resorption of bone by stimulating osteoclasts and promotes the release of calcium and phosphate into the circulation. At the kidney, calcium is primarily absorbed in concert with sodium in the proximal convoluted tubule, but fine adjustments occur more distally. PTH acts to limit calcium excretion at the distal convoluted tubule via an active transport mechanism. PTH also inhibits phosphate reabsorption (at the proximal convoluted tubule) and bicarbonate reabsorption. It also inhibits the Na+/H+ antiporter, which results in a mild metabolic acidosis in hyperparathyroid states. PTH and hypophosphatemia also enhance 1-hydroxylation of 25-hydroxyvitamin D, which is responsible for its indirect effect of increasing intestinal calcium absorption.

Calcitonin. Calcitonin is produced by thyroid C cells and functions as an antihypercalcemic hormone by inhibiting osteoclast-mediated bone resorption. Calcitonin production is stimulated

Figure 38-26. Normal parathyroid histology showing chief cells interspersed with adipose cells.
Vitamin D. Vitamin D refers to vitamin D$_2$ and vitamin D$_3$, both of which are produced by photolysis of naturally occurring sterol precursors. Vitamin D$_3$ is available commercially in pharmaceutical preparations, whereas vitamin D$_2$ is the most important physiologic compound and is produced from 7-dehydrocholesterol, which is found in the skin. Vitamin D is metabolized in the liver to its primary circulating form, 25-hydroxyvitamin D. Further hydroxylation in the kidney results in 1,25-dihydroxy vitamin D, which is the most metabolically active form of vitamin D. Vitamin D stimulates the absorption of calcium and phosphate from the gut and the resorption of calcium from the bone.

Hyperparathyroidism

Hyperfunction of the parathyroid glands may be classified as primary, secondary, or tertiary. PHPT arises from increased PTH production from abnormal parathyroid glands and results from a disturbance of normal feedback control exerted by serum calcium. Elevated PTH levels may also occur as a compensatory response to hypocalcemic states resulting from chronic renal failure or GI malabsorption of calcium. This secondary HPT can be reversed by correction of the underlying problem (e.g., kidney transplantation for chronic renal failure). However, chronically stimulated glands may occasionally become autonomous, resulting in persistence or recurrence of hypercalcemia after successful renal transplantation, resulting in tertiary HPT.

Primary Hyperparathyroidism. PHPT is a common disorder, affecting 100,000 individuals annually in the United States. PHPT occurs in 0.1% to 0.3% of the general population and is more common in women (1:500) than in men (1:2000). Increased PTH production leads to hypercalcemia via increased GI absorption of calcium, increased production of vitamin D$_3$, and reduced renal calcium clearance. PHPT is characterized by increased parathyroid cell proliferation and PTH secretion that is independent of calcium levels.

Etiology. The exact cause of PHPT is unknown, although exposure to low-dose therapeutic ionizing radiation and familial predisposition account for some cases. Various diets and intermittent exposure to sunshine may also be related. Other causes include renal leak of calcium and declining renal function with age as well as alteration in the sensitivity of parathyroid glands to suppression by calcium. The latency period for development of PHPT after radiation exposure is longer than that for the development of thyroid tumors, with most cases occurring 30 to 40 years after exposure. Patients who have been exposed to radiation have similar clinical presentations and calcium levels when compared to patients without a history of radiation exposure. However, the former tends to have higher PTH levels and a higher incidence of concomitant thyroid neoplasms. Lithium therapy has been known to shift the set point for PTH secretion in parathyroid cells, thereby resulting in elevated PTH levels and mild hypercalcemia. Lithium stimulates the growth of abnormal parathyroid glands in vivo and also in susceptible patients in vivo. PHPT results from the enlargement of a single gland or parathyroid adenoma in approximately 80% of cases, multiple adenomas or hyperplasia in 15% to 20% of patients, and parathyroid carcinoma in 1% of patients. Existence of two enlarged glands or double adenomas is supported by biochemical (calcium and PTH), intraoperative PTH (IOPHT), molecular, and histologic data. This entity is less common in younger patients but accounts for up to 10% of older patients with PHPT. It should be emphasized that when more than one abnormal parathyroid gland is identified preoperatively or intraoperatively, the patient has hyperplasia (all glands abnormal) until proven otherwise.

Genetics. Most cases of PHPT are sporadic. However, PHPT also occurs within the spectrum of a number of inherited
disorders such as MEN1, MEN2A, isolated familial HPT, and familial HPT with jaw-tumor syndrome. All of these syndromes are inherited in an autosomal dominant fashion. PHPT is the earliest and most common manifestation of MEN1 and develops in 80% to 100% of patients by age 40 years old. These patients also are prone to pancreatic neuroendocrine tumors and pituitary adenomas and, less commonly, to adrenocortical tumors, lipomas, skin angiomas, and carcinoid tumors of the bronchus, thymus, or stomach. About 50% of patients develop gastrinomas, which often are multiple and metastatic at diagnosis. Insulinomas develop in 10% to 15% of cases, whereas many patients have nonfunctional pancreatic endocrine tumors. Prolactinomas occur in 10% to 50% of MEN1 patients and constitute the most common pituitary lesion. MEN1 has been shown to result from germline mutations in the \textit{MEN1} gene, a tumor suppressor gene located on chromosome 11q12-13 that encodes menin, a protein that is postulated to interact with the transcription factors JunD and nuclear factor-κB in the nucleus, in addition to replication protein A and other proteins.\textsuperscript{70} Most \textit{MEN1} mutations result in a nonfunctional protein and are scattered throughout the translated nine exons of the gene. This makes presymptomatic screening for mutation carriers difficult. \textit{MEN1} mutations also have been found in kindreds initially suspected to represent isolated familial HPT. HPT develops in about 20% of patients with MEN2A and generally is less severe. MEN2A is caused by germline mutations of the \textit{RET} proto-oncogene located on chromosome 10. In contrast to MEN1, genotype-phenotype correlations have been noted in this syndrome in that individuals with mutations at codon 634 are more likely to develop HPT. Patients with the familial HPT with jaw-tumor syndrome have an increased predisposition to parathyroid carcinoma. This syndrome maps to a tumor suppressor locus \textit{HRPT2} (CDC73 or parafibromin) on chromosome 1. Patients belonging to isolated HPT kindreds also appear to demonstrate linkage to \textit{HRPT2}. More recently, a subset of patients with MEN-1 phenotype in the absence of \textit{MENIN} mutations were found to harbor inactivating mutation

Figure 38-28. Regulation of calcium homeostasis. The calcium-sensing receptor (CASR) is expressed on the surface of the parathyroid cell and senses fluctuations in the concentration of extracellular calcium. Activation of the receptor is thought to increase intracellular calcium levels, which, in turn, inhibit parathyroid hormone (PTH) secretion via posttranslational mechanisms. Increased PTH secretion leads to an increase in serum calcium levels by increasing bone resorption and enhancing renal calcium reabsorption. PTH also stimulates renal 1-α-hydroxylase activity, leading to an increase in 1,25-dihydroxy vitamin D, which also exerts a negative feedback on PTH secretion. PKC = protein kinase C; PLC = phospholipase C. (Reproduced with permission from Carling T: Molecular pathology of parathyroid tumors, Trends Endocrinol Metab. 2001 Mar;12(2):53-58.)
in the tumor suppressor gene \textit{CDKN1B} on chromosome 12p13 and given the diagnosis of MEN4.\textsuperscript{71-72} \textit{CDKN1B} encodes p27kip1, which is involved in cyclin D1 signaling. Approximately 25% to 40% of sporadic parathyroid adenomas and some hyperplastic parathyroid glands have loss of heterozygosity (LOH) at 11q13, the site of the \textit{MEN1} gene. The parathyroid adenoma 1 oncogene (\textit{PRAD1} or \textit{CCND1}), which encodes cyclin D1, a cell cycle control protein, is overexpressed in about 18% of parathyroid adenomas. This was demonstrated to result from a rearrangement on chromosome 11 that places the \textit{PRAD1} gene under the control of the PTH promoter. Sporadic parathyroid tumors also appear to carry alterations in cyclin dependent kinase inhibitor encoding genes, in particular somatic inactivating mutations of \textit{CDKN1B}. Other chromosomal regions deleted in parathyroid adenomas and possibly reflecting loss of tumor suppressor genes include 1p, 6q, and 15q, whereas amplified regions suggesting oncogenes have been identified at 16p and 19p. \textit{RET} mutations are rare in sporadic parathyroid tumors. Sporadic parathyroid cancers are characterized by uniform loss of the tumor suppressor gene \textit{RB}, which is involved in cell cycle regulation, and 60% have \textit{HRPT2} (CDC73) mutations. These alterations are rare in benign parathyroid tumors and may have implications for diagnosis. The \textit{p53} tumor suppressor gene is also inactivated in a subset (30%) of parathyroid carcinomas.\textsuperscript{73}

\textbf{Clinical Manifestations} Patients with PHPT formerly presented with the “classic” pentad of symptoms (i.e., kidney stones, painful bones, abdominal groans, psychic moans, and fatigue overtones). With the advent and widespread use of automated blood analyzers in the early 1970s, there has been an alteration in the “typical” patient with PHPT. They are more likely to be minimally symptomatic or asymptomatic. Currently, most patients present with weakness, fatigue, polydipsia, polyuria, nocturia, bone and joint pain, constipation, decreased appetite, nausea, heartburn, pruritus, depression, and memory loss. Patients with PHPT also tend to score lower than healthy controls when assessed by general multidimensional health assessment tools such as the Medical Outcomes Study Short-Form Health Survey (SF-36) and other specific questionnaires. Furthermore, these symptoms and signs improve in most, but certainly not all, patients after parathyroidectomy. Truly “asymptomatic” PHPT appears to be rare, occurring in <5% of patients, as determined by prospectively administered questionnaires. Complications of PHPT are described in the following section.

\textbf{Renal Disease.} Approximately 80% of patients with PHPT have some degree of renal dysfunction or symptoms. Kidney stones were previously reported in up to 80% of patients but now occur in about 20% to 25%. The calculi are typically composed of calcium phosphate or oxalate. In contrast, PHPT is found to be the underlying disorder in only 3% of patients presenting with nephrolithiasis. \textit{Nephrocalcinosis}, which refers to renal parenchymal calcification, is found in <5% of patients and is more likely to lead to renal dysfunction. Chronic hypercalcemia also can impair concentrating ability, thereby resulting in polyuria, polydipsia, and nocturia. The incidence of hypertension is variable but has been reported to occur in up to 50% of patients with PHPT. Hypertension appears to be more common in older patients and correlates with the magnitude of renal dysfunction and, in contrast to other symptoms, is least likely to improve after parathyroidectomy.

\textbf{Bone Disease.} Bone disease, including osteopenia, osteoporosis, and osteitis fibrosa cystica, is found in about 15% of patients with PHPT. Increased bone turnover, as found in patients with osteitis fibrosa cystica, can be determined by documenting an elevated blood alkaline phosphatase level. Advanced PHPT with osteitis fibrosa cystica now occurs in <5% of patients. It has pathognomonic radiologic findings, which are best seen on X-rays of the hands and are characterized by subperiosteal resorption (most apparent on the radial aspect of the middle phalanx of the second and third fingers), bone cysts, and tufting of the distal phalanges (Fig. 38-29). The skull also may be affected and appears mottled with a loss of definition of the inner and outer cortices. Brown or osteoclastic tumors and bone cysts also may be present. Severe bone disease, resulting in bone pain and tenderness and/or pathologic fractures, is rarely observed nowadays. However, reductions of bone mineral density (BMD) with osteopenia and osteoporosis are more common. Patients with normal serum alkaline phosphatase levels almost never have clinically apparent osteitis fibrosa cystica. HPT typically results in a loss of bone mass at sites of cortical bone such as the radius and relative preservation of cancellous bone such as that located at the vertebral bodies. Patients with PHPT, however, also may

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{hand-xray.png}
\caption{X-ray of the hand showing subperiosteal bone resorption most apparent along the radial aspect of the middle phalanx, characteristic of osteitis fibrosa cystica.}
\end{figure}
have osteoporosis of the lumbar spine that improves dramatically following parathyroidectomy. Fractures also occur more frequently in patients with PHPT, and the incidence of fractures also decreases after parathyroidectomy. Bone disease correlates with serum PTH and vitamin D levels.

Gastrointestinal Complications. PHPT has been associated with peptic ulcer disease. In experimental animals, hypergastrinemia has been shown to result from PTH infusion into blood vessels supplying the stomach, independent of its effects on serum calcium. An increased incidence of pancreatitis also has been reported in patients with PHPT, although this appears to occur only in patients with profound hypercalcemia (Ca\(^{2+}\) >12.5 mg/dL). Patients with PHPT also have an increased incidence of choledocholithiasis, presumably due to an increase in biliary calcium, which leads to the formation of calcium bilirubinate stones.

Neuropsychiatric Complications. Severe hypercalcemia may lead to various neuropsychiatric manifestations such as florid psychosis, obtundation, or coma. Other findings such as depression, anxiety, and fatigue are more commonly observed in patients with only mild hypercalcemia. The etiology of these symptoms is not known. Studies demonstrate that levels of certain neurotransmitters (monoamine metabolites 5-hydroxyindoleacetic acid and homovanillic acid) are reduced in the cerebrospinal fluid of patients with PHPT when compared to controls. Electroencephalogram abnormalities also occur in patients with primary and secondary HPT and normalize following parathyroidectomy.

Other Features. PHPT also can lead to fatigue and muscle weakness, which is prominent in the proximal muscle groups. Although the exact etiology of this finding is not known, muscle biopsy studies show that weakness results from a neuromyopathy, rather than a primary myopathic abnormality. Patients with HPT also have an increased incidence of chondrocalcinosis, gout, and pseudogout, with deposition of uric acid and calcium pyrophosphate crystals in the joints. Calcification at ectopic sites such as blood vessels, cardiac valves, and skin also has been reported, as has hypertrophy of the left ventricle independent of the presence of hypertension. There is also evidence for subtle cardiovascular manifestations in mild disease, such as changes in endothelial function, increased vascular stiffness, and perhaps subtle diastolic dysfunction. Several large studies from Europe also suggest that PHPT is associated with increased death rates from cardiovascular disease and cancer even in patients with mild HPT, although this finding was not substantiated in North American studies.

Physical Findings Parathyroid tumors are seldom palpable, except in patients with profound hypercalcemia or parathyroid cancer. A palpable neck mass in a patient with PHPT is more likely to be thyroid in origin or a parathyroid cancer. Patients also may demonstrate evidence of band keratopathy, a deposition of calcium in Bowman’s membrane just inside the iris of the eye. This nonspecific condition generally is caused by chronic eye diseases such as uveitis, glaucoma, and trauma but also may occur in the presence of conditions associated with high calcium or phosphate levels. Fibro-osseous jaw tumors, and/or the presence of familial disease in patients with PHPT and jaw tumors, if present, should alert the physician to the possibility of parathyroid carcinoma.

Differential Diagnosis Hypercalcemia may be caused by a multitude of conditions, as listed in Table 38-9. PHPT and malignancy account for >90% of all cases of hypercalcemia. PHPT is more common in the outpatient setting, whereas malignancy is the leading cause of hypercalcemia in hospitalized patients. PHPT can virtually always be distinguished from other diseases causing hypercalcemia by a combination of history, physical examination, and appropriate laboratory investigations.

Hypercalcemia associated with malignancy includes three distinct syndromes. Although bone metastases may cause hypercalcemia, patients with solid tumors of the lung, breast, kidney, head and neck, and ovary often have humoral hypercalcemia of malignancy, without any associated bony metastases. In addition, hypercalcemia also may be associated with hematologic malignancies such as multiple myeloma. Humoral hypercalcemia of malignancy is known to be mediated primarily by PTH-related peptide (PTHrP), which also plays a role in the hypercalcemia associated with bone metastases and multiple myeloma.

Thiazide diuretics cause hypercalcemia by decreasing renal clearance of calcium. This corrects in normal patients within days to weeks after discontinuing the diuretic, but patients with PHPT continue to be hypercalcemic. Thiazide diuretics can, therefore, exacerbate underlying PHPT and can be used to unmask PHPT in patients with borderline hypercalcemia. Familial hypocalciuric hypercalcemia 1 (FHH1) is a rare autosomal dominant condition with nearly 100% penetrance and results from inherited heterozygous mutations in the CASR gene located on chromosome 3. Homozygous germline mutations at this locus result in neonatal severe primary hyperparathyroidism and calcemia, a condition that can rapidly prove fatal. Patients with FHH1 generally have lifelong hypercalcemia, which is not corrected by parathyroidectomy. A milder form of the disease known as familial hypocalciuric hypercalcemia results from germline inactivating mutations in the intracytoplasmic tail domain of the CaSR gene. These patients have an appropriate hypercalciuric response to elevated calcium and PTH in addition to hypermagnesemia and hyperphosphaturia. Some cases benefit from parathyroidectomy. Recently two new types of FHH (2 and 3) have been described. These are associated with germline inactivating mutations of GNA11(19p13.3) and AP2S1(19q12.2) genes. Both mutations cause hypocalciuric hypercalcemia through aberrant inactivation of CaSR signaling. Although clinical presentation in FHH2 is similar to FHH1, those with FHH3 tend to have higher PTH levels and osteomalacia.  

**Table 38-9**

<table>
<thead>
<tr>
<th>Differential diagnosis of hypercalcemia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperparathyroidism</td>
</tr>
<tr>
<td>Malignancy—hematologic (multiple myeloma), solid tumors (due to PTHrP)</td>
</tr>
<tr>
<td>Endocrine diseases—hyperthyroidism, Addisonian crisis, VIPoma</td>
</tr>
<tr>
<td>Granulomatous diseases—sarcoïdosis, tuberculosis, berylliosis, histoplasmosis</td>
</tr>
<tr>
<td>Milk-alkali syndrome</td>
</tr>
<tr>
<td>Drugs—thiazide diuretics, lithium, vitamin A or D intoxication</td>
</tr>
<tr>
<td>Familial hypocalciuric hypercalcemia</td>
</tr>
<tr>
<td>Paget’s disease</td>
</tr>
<tr>
<td>Immobilization</td>
</tr>
</tbody>
</table>

PTHrP = parathyroid hormone-related protein; VIP = vasoactive intestinal peptide.
Hypercalcemia also is found in approximately 10% of patients with sarcoidosis secondary to increased 25-hydroxy vitamin D 1-hydroxylase activity in lymphoid tissue and pulmonary macrophages, which is not subject to inhibitory feedback control by serum calcium. Thyroid hormone also has bone-resorption properties, thus causing hypercalcemia in thyrotoxic states, especially in immobilized patients. Hemoconcentration appears to be an important factor in the hypercalcemia associated with adrenal insufficiency and pheochromocytoma, although the latter patients may have associated parathyroid tumors (MEN2A), and some pheochromocytomas are known to secrete PTHrP. Other endocrine lesions such as vasoactive intestinal peptide–secreting tumors may be associated with hypercalcemia due to increased secretion of PTHrP. Milk-alkali syndrome requires the ingestion of large quantities of calcium with an absorbable alkali such as that used in the treatment of peptic ulcer disease with antacids. Ingestions of large quantities of vitamins D and A are infrequent causes of hypercalcemia, as is immobilization.

**Diagnostic Investigations**

**Biochemical Studies.** The presence of an elevated serum calcium and intact PTH or two-site PTH levels, without hypocalciuria, establishes the diagnosis of PHPT with virtual certainty. These sensitive PTH assays use immunoradiometric or immunoluminescent techniques and can reliably distinguish PHPT from other causes of hypercalcemia. Furthermore, they do not cross-react with PTHrP (Fig. 38-30). In patients with metastatic cancer and hypercalcemia, intact PTH levels help to determine whether the patient also has concurrent PHPT. Although extremely rare, a patient with hypercalcemia may have a tumor that secretes PTH. FNAB of such a tumor for PTH levels or selective venous catheterization of the veins draining such tumors can help clarify the diagnosis.

Patients with PHPT also typically have decreased serum phosphate (~50%) and elevated 24-hour urinary calcium concentrations (~60%). A mild hyperchloremic metabolic acidosis also is present (80%), thereby leading to an elevated chloride-to-phosphate ratio (~33). Urinary calcium levels need not be measured routinely, except in patients who have not had previously documented normocalcemia or have a family history of hypercalcemia to rule out FHH. In patients with FHH, 24-hour urinary calcium excretion is characteristically low (<100 mg/d). Furthermore, the serum calcium-to-creatinine clearance ratio (24-hour urine calcium/plasma total calcium/24-hour urine creatinine/plasma creatinine) usually is <0.01 in patients with FHH, whereas it is typically >0.02 in patients with PHPT, although there are exceptions to this. Other biochemical features of PHPT are listed in Table 38-10. Elevated levels of alkaline phosphatase may be found in approximately 10% of patients with PHPT and are indicative of high-turnover bone disease. These patients are prone to developing postoperative hypocalcemia due to bone hunger. Serum and urine protein electrophoresis may be necessary to exclude multiple myeloma.

**Table 38-10**

**Biochemical features of primary hyperparathyroidism**

<table>
<thead>
<tr>
<th>SERUM TESTS</th>
<th>ALTERATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calcium</td>
<td>Increased, except in normocalcemic primary hyperparathyroidism</td>
</tr>
<tr>
<td>Intact PTH</td>
<td>Increased or inappropriately high</td>
</tr>
<tr>
<td>Chloride</td>
<td>Increased or high normal</td>
</tr>
<tr>
<td>Phosphate</td>
<td>Decreased or low normal</td>
</tr>
<tr>
<td>Chloride-to-phosphate ratio</td>
<td>Increased (usually &gt;33)</td>
</tr>
<tr>
<td>Magnesium</td>
<td>Unchanged or decreased (in patients with osteitis fibrosa cystica)</td>
</tr>
<tr>
<td>Uric acid</td>
<td>Normal or increased</td>
</tr>
<tr>
<td>Alkaline phosphatase</td>
<td>Normal or increased (in the presence of high turnover bone disease)</td>
</tr>
<tr>
<td>Acid-base status</td>
<td>Mild hyperchloremic metabolic acidosis</td>
</tr>
<tr>
<td>Calcium-to-creatinine clearance ratio</td>
<td>Generally &gt;0.02 (vs. &lt;0.01 in FHH) but there are exceptions</td>
</tr>
<tr>
<td>1,25-dihydroxy vitamin D</td>
<td>Normal or increased</td>
</tr>
</tbody>
</table>

**Urine tests**

<table>
<thead>
<tr>
<th>TEST</th>
<th>ALTERATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>24-h urinary calcium</td>
<td>Normal or increased</td>
</tr>
</tbody>
</table>

BFHH = benign familial hypocalciuric hypercalcemia; PTH = parathyroid hormone.

---

Figure 38-30. Intact parathyroid hormone (PTH) measurement allows differentiation between the various causes of hypercalcemia. (Reproduced with permission from Endres DB, Villanueva R, Sharp CF, et al. Measurement of parathyroid hormone. Endocrinol Metab Clin North Am. 1989 Sep;18(3):611-629.)
Occasionally, patients present with normocalcemic PHPT due to vitamin D deficiency, a low serum albumin, excessive hydration, a high-phosphate diet, or a low normal blood calcium set point. These patients have increased total PTH levels with or without increased blood ionized calcium levels and must be distinguished from patients with renal leak hypercalcemia who also have increased PTH levels due to excessive calcium loss in the urine. This can be accomplished by administering thiazide diuretics. In patients with idiopathic hypercalcemia, the urinary calcium level falls, and the secondary increase in the blood PTH level also decreases to normal, whereas patients with normocalcemic HPT continue to have elevated urine calcium and blood PTH levels and may, in fact, become hypercalcemic.

**Radiologic Tests.** In patients with profound hypercalcemia or PHPT associated with vitamin D deficiency, hand and skull X-rays may demonstrate osteitis fibrosa cystica, but this is rare in current clinical practice. BMD studies using dual-energy absorptiometry are being increasingly used to assess the effects of PHPT on bone. PHPT primarily leads to one loss at cortical sites such as the distal radius while bone density is preserved at sites such as the lumbar spine. Current evaluation of patients with PHPT includes vertebral imaging by X-ray or vertebral fracture assessment (VFA) or CT scan in addition to BMD studies. Measurement of trabecular bone score (TBS) is optional. In addition, renal imaging by ultrasound, X-ray, or CT scan is also recommended. Parathyroid localization studies are not used to confirm the diagnosis of PHPT, but rather to aid in identifying the location of the offending gland(s), as discussed later in “Preoperative Localization Tests.”

**Treatment**

**Indications for Parathyroidectomy and Role of Medical Management.** Most authorities agree that patients who have developed complications and have “classic” symptoms of PHPT should undergo parathyroidectomy. However, the treatment of patients with asymptomatic PHPT has been the subject of controversy, due, in part, to the fact that there is little agreement on what constitutes an asymptomatic patient.

At the National Institutes of Health consensus conference in 1990, “asymptomatic” PHPT was defined as “the absence of common symptoms and signs of PHPT, including no bone, renal, gastrointestinal, or neuromuscular disorders.” To determine the best course of action for these patients, it is important to consider the natural history of untreated PHPT and the outcomes of treatment options, both medical and surgical.

With respect to the natural history, the panel advocated nonoperative management of these patients with mild PHPT based on observational studies, which suggested relative stability of biochemical parameters over time. However, the consensus panel considered certain patients to be candidates for surgery based on testing or other information indicating end-organ effects or a higher likelihood of disease progression, and this led to the establishment of initial guidelines for parathyroidectomy. Subsequently, another observational study on the natural history of treated versus untreated HPT was published by Silverberg and colleagues. Their cohort of 52 patients with asymptomatic HPT followed without surgery, levels of serum and urinary calcium, PTH, alkaline phosphatase, and vitamin D metabolites remained relatively stable over a 10-year period in most patients. Average bone mass also remained relatively stable. However, the study also reported development of a new indication for surgery in 14 (27%) of 52 of their asymptomatic patients and, because approximately 50% of their patients were initially treated surgically, overall, about 75% of patients were underwent parathyroidectomy. Age <50 years was predictive of progression, and patients undergoing parathyroidectomy showed not only normalization of calcium and PTH levels but also improved BMD at the spine and hip. Based on these and other studies, the guidelines were reassessed at a second workshop on asymptomatic PHPT held at the National Institutes of Health in 2002.

Since that time, additional studies have provided further insights into the natural history of treated and untreated HPT. Three of these were randomized, controlled, prospective studies ranging in duration from 1 to 3.5 years. One was an observational study (a continuation of the Columbia University PHPT Project) but was notable for its long duration of follow-up of 15 years. These studies confirmed the relative stability of various biochemical indices, thus validating the need for guidelines. However, the long-term study suggested that the stability was not indefinite as calcium levels tended to rise in years 13 to 15. In addition, the study also demonstrated that bone density measurements remained stable for 8 to 10 years, but cortical bone density worsened after year 10. More concerning was the fact that 60% of patients lost >10% of their BMD over the 15-year observation period. Furthermore, whether patients met the 2002 guidelines for surgery did not appear to predict the risk of progressive disease, with 40% of patients undergoing follow-up eventually needing surgery. Although there are no randomized trials, registry data also suggest that fracture risk is increased for PHPT up to 10 years prior to diagnosis and treatment.

Medical options for treating PHPT and its complications include antiresorptive treatments such as bisphosphonates, hormone replacement therapy (HRT), and selective estrogen receptor modulators such as raloxifene. Bisphosphonates and HRT are reasonable options in patients for whom skeletal protection is needed, as evidence from randomized, placebo-controlled trials indicates that these medications are very effective at decreasing bone turnover and increasing BMD in PHPT, with the effects being comparable to patients undergoing parathyroidectomy. Caution needs to be exercised due to the nonskeletal effects of HRT, and hence, bisphosphonates are preferred. There are no clinical studies regarding the effects of raloxifene on BMD in HPT, and none of these agents affects calcium or PTH levels. More recently, calcimimetics (modifiers of the sensitivity of the CASR) have been used in randomized, multicenter controlled trials and have been shown to decrease both serum calcium and PTH levels in both symptomatic and asymptomatic PHPT patients. Unfortunately, bone density failed to improve in medically treated patients. Although this therapy shows promise, long-term outcome data are lacking, and their routine use is not advocated at this time, except in patients who are very poor operative risks or refuse surgery.

Successful parathyroidectomy results in resolution of osteitis fibrosa cystica and decreased formation of renal stones in symptomatic (classic) patients. In addition, it results in improved BMD (6% to 8% in the first year and up to 12% to 15% at 15 years) and fracture risk (by 50% at hip and upper arm and 30% overall) after adjustment for age, sex, and previous fractures over a 20-year observation period. There are also data to show that it improves a number of the nonspecific manifestations of PHPT such as fatigue, polydipsia, polyuria and nocturia.
bone and joint pain, constipation, nausea, and depression in many patients. This also has been demonstrated using symptom questionnaires and various standardized general quality-of-life assessments such as the SF-36 and a specific parathyroidectomy assessment of symptoms scale. The increased death rate in patients with PHPT appears to be reversible by successful parathyroidectomy, at least in some studies. Lastly, parathyroidectomy can be accomplished with >95% success rates with minimal morbidity, even in elderly patients and is the only curative treatment option for PHPT. Previous investigations have also documented that parathyroidectomy is more cost-effective than medical management or follow-up.

Given these findings, it is recommended that parathyroidectomy should be offered to virtually all patients except those in whom the operative risks are prohibitive. This is also acknowledged by the panel of the latest workshop, which stated that “even though patients may not meet the guidelines for surgical intervention, it is always a reasonable option in those who do not have medical contraindications.” This was first stated in the 2008 guidelines and reiterated in the most recent revision in 2014, which advised parathyroidectomy for patients with smaller elevations in serum calcium levels (>1 mg/dL above the upper limit of normal) and if BMD measured at any of three sites (radius, spine, or hip) is greater than 2.5 standard deviations below those of gender- and race-matched, not age-matched, controls (i.e., peak bone density or T score [rather than Z score] <2.5). In addition, patients <50 years of age were advised to undergo parathyroidectomy. Parathyroidectomy is also indicated for creatinine clearance <60 cc/minute and urine calcium >400 mg/day in the presence of increased stone risk by biochemical stone risk analysis. The significant changes from the previous guidelines pertain to the fact that (a) patients with nephrolithiasis or nephrocalcinosis by X-ray, ultrasound, or CT scan and (b) those with vertebral fracture by X-ray, CT, MRI, or VFA are also candidates for parathyroidectomy. The current guidelines are summarized in Table 38-11.

It is important to point out that the neurocognitive and neuropsychological aspects of PHPT remain a topic of controversy with respect to the guidelines for parathyroidectomy. Although there were more studies since the previous iteration of the guidelines, there were concerns that while some lacked adequate controls and were plagued by problems related to the instruments used to quantify these nonspecific symptoms, others showed variability in improvement of neurocognitive symptoms following parathyroidectomy. Similarly, uncertainty is also present concerning the cardiovascular consequences of mild HPT. Therefore, the workshop panel emphasizes that these criteria alone should not be used as guidelines for surgical intervention.

Since there are no definitive criteria to indicate which patients with mild PHPT will develop progressive disease, more clinical studies are required. Patients who do not undergo surgery should undergo routine follow-up as outlined in the recent workshop summary statement, consisting of annual calcium and serum creatinine measurements, and measurements of BMD at three sites every 1 to 2 years.

Table 38-11

<table>
<thead>
<tr>
<th>Indications for parathyroidectomy in patients with asymptomatic primary HPT (2014 NIH consensus conference guidelines)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Serum calcium &gt;1 mg/dL above the upper limits of normal</td>
</tr>
<tr>
<td>• GFR &lt;60 mL/min; 24-h urine for calcium &gt;400 mg/d (&gt;10 mmol/d) and increased stone risk by biochemical stone risk analysis</td>
</tr>
<tr>
<td>• Presence of nephrolithiasis or nephrocalcinosis by X-ray, ultrasound, or CT</td>
</tr>
<tr>
<td>• Substantially decreased bone mineral density at the lumbar spine, total hip, femoral neck, or distal radius (&gt;2.5 SD below peak bone mass, T score &lt;−2.5; vertebral fracture by X-ray, CT, MRI, or VFA)</td>
</tr>
<tr>
<td>• Age &lt;50 y</td>
</tr>
<tr>
<td>• Long-term medical surveillance not desired or possible</td>
</tr>
</tbody>
</table>

GFR = glomerular filtration rate; HPT = hyperparathyroidism; NIH = National Institutes of Health; SD = standard deviation; VFA = vertebral fracture assessment

Preoperative Localization Tests. Localization studies may be classified into noninvasive or invasive modalities. These studies have variable performance characteristics, which, in turn, vary with operator and institutional experience, as outlined in Table 38-12. Localization studies have permitted surgeons to perform more limited operations, some of them under local anesthesia. These “minimally invasive” procedures include unilateral and focused neck exploration, radio-guided parathyroidectomy, and several endoscopic or video-assisted approaches. The use of localization studies has been shown in some studies to be associated with lower morbidity rates (hypoparathyroidism and RLN injury) and decreased operative times, reduced duration of hospital stay, and improved cosmetic outcomes, while maintaining success rates similar to those obtained with traditional bilateral neck explorations. Some studies also show that use of localization studies may be more cost-effective. Overall, it has become routine to localize hyperfunctioning parathyroid glands before parathyroidectomy. It is important to point out that imaging is not a diagnostic approach, and the decision for exploration should be made before any imaging is performed. $^{99m}$Tc-labeled sestamibi (Fig. 38-31A) is the most widely used and accurate modality with a sensitivity >80% for detection of parathyroid adenomas. Sestamibi (Cardiolite) initially was introduced for cardiac imaging and is concentrated in mitochondria-rich tissue. It was subsequently noted to be useful for parathyroid localization due to the delayed washout of the radiouclide from hypercellular parathyroid tissue compared to thyroid tissue. Sestamibi scans generally are complemented by neck ultrasound (Fig. 38-31B), which can identify adenomas with >75% sensitivity in experienced centers and is most useful in identifying intrathyroidal parathyroids. Single-photon emission CT, particularly when used with CT, has been shown to be superior to other nuclear medicine-based imaging. Specifically, single-photon emission CT can indicate whether an adenoma is located in the anterior or posterior mediastinum (aortopulmonary window), thus enabling the surgeon to modify the operative approach accordingly. CT and MRI scans are less sensitive than sestamibi scans, but they are helpful in localizing large paraesophageal and mediastinal glands. More recently, four-dimensional CT (4D-CT) has shown utility in parathyroid localization. This technique incorporates the perfusion of contrast in hyperfunctioning parathyroid tissue over time, thus providing functional information in addition to the anatomic information provided by conventional three-dimensional CT imaging. In one study, 4D-CT showed improved sensitivity of 88% compared to...
that of sestamibi (65%) and ultrasound (57%) for lateralization of the enlarged gland and also showed superiority when localization to the correct quadrant was examined. A combination of 4D-CT and ultrasound has been reported to have a positive predictive value of 92% for single-gland disease and 75% for multiple-gland disease.

IOPTH was initially introduced in 1993 and is used to determine the adequacy of parathyroid resection (Fig. 38-32). IOPTH measurements, like localization studies, are less reliable in multiglandular disease. Bilateral internal jugular vein sampling has also been used to lateralize tumors intraoperatively but is less accurate.

**Operative Approaches** Unilateral parathyroid exploration was first carried out using intraoperative staining of a biopsy from the normal parathyroid gland with Sudan black dye to rule out a double adenoma. Initially, the choice of side to be explored was random, but the introduction of preoperative localization studies has enabled a more directed approach. In contrast, the focused approach identifies only the enlarged parathyroid gland, and no attempts are made to locate other parathyroid glands. Unilateral neck explorations have several advantages over bilateral neck exploration, including reduced operative times and complications, such as injury to the RLN and hypoparathyroidism. However,

| **Table 38-12** Commonly used parathyroid localization studies |
|-----------------|-----------------|-----------------|
| **STUDY**       | **ADVANTAGES**   | **DISADVANTAGES** |
| Preoperative, noninvasive | | |
| Sestamibi-technetium-99m scan | Allows planar and SPECT imaging | False-positive tests due to thyroid neoplasms, lymphadenopathy |
| Ultrasound | Identification of juxta- and intrathyroidal tumors | False-positive results due to thyroid nodules, cysts, lymph nodes, esophageal lesions |
| CT scan | Localization of ectopic (mediastinal) glands | Not useful for juxta- or intrathyroidal glands, False-positive results from lymph nodes, Relatively high cost, Radiation exposure, Requires IV contrast, Interference from shoulders and metallic clips |
| MRI scan | Localization of ectopic tumors | Expensive |
| | No radiation exposure | False-positive results from lymph nodes and thyroid nodules |
| | No IV contrast | Cannot be used in claustrophobic patients |
| Four-dimensional CT scan | Structural and functional information | Similar to CT scan |
| Preoperative, invasive | | |
| FNAB | Can distinguish parathyroid tumor from lymphadenopathy using PTH assay | Experienced cytologist needed |
| Angiogram | Provides a road map for selective venous sampling | Expensive |
| | Treatment of mediastinal tumors by embolization | Experienced radiologist needed |
| Venous sampling | Useful to lateralize tumor in equivocal cases or negative localization studies | Expensive, experienced radiologist needed |
| Intraoperative | | |
| PTH assay | Immediate confirmation of tumor removal | Expensive |
| | | Increased operative time, decreased accuracy in multiple-gland disease |

CT = computed tomography; FNAB = fine-needle aspiration biopsy; IV = intravenous; MRI = magnetic resonance imaging; PTH = parathyroid hormone; SPECT = single-photon emission computed tomography.
most existing studies comparing the two approaches are retrospective and do not analyze the results on an intent-to-treat basis. Another argument against a unilateral exploration is the risk of missing another adenoma on the opposite side of the neck. The incidence of double adenomas has been reported to range from 0% to 10%, with an increased incidence in elderly patients. The risk of missing a second adenoma is higher in populations with a higher incidence of multiple adenomas, such as those with familial HPT, MEN syndromes, and the elderly. Another difficulty inherent with unilateral exploration is the inability to discern whether the combination of an abnormal gland and a normal gland on the initial side constitutes a single adenoma or asymmetric hyperplasia. A recently published update on the 5-year results of a randomized trial comparing unilateral versus bilateral neck exploration did not note any difference in the rates of recurrent or persistent disease in the two groups of patients. These issues will only be resolved by a large, prospective, multicenter study or improved molecular analytic techniques.
Radio-guided parathyroidectomy takes advantage of the ability of parathyroid tumors to retain 99mTc-sestamibi. Before surgery, 1 to 2 mCi of the isotope is injected, and a hand-held gamma probe is used to guide the identification of the enlarged gland, taking care to ensure the equilibration of radioactivity counts in all quadrants. Reported advantages include easier localization, particularly in reoperative cases, and the ability to perform the procedure under local anesthesia or sedation using smaller incisions. Many studies demonstrated the feasibility of this technique; however, it is rarely used now, largely because it offers little advantage over preoperative sestamibi scans and is associated with increased operative times. Like preoperative scanning, it also has reduced accuracy in the presence of multiglandular disease.

Endoscopic approaches include both video-assisted and total endoscopic techniques. Total endoscopic parathyroidectomy was first described by Gagner in 1996, and several other investigators have since reported on this technique. Although port placements are variable, as is the case with endoscopic thyroidectomy, they all involve creation of a working space in the neck using CO₂ insufflation, with the reported advantages being superior cosmesis and excellent visualization. Although feasible, these techniques also have been associated with increased operating times, more personnel, and greater expense, and have, in general, not been useful for patients with multiglandular disease, a large thyroid mass, or previous neck surgery and irradiation. Their greatest use has been in patients with tumors at ectopic sites such as the mediastinum where thoracoscopic parathyroidectomy is an excellent alternative to sternotomy. A bloodless field is important to allow identification of parathyroid glands. The middle thyroid veins are ligated and divided, thus enabling medial and anterior retraction of the thyroid lobe, with the aid of a peanut sponge or placement of 2-0 silk sutures into the thyroid. The space between the carotid sheath and thyroid is then opened by gentle sharp and blunt dissection, from the cricoid cartilage superiorly to the thy- mus inferiorly and the RLN is identified. Approximately 85% of the parathyroid glands are found within 1 cm of the junction of the inferior thyroid artery and RLNs. The upper parathyroid glands usually are superior to this junction and dorsal (posterior) to the nerve, whereas the lower glands are located inferior to the junction and ventral (anterior) to the recurrent nerve. Because parathyroid glands are partly surrounded by fat, any fat lobule at typical parathyroid locations should be explored because the normal or abnormal parathyroid gland may be concealed in the fatty tissue. The thin fascia overlying a “suspicious” fat lobule should be incised using a sharp curved hemostat and scalpel. This maneuver often causes the parathyroid gland to “pop” out. Alternatively, gentle, blunt peanut sponge dissection between the carotid sheath and the thyroid gland often reveals a “float” sign, suggesting the site of the abnormal parathyroid gland. Normal parathyroids are light beige and only slightly darker or brown compared to adjacent fat.

Parathyroid tissue needs to be distinguished from normal or brown fat tissue, thyroid nodules, lymph nodes, and ectopic thymus. Lymph nodes generally are light beige to whitish gray in color, glassy, and multiple in number, whereas thyroid nodules generally are more vascular, firm, dark or reddish brown in color, and have a more variegated appearance. Intraoperatively,
a suspicious nodule may be aspirated using a fine needle attached to a syringe containing 1 cc of saline. Very high PTH levels in the aspirate have been shown to be diagnostic in the intraoperative identification of parathyroid glands. Several characteristics such as size (>7 mm), weight, and color are used to distinguish normal from hypercellular parathyroid glands. Hypercellular glands generally are darker, more firm, and more vascular than normocellular glands. No single method is 100% reliable, and therefore, the parathyroid surgeon must rely on experience and, at times, advice from a pathologist to help distinguish normal from hypercellular glands. Although several molecular studies have shown use in distinguishing parathyroid adenomas from hyperplasia, this determination also must be made by the surgeon intraoperatively by documenting the presence of a normal parathyroid gland.

**Location of Parathyroid Glands.** The majority of lower parathyroid glands are found in proximity to the lower thyroid pole (Fig. 38-33A). If not found at this location, the thyrothymic ligament and thymus should be mobilized. The upper end of the cervical thymus is gently grasped with a right-angle clamp, and the distal portion is bluntly dissected from perithymic fat with a peanut sponge. One can then “walk down” the thymus with successive right-angle clamps (Fig. 38-33B). Applying light tension along with a “twisting” motion helps to free the upper thymus. The carotid sheath also should be opened from the bifurcation to the base of the neck if the parathyroid tumor cannot be found. If these maneuvers are unsuccessful, an intrathyroidal gland should be sought by using intraoperative ultrasound, incising the thyroid capsule on its posterolateral surface, or by performing an ipsilateral thyroid lobectomy and “bread-loafing” the thyroid lobe. Preoperative or intraoperative ultrasonography can be useful for identifying intrathyroidal parathyroid glands. Rarely, the third branchial pouch may maldescend and be found high in the neck (undescended parathyroid), anterior to the carotid bulb, along with the missing parathyroid gland. Upper parathyroid glands are more consistent in position and usually are found near the junction of the upper and middle thirds of the gland, at the level of the cricoid cartilage (Fig. 38-33C). Ectopic upper glands may be found in carotid sheath, tracheoesophageal groove, retroesophageal, or in the posterior mediastinum. The locations of ectopic upper and lower parathyroid glands are shown in Fig. 38-34. Every attempt must be made to identify all four glands. Treatment depends on the number of abnormal glands.

1. A single adenoma is presumed to be the cause of a patient’s PHPT if only one parathyroid tumor is identified and the other parathyroid glands are normal, a situation present in about 80% of patients with PHPT. Adenomas typically have an atrophic rim of normal parathyroid tissue, but this characteristic may be absent. The adenoma is dissected free of surrounding tissue, taking care to stay immediately adjacent to the tumor, without fracturing it. The vascular pedicle is clamped, divided, and ligated. Care should be taken to not rupture the parathyroid gland to decrease the risk of parathyromatosis. If there is any question about the presumed normal glands, one of them should be biopsied and examined by frozen section.

2. If two abnormal and two normal glands are identified, the patient has double adenomas. Triple adenomas are present if three glands are abnormal and one is normal. Multiple adenomas are more common in older patients with an incidence of up to 10% in patients >60 years old. The abnormal glands should be excised, provided the remaining glands are confirmed as such, thus excluding asymmetric hyperplasia after biopsy and frozen section.

3. If all parathyroid glands are enlarged or hypercellular, patients have parathyroid hyperplasia that has been shown to occur in about 15% of patients in various series. These glands are often lobulated, usually lack the rim of normal
parathyroid gland seen in adenomas, and may be variable in size. It often is difficult to distinguish multiple adenomas from hyperplasia with variable gland size. Hyperplasia may be of the chief cell (more common), mixed, or clear cell type. Patients with hyperplasia may be treated by subtotal parathyroidectomy or by total parathyroidectomy and autotransplantation, with the choice of procedure being determined by rates of recurrence, postoperative hypocalcemia, and failure rates of autotransplanted tissue. Initial studies demonstrated equivalent cure rates and postoperative hypocalcemia for the two techniques, with the latter having the added advantage of avoiding recurrence in the neck. However, autotransplanted tissue may fail to function in about 5% of cases.

All four parathyroid glands are identified and carefully mobilized. For patients with hyperplasia, a titanium clip is placed across the most normal gland, leaving a 50-mg remnant and taking care to avoid disturbing the vascular pedicle and that the gland is resected with a sharp scalpel. If possible, it is preferable to subtotally resect an inferior gland, which is more easily accessible in case of recurrence due to its anterior location with respect to the RLN. The resected parathyroid tissue is confirmed by frozen section or PTH assay. If the remnant appears to be viable, the remaining glands are resected. If there is any question as to the viability of the initially subtotally resected gland, another gland is chosen for subtotal resection, and the initial remnant is removed. Whenever multiple parathyroids are resected, it is preferable to cryopreserve tissue, so that it may be autotransplanted should the patient become hypoparathyroid. Parathyroid tissue usually is transplanted into the nondominant forearm. A horizontal skin incision is made overlying the brachioradialis muscle a few centimeters below the antecubital fossa. Pockets are made in the belly of the muscle, and one to two pieces of parathyroid tissue measuring 1 mm each are placed into each pocket. A total of 12 to 14 pieces are transplanted. Autotransplanted tissue also has been reported to function when transplanted into fat.

**Indications for Sternotomy**

A sternotomy is usually not recommended at the initial operation, unless the calcium level is >13 mg/dL. Rather, it is preferred to biopsy the normal glands and subsequently close the patient’s neck and obtain localizing studies, if they were not obtained previously. Intraoperative PTH assay during the operation from large veins may be helpful. Using highly selective venous catheterization postoperatively also may be needed when noninvasive localization studies are negative, equivocal, or conflicting. Lower parathyroid glands tend to migrate into the anterior mediastinum in the thymus or perithymic fat and usually can be approached via a cervical incision. A sternotomy is needed to deliver these tumors in approximately 5% of cases. Generally, the gland can be approached by a partial sternotomy to the third intercostal space. The midline sternotomy can be extended to the left or right side as required. Upper glands tend to migrate to the posterior mediastinum in the tracheoesophageal groove. Mediastinal glands also may be found in the aortopulmonary window or pericardium, or attached to the ascending aorta, aortic arch, or its branches.

**Special Situations**

**Normocalcemic Hyperparathyroidism.** This disorder is becoming increasingly recognized in clinical practice (prevalence from 0.5% to 16%) and is defined by the presence of an elevated PTH level with repeatedly normal calcium (including ionized calcium) levels. In addition, other secondary causes of elevated PTH should be ruled out, namely, vitamin D deficiency, osteomalacia, hypercalcemia (renal leak), and renal insufficiency. Data regarding the natural history of this disorder are limited. In a series of 37 patients, Lowe and colleagues showed that 19% of patients became frankly hypercalcemic within 3 years. In addition, 57% developed osteoporosis, 11% developed fragility fractures, and 14% developed nephrolithiasis. Although the study had some limitations, it led the authors to suggest that normocalcemic HPT may represent a variant of “symptomatic” PHPT and may not be an early form of “asymptomatic” disease. Limited studies show that
Parathyroidectomy is more likely to be unsuccessful in these patients. In the absence of strong data, no guidelines are available for this entity. As such, most clinicians follow a conservative course unless patients progress to the classic hypercalcemic form or develop nephro lithiasis, reduced bone mineral density, or fragility fractures.

**Parathyroid Carcinoma.** Parathyroid cancer accounts for approximately 1% of PHPT cases. It may be suspected preoperatively by the presence of severe symptoms, serum calcium levels >14 mg/dL, significantly elevated PTH levels (five times normal), and a palpable parathyroid gland. Local invasion is quite common; approximately 15% of patients have lymph node metastases, and 33% have distant metastases at presentation. Intraoperatively, parathyroid cancer is suspected by the presence of a large, gray-white to gray-brown parathyroid tumor that is adherent to or invasive into surrounding tissues like muscle, thyroid, RLN, trachea, or esophagus. Enlarged lymph nodes also may be present. Frozen sections are generally unreliable. Accurate diagnosis necessitates histologic examination. The major diagnostic criteria include vascular or capsular invasion, trabecular or fibrous stroma, and frequent mitoses. It is, however, important to emphasize that these classic findings are not as frequently noted as previously reported, and some may be found in benign adenomas as well.

Treatment of parathyroid cancer consists of neck exploration, with en bloc excision of the tumor and the ipsilateral thyroid lobe, in addition to the removal of contiguous lymph nodes (tracheoesophageal, paratracheal, and upper mediastinal). The recurrent nerve is not sacrificed unless it is directly involved with tumor. Adherent soft tissue structures (strap muscles or other soft tissues) should also be resected. Modified radical neck dissection is recommended in the presence of lateral lymph node metastases. Prophylactic neck dissection is not advised. If the diagnosis is made postoperatively, a decision must be made regarding the adequacy of initial surgery based on a review of operative notes, pathology reports, localization studies, and calcium and PTH levels. If any question exists, histologic review by another experienced pathologist can be helpful. Additional procedures can include ipsilateral thyroid lobectomy with resection of contiguous structures and lymph nodes if the features are typical or the patient remains hypercalcemic. Patients with equivocal pathologic findings and normocalcemia may be monitored closely. Reoperation is indicated for locally recurrent or metastatic disease to control hypercalcemia. Adjuvant radiation therapy should be considered in patients at high risk of local recurrence such as those with close or positive margins, invasion of surrounding structures, or tumor rupture. Radiation may also be used as primary therapy in unresectable disease or for palliation of bone metastases. Chemotherapy is not very effective. Bisphosphonates have shown some effectiveness in treating hypercalcemia associated with parathyroid carcinoma. Cinacalcet hydrochloride, a calcimimetic, can reduce PTH levels by directly binding to the CASR cells on the parathyroid gland and has been shown to be useful in controlling hypercalcemia in patients with refractory parathyroid carcinoma. Other promising approaches include antiparathyroid hormone immunotherapy, octreotide, and the telomerase inhibitor azidothymidine, but additional investigations are needed in this area.

**Familial Hyperparathyroidism.** PHPT may occur as a component of various inherited syndromes such as MEN1 and MEN2A. Inherited PHPT also can occur as isolated familial HPT (non-MEN) or familial HPT with jaw tumors. The diagnosis of familial HPT is known or suspected in approximately 85% of patients preoperatively. Furthermore, patients with hereditary HPT generally have a higher incidence of multiglandular disease, supernumerary glands, and recurrent or persistent disease. Therefore, these patients warrant a more aggressive approach and are not candidates for various focused surgical approaches. Although not absolutely necessary, preoperative sestamibi scan and ultrasound can be obtained in patients with inherited HPT to identify potential ectopic glands. A standard bilateral neck exploration is performed, along with a bilateral cervical thymectomy, regardless of the results of localization studies. Both subtotal parathyroidectomy and total parathyroidectomy with autotransplantation are appropriate, and parathyroid tissue also should be cryopreserved. If an adenoma is found in patients with familial HPT, the adenoma and the ipsilateral normal parathyroid glands are resected. The normal-appearing glands on the contralateral side are biopsied and marked, so that only one side of the neck will need to be explored in the event of recurrence. Patients with MEN2A require total thyroidectomy and central neck dissection for prevention/treatment of MTC, a procedure that places the parathyroids at risk. Moreover, HPT is less aggressive in these patients. Hence, only abnormal parathyroid glands need to be resected at neck exploration. The other normal parathyroid glands should be marked with a clip.

**Neonatal Hyperparathyroidism.** Infants with neonatal HPT present with severe hypercalcemia, lethargy, hypotonia, and mental retardation. This disorder is associated with homozygous mutations in the CASR gene. As indicated earlier, urgent total parathyroidectomy (with autotransplantation and cryopreservation) and thymectomy are indicated. Subtotal resection is associated with high recurrence rates.

**Parathyromatosis.** Parathyromatosis is a rare condition characterized by the finding of multiple nodules of hyperfunctioning parathyroid tissue throughout the neck and mediastinum, usually following a previous parathyroidectomy. The true etiology of parathyromatosis is not known. It is postulated to arise either from overgrowth of congenital parathyroid rests (ontogenous parathyromatosis) or seeding at surgery from rupture of parathyroid tumors or subtotal resection of hyperplastic glands. Parathyromatosis represents a rare cause of persistent or recurrent HPT and can be identified intraoperatively. Aggressive local resection of these deposits can result in normocalcemia but is rarely curative. Some studies suggest that these patients have low-grade carcinoma because of invasion into muscle and other structures distant from the resected parathyroid tumor.

**Postoperative Care and Follow-Up.** Patients who have undergone parathyroidectomy are advised to undergo calcium level checks 2 weeks postoperatively, at 6 months, and then annually. Recurrences are rare (<1%), except in patients with familial HPT. Recurrence rates of 15% at 2 years and 67% at 8 years have been reported for MEN1 patients.

**Persistent and Recurrent Hyperparathyroidism.** Persistence is defined as hypercalcemia that fails to resolve after parathyroidectomy and is more common than recurrence, which refers to HPT occurring after an intervening period of at least 6 months of biochemically documented normocalcemia. Recurrent disease is far less common than persistent HPT; however, both occur more frequently in the setting of familial HPT and MEN1, in particular. The most common causes for both these states include ectopic parathyroids, unrecognized
hyperplasia, or supernumerary glands. More rare causes include parathyroid carcinoma, missed adenoma in a normal position, incomplete resection of an abnormal gland, parathyromatosis, or an inexperienced surgeon. The most common sites of ectopic parathyroid glands in patients with persistent or recurrent HPT are paraesophageal (28%), mediastinal (26%), intrathyroidal (11%), carotid sheath (9%), and high cervical or undescended (2%) (Fig. 38-35).

Once the diagnosis of persistent or recurrent HPT is suspected, it should be confirmed by the necessary biochemical tests. Other causes of an elevated serum PTH such as renal insufficiency, renal calcium leak, and GI tract abnormalities should be considered. A detailed family history should be performed to screen for familial disease, as this will influence the operative approach. In particular, a 24-hour urine collection should be performed to rule out FHH. In redo-parathyroid surgery, the glands are more likely to be in ectopic locations, and postoperative scarring tends to make the procedure more technically demanding. Cure rates are generally lower (80–90% compared with 95–99% for initial operation), and risk of injury to RLNs and permanent hypocalcemia are higher. Therefore, an evaluation of severity of HPT and the patient’s anesthetic risk (using the American Society of Anesthesiology classification of physical status or the Goldman cardiac index) is important. There are no published guidelines directly applicable to this group of patients. In general, patients with significant and ongoing problems such as recurrent kidney stones, a markedly elevated calcium level, or ongoing bone loss will need reexploration. Patients in whom the diagnosis remains in question or those with equivocal or minimal symptoms may be considered for conservative management. Preoperative localization studies are routinely performed. Noninvasive studies such as a sestamibi scan and ultrasound are obtained, supplemented by 4D-CT scans. If these studies are negative, discordant, or equivocal, obtaining an ultrasound-guided aspirate of a suspicious cervical lesion or a highly selective venous catheterization for PTH levels (by an experienced angiographer) is recommended. Previous operative notes and pathology reports should be carefully reviewed and reconciled with the information obtained from localization studies before any neck reexploration. An algorithm for the treatment of patients with recurrent and persistent HPT is shown in Fig. 38-36.

Generally, these patients are approached with a focused exploration. The lateral approach is frequently used and can be achieved via the previous incision. The plane between the sternocleidomastoid and strap muscle is opened and allows for early identification of the RLN. Parathyroid tissue is cryopreserved routinely. Use of adjuncts, such as measuring intraoperative PTH levels, is critical to ensure adequate resection and avoid potentially harmful additional explorations. In case of difficult reexplorations, additional techniques such as bilateral internal jugular vein sampling for PTH, thyroid lobectomy on the side of the missing gland, cervical thymectomy, and ligation of the ipsilateral inferior thyroid artery (after lobectomy, to cause infarction of the missing gland) may be needed. Blind mediastinal exploration is not recommended. In patients who are denied, refuse, or fail exploration, medical options such as cinacalcet may be considered.

Figure 38-35. Anatomic location of ectopic parathyroid glands. Numbers represent number of glands found in each location, with a total of 54. (Reproduced with permission from Shen W, Düren M, Morita E, et al: Reoperation for persistent or recurrent primary hyperparathyroidism, Arch Surg. 1996 Aug;131(8):861-867.)
Recurrence or persistent HPT

1) Confirm diagnosis
2) Rule out FHH
3) Review operative notes and pathology

Noninvasive localization studies

- Positive
- Negative

Parathyroidectomy
Selective venous catheterization for PTH

- Positive
- Negative

Is tumor localized?

- Yes
- No

Parathyroidectomy
Medical therapy

Figure 38-36. Management of recurrent and persistent hyperparathyroidism (HPT). FHH = familial hypocalciuric hypercalcemia; PTH = parathyroid hormone.

Hypercalcemic Crisis. Patients with PHPT may occasionally present acutely with nausea, vomiting, fatigue, muscle weakness, confusion, and a decreased level of consciousness—a complex referred to as hypercalcemic crisis. These symptoms result from severe hypercalcemia from uncontrolled PTH secretion, worsened by polyuria, dehydration, and reduced kidney function and may occur with other conditions causing hypercalcemia. Calcium levels are markedly elevated and may be as high as 16 to 20 mg/dL. Parathyroid tumors tend to be large or multiple and may be palpable. Patients with parathyroid cancer or familial HPT are more likely to present with hypercalcemic crisis.

Treatment consists of therapies to lower serum calcium levels followed by surgery to correct HPT. The mainstay of therapy involves rehydration with a 0.9% saline solution to keep urine output >100 cc/h. Once urine output is established, diuretics and furosemide (which increases renal calcium clearance) are begun. If these methods are unsuccessful, other drugs may be used to lower serum calcium levels as outlined in Table 38-13. Occasionally, in life-threatening cases, hemodialysis may be of benefit.

Secondary Hyperparathyroidism. Secondary HPT commonly occurs in patients with chronic renal failure but also may occur in those with hypocalcemia secondary to inadequate calcium or vitamin D intake or malabsorption. The pathophysiology of HPT in chronic renal failure is complex and appears to be related to hyperphosphatemia (and resultant hypocalcemia), deficiency of 1,25-dihydroxy vitamin D due to loss of renal tissue, low calcium intake, decreased calcium absorption, and abnormal parathyroid cell response to extracellular calcium or vitamin D in vitro and in vivo. Patients generally are hypocalcemic or normocalcemic. Aluminum hydroxide, which often was used as a phosphate binder, has been shown to contribute to the osteomalacia observed in this disease. These patients generally are treated medically with a low-phosphate diet, phosphate binders, adequate intake of calcium and 1, 25-dihydroxy vitamin D, and a high-calcium, low-aluminum dialysis bath. Calcimimetics have been shown to control parathyroid hyperplasia and osteitis fibrosa cystica associated with secondary HPT in animal studies and to decrease plasma PTH and total and ionized calcium levels in humans.

As the indications for parathyroidectomy were not well established, surgical treatment was traditionally recommended for patients with bone pain, pruritus, and (a) a calcium-phosphate product ≥70, (b) calcium >11 mg/dL with markedly elevated PTH, (c) calciphylaxis, (d) progressive renal osteodystrophy, and (e) soft tissue calcification and tumoral calcinosis, despite maximal medical therapy. Following the introduction of calcimimetics, there appears to have been a reduction in parathyroidectomy rates. Parathyroidectomy has been reported to maintain biochemical targets for up to 5 years and improve bone density, fracture risk, calcinosis, hemoglobin levels, and even long-term survival. Studies also report that in a large series of patients on hemodialysis, calcimimetics increased the likelihood of achieving goal PTH (≤300 pg/mL), calcium, phosphate, and Ca × PO₄ product, in addition to reducing the risk of fractures and cardiovascular complications.

In the absence of randomized trials comparing medical therapy with parathyroidectomy, current recommendations from the National Kidney Foundation’s Kidney Disease Quality Outcomes Initiative (KDOQI) advise parathyroidectomy for patients on maximal medical therapy with (a) severe HPT (defined as PTH >800 pg/mL), (b) hypercalcemia, (c) osteoporosis or pathologic bone fracture, (d) Symptoms and signs such as pruritus, bone pain, severe vascular calcifications, myopathy, and (e) calciphylaxis. Calciphylaxis is a rare, limb- and life-threatening complication of secondary HPT characterized by painful (sometimes throbbing), violaceous, and mottled lesions usually on the extremities, which often become necrotic and progress to non-healing ulcers, gangrene, sepsis, and death. Skin biopsy can be helpful to make the diagnosis. These are critically ill, high-risk patients, but successful parathyroidectomy sometimes relieves symptoms. Not all patients with calciphylaxis will have high PTH levels, and parathyroidectomy should not be undertaken in the absence of documented hyperparathyroidism. Assessment of parathyroid mass is thought to be an important factor for predicting the response to medical management. Therefore, some groups recommend parathyroidectomy if the glands are >1 cm (or >500 mm³) on ultrasound. These glands are more likely to have developed nodular hyperplasia and hence might be refractory to medical management.

Patients should undergo routine dialysis the day before surgery to correct electrolyte abnormalities. Localization studies are not necessary but can identify ectopic parathyroid glands. A bilateral neck exploration is indicated. The parathyroid glands in secondary HPT are characterized by asymmetric enlargement and nodular hyperplasia. These patients may be treated by subtotal resection, leaving about 50 mg of the most normal parathyroid gland or total parathyroidectomy and autotransplantation.
of parathyroid tissue into the brachioradialis muscle of the non-
dominant forearm, with parathyroid cryopreservation. Upper
thymectomy usually is performed because 15% to 20% of
patients have one or more parathyroid glands situated in the thy-
mus or perithymic fat. Some groups recommend total parathy-
roidectomy without autotransplantation because it is associated
with a lower rate of recurrence.95 While it may be preferable in
patients with calciphylaxis, this procedure is contraindicated in
some patients with chronic renal failure.

Tertiary Hyperparathyroidism. Generally, renal transplan-
tation is an excellent method of treating secondary HPT, but
some patients develop autonomous parathyroid gland function
and tertiary HPT. Tertiary HPT can cause problems similar to
PHPT, such as pathologic fractures, bone pain and worsened
bone disease, renal stones, peptic ulcer disease, pancreatitis,
and mental status changes. The transplanted kidney is also at
risk from tubulointerstitial calcification and volume depletion.
Similar to patients with secondary HPT, many patients with ter-
tiary HPT are being treated with cinacalcet. Although the drug
is effective and well-tolerated in these patients, the long-term
effects on kidney allograft function are not known, and many
of these patients have persistence of their hypercalcemic symp-
toms. On the other hand, parathyroidectomy has been shown to
lead to a more immediate and dramatic reduction in hypercal-
cemic symptoms. As such, operative intervention is indicated if
autonomous PTH secretion persists for >1 year after a successful
transplant in patients with hypophosphatemia, low BMD/severe
osteopenia, symptoms, and signs such as fatigue, pruritis, bone
pain, peptic ulcer disease or nephrocalcinosis, provided they
are deemed operative candidates.96 All parathyroid glands
should be identified. The traditional surgical management of
these patients consisted of subtotal or total parathyroid-
ectomy with autotransplantation and an upper thymectomy.
Some authors suggest that these patients derive similar benefit
from excision of only obviously enlarged glands, while avoid-
ing the higher risks of hypocalcemia associated with the for-
mer approach. Others recommend that all parathyroid glands
be identified and subtotal parathyroidectomy be performed as
long-term follow-up studies show that limited excisions in these
patients are associated with an up to fivefold increased risk of
recurrent or persistent disease. Further studies are needed to
define the best operative approach for these patients.

Complications of Parathyroid Surgery. Parathyroidectomy
can be accomplished successfully in >95% of patients with
minimal mortality and morbidity, provided the procedure is
performed by a surgeon experienced in parathyroid surgery.

<table>
<thead>
<tr>
<th>MEDICATION</th>
<th>DOSAGE AND ADMINISTRATION</th>
<th>MECHANISM, ONSET OF ACTION, AND DURATION</th>
<th>SIDE EFFECTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bisphosphonates (pamidronate,</td>
<td>60–90 mg IV over 4–24 h</td>
<td>Inhibits osteoclastic bone resorption; rapid onset, 2–3 d</td>
<td>May cause local pain and swelling, low-grade fever, lymphopenia, electrolyte abnormalities, osteonecrosis of the jaw in some patients (iv use)</td>
</tr>
<tr>
<td>Gallium nitrate</td>
<td>200 mg/m² BSA/d IV for 5 d</td>
<td>Reduces urinary calcium excretion; onset of action delayed (5–7 d)</td>
<td>Nephrotoxicity, nausea, vomiting, hypotension, anemia, hypophosphatemia</td>
</tr>
<tr>
<td>Mithramycin (plicamycin)</td>
<td>25 μg/kg/d IV for 3–4 d</td>
<td>Inhibits osteoclasts RNA secretion; rapid onset of action (12 h); peaks at 48–72 h and lasts days to several weeks</td>
<td>May cause renal, hepatic, and hematologic complications, nausea and vomiting</td>
</tr>
<tr>
<td>Glucocorticoids</td>
<td>100 mg IV q8h</td>
<td>Delayed onset of action (7–10 d); useful for hematologic malignancies, sarcoidosis, vitamin D intoxication, hyperthyroidism</td>
<td>Hypertension, hyperglycemia</td>
</tr>
<tr>
<td>Calcimimetics (cinacalcet)</td>
<td>Up to 90 mg 3 o 4 times per day</td>
<td>Useful in patients with parathyroid carcinoma and patients with chronic renal failure</td>
<td>Gastrointestinal complaints, hypotension, hypocalcemia</td>
</tr>
</tbody>
</table>

BSA = body surface area; IM = intramuscular; IV = intravenous; SC = subcutaneous.
Specific complications include transient and permanent vocal cord palsy and hypoparathyroidism. The latter is more likely to occur in patients who undergo four-gland exploration with biopsies, subtotal resection with an inadequate remnant, or total parathyroidectomy with a failure of autotransplanted tissue. Furthermore, hypocalcemia is more likely to occur in patients with high-turnover bone disease as evidenced by elevated preoperative alkaline phosphatase levels. Vocal cord paralysis and hypoparathyroidism are considered permanent if they persist for >6 months. Fortunately, these complications are rare, occurring in approximately 1% of patients undergoing surgery by experienced parathyroid surgeons.

Patients with symptomatic hypocalcemia or those with calcium levels <8 mg/dL are treated with oral calcium supplementation (up to 1-2 g every 4 hours). 1,25-Dihydroxy vitamin D (calcitriol [Rocaltril] 0.25–0.5 μg twice a day) may also be required, particularly in patients with severe hypercalcemia and elevated serum alkaline phosphatase levels, preoperatively and with osteitis fibrosa cystica. Intravenous calcium supplementation rarely is needed, except in cases of severe, symptomatic hypocalcemia.

**Hypoparathyroidism**

Hypocalcemia can be the result of a multitude of conditions, which are listed in Table 38-14. The parathyroid glands may be congenitally absent in DiGeorge syndrome, which also is characterized by lack of thymic development and, therefore, a thymus-dependent lymphoid system. By far, the most common cause of hypoparathyroidism is thyroid surgery, particularly total thyroidectomy with a concomitant central neck dissection. Patients often develop transient hypocalcemia due to ischemia of the parathyroid glands; permanent hypoparathyroidism is rare. Hypoparathyroidism also may occur after parathyroid surgery, which is more likely if patients undergo a subtotal resection or total parathyroidectomy with parathyroid autotransplantation.

Acute hypocalcemia results in decreased ionized calcium and increased neuromuscular excitability. Patients initially develop circumoral and fingertip numbness and tingling. Mental symptoms include anxiety, confusion, and depression. Physical examination reveals positive Chvostek’s sign (contraction of facial muscles elicited by tapping on the facial nerve anterior to the ear) and Trousseau’s sign (carpopedal spasm that is elicited by occluding blood flow to the forearm with a blood pressure cuff for 2–3 minutes). Tetany, which is characterized by tonic-clonic seizures, carpopedal spasm, and laryngeal stridor, may prove fatal and should be avoided. Most patients with postoperative hypocalcemia can be treated with oral calcium and vitamin D supplements; IV calcium infusion is rarely required except in patients with preoperative osteitis fibrosa cystica.

**ADRENAL**

**Historical Background**

Eustachius provided the first accurate anatomic account of the adrenals in 1563. The anatomic division of the adrenals into the cortex and medulla was described much later, by Cuvier in 1805. Subsequently, Thomas Addison in 1855 described the features of adrenal insufficiency, which still bear his name. DeCrecchio provided the first description of congenital adrenal hyperplasia (CAH) occurring in a female pseudohermaphrodite in 1865. Pheochromocytomas were first identified by Frankel in 1885, but were not named as such until 1912 by Pick, who noted the characteristic chromaffin reaction of the tumor cells. Adrenaline was identified as an agent from the adrenal medulla that elevated blood pressure in dogs and was subsequently named epinephrine in 1897. The first successful adrenalectomies for pheochromocytoma were performed by Roux in Switzerland and Charles Mayo in the United States.

In 1932, Harvey Cushing described 11 patients who had moon facies, truncal obesity, hypertension, and other features of the syndrome that now bears his name. Although several individuals prepared adrenocortical extracts to treat adrenalec-tomized animals, cortisol was first synthesized by Kendall. Aldosterone was identified in 1952, and the syndrome resulting from excessive secretion of this mineralocorticoid was first described in 1955 by Conn.

**Embryology**

The adrenal or suprarenal glands are two endocrine organs in one; an outer cortex and an inner medulla, each with distinct embryologic, anatomic, histologic, and secretory features. The cortex originates around the fifth week of gestation from the mesodermal tissue near the gonads on the adrenogenital ridge (Fig. 38-37). Therefore, ectopic adrenocortical tissue may be found in the ovaries, spermatic cord, and testes. The cortex differentiates further into a thin, definitive cortex and a thicker, inner fetal cortex. The latter is functional and produces fetal adrenal steroids by the eighth week of gestation, but undergoes involution after birth, resulting in a decrease in adrenal weight during the first three postpartum months. The definitive cortex persists after birth to form the adult cortex over the first 3 years of life. In contrast, the adrenal medulla is ectodermal in origin and arises from the neural crest. At around the same time as cortical development, neural crest cells migrate to the para-aortic and paravertebral areas and toward the medial aspect of the developing cortex to form the medulla. Most extra-adrenal neural tissue regresses but may persist at several sites. The largest of these is located to the left of the aortic bifurcation near the inferior mesenteric artery origin and is designated as the organ of Zuckerkandl. Adrenal medullary tissue also may be found in neck, urinary bladder, and para-aortic regions. Several

<table>
<thead>
<tr>
<th>Table 38-14</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conditions causing hypocalcemia</strong></td>
</tr>
<tr>
<td><strong>Hypoparathyroidism</strong></td>
</tr>
<tr>
<td>• Surgical</td>
</tr>
<tr>
<td>• Neonatal</td>
</tr>
<tr>
<td>• Familial</td>
</tr>
<tr>
<td>• Heavy metal deposition</td>
</tr>
<tr>
<td>• Magnesium depletion</td>
</tr>
<tr>
<td><strong>Resistance to the action of parathyroid hormone</strong></td>
</tr>
<tr>
<td>• Pseudohypoparathyroidism</td>
</tr>
<tr>
<td>• Renal failure</td>
</tr>
<tr>
<td>• Medications—calcitonin, bisphosphonates, mithramycin</td>
</tr>
<tr>
<td><strong>Failure of normal 1,25-dihydroxy vitamin D production</strong></td>
</tr>
<tr>
<td><strong>Resistance to the action of 1,25-dihydroxy vitamin D</strong></td>
</tr>
<tr>
<td>• Acute hyperphosphatemia</td>
</tr>
<tr>
<td>• Acute pancreatitis</td>
</tr>
<tr>
<td>• Massive blood transfusion (citrate overload)</td>
</tr>
<tr>
<td>• “Hungry bones”</td>
</tr>
</tbody>
</table>
Anatomy

The adrenal glands are paired, retroperitoneal organs located superior and medial to the kidneys at the level of the eleventh ribs. The normal adrenal gland measures 5 × 3 × 1 cm and weighs 4 to 5 g. The right gland is pyramidal shaped and lies in close proximity to the right hemidiaphragm, liver, and inferior vena cava (IVC). The left adrenal is closely associated with the aorta, spleen, and tail of the pancreas. Each gland is supplied by three groups of vessels—the superior adrenal arteries derived from the inferior phrenic artery, the middle adrenal arteries derived from the aorta, and the inferior adrenal arteries derived from the renal artery. Other vessels originating from the intercostal and gonadal vessels may also supply the adrenals. These arteries branch into about 50 arterioles to form a rich plexus beneath the glandular capsule and require careful dissection, ligation, and division during adrenalectomy. In contrast to the arterial supply, each adrenal usually is drained by a single, major adrenal vein. The right adrenal vein is usually short and drains into the IVC, whereas the left adrenal vein is longer and empties into the left renal vein after joining the inferior phrenic vein. Accessory veins occur in 5% to 10% of patients—on the right, these vessels may drain into the right hepatic vein or the right renal vein; on the left, accessory veins may drain directly into the left renal vein. The anatomic relationships of the adrenals and surrounding structures are depicted in Fig. 38-38.

The adrenal cortex appears yellow due to its high lipid content and accounts for about 80% to 90% of the gland’s volume. Histologically, the cortex is divided into three zones—the zona glomerulosa, zona fasciculata, and zona reticularis. The outer area of the zona glomerulosa consists of small cells and is the site of production of the mineralocorticoid hormone, aldosterone. The zona fasciculata is made up of larger cells, which often appear foamy due to multiple lipid inclusions, whereas the zona reticularis cells are smaller. These latter zones are the site of production of glucocorticoids and adrenal androgens. The adrenal medulla constitutes up to 10% to 20% of the gland’s volume and is reddish-brown in color. It produces the catecholamine hormones epinephrine and norepinephrine. The cells of the adrenal medulla are arranged in cords and are polyhedral in shape. They are often referred to as chromaffin cells because they stain specifically with chromium salts.

Adrenal Physiology

Cholesterol, derived from the plasma or synthesized in the adrenal, is the common precursor of all steroid hormones derived from the adrenal cortex. Cholesterol initially is cleaved within mitochondria to 5-6-pregnanolone, which in turn is transported to the smooth endoplasmic reticulum where it forms the substrate for various biosynthetic pathways leading to steroidogenesis (Fig. 38-39).

Mineralocorticoids. The major adrenal mineralocorticoid hormones are aldosterone, 11-deoxycorticosterone (DOC), and cortisol. Cortisol has minimal effects on the kidney due to hormone degradation. Aldosterone secretion is regulated primarily by the renin-angiotensin system. Decreased renal blood flow, decreased plasma sodium, and increased sympathetic tone all stimulate the release of renin from juxtaglomerular cells. Renin, in turn, leads to the production of angiotensin I from its precursor angiotensinogen. Angiotensin I is cleaved by pulmonary angiotensin-converting enzyme (ACE) to angiotensin II; the latter is not only a potent vasoconstrictor, but it also leads to increased aldosterone synthesis and release. Hyperkalemia is another potent stimulator of aldosterone synthesis, whereas ACTH, pituitary pro-opiomelanocortin, and antidiuretic hormone are weak stimulators.

Aldosterone is secreted at a rate of 50 to 250 μg/d (depending on sodium intake) and circulates in plasma chiefly as a complex with albumin. Small amounts of the hormone bind to corticosteroid-binding globulin, and approximately 30% to 50% of secreted aldosterone circulates in a free form. The hormone has a half-life of only 15 to 20 minutes and is rapidly cleared via the liver and kidney. A small quantity of free aldosterone also is excreted in the urine. Mineralocorticoids cross the cell membrane and bind to cytosolic receptors. The receptor-ligand complex subsequently is transported into the nucleus where it induces the transcription and translation of specific genes. Aldosterone functions mainly to increase sodium reabsorption and potassium and hydrogen ion excretion at the level of the renal distal convoluted tubule. Less commonly, aldosterone increases sodium absorption in salivary glands and GI mucosal surfaces.
**Glucocorticoids.** The secretion of cortisol, the major adrenal glucocorticoid, is regulated by ACTH secreted by the anterior pituitary, which, in turn, is under the control of corticotrophin-releasing hormone (CRH) secreted by the hypothalamus. ACTH is a 39-amino-acid protein, which is derived by cleavage from a larger precursor, pro-opiomelanocortin. ACTH is further cleaved into \( \alpha \)-melanocyte-stimulating hormone and corticotrophin-like intermediate peptide. ACTH not only stimulates the secretion of glucocorticoids, mineralocorticoids, and adrenal androgens, but is also trophic for the adrenal glands. ACTH secretion may be
stimulated by pain, stress, hypoxia, hypothermia, trauma, and hypoglycemia. ACTH secretion fluctuates, peaking in the morning and reaching nadir levels in the late afternoon. Thus, there is a diurnal variation in the secretion of cortisol, with peak cortisol excretion also occurring in the early morning and declining during the day to its lowest levels in the evening (Fig. 38-40). Cortisol controls the secretion of both CRH and ACTH via a negative feedback loop. A similar mechanism leads to the inhibition of CRH secretion by ACTH.

Cortisol is transported in plasma bound primarily to corticosteroid-binding globulin (75%) and albumin (15%). Approximately 10% of circulating cortisol is free and is the biologically active component. The plasma half-life of cortisol is 60 to 90 minutes and is determined by the extent of binding and rate of inactivation. Cortisol is converted to di- and tetrahydrocortisol and cortisone metabolites in the liver and the kidney. The majority (95%) of cortisol and cortisone metabolites are conjugated with glucuronic acid in the liver, thus facilitating their renal excretion. A small amount of unmetabolized cortisol is excreted unchanged in the urine.

Glucocorticoid hormones enter the cell and bind cytosolic steroid receptors. The activated receptor-ligand complex is then transported to the nucleus where it stimulates the transcription of specific target genes via a “zinc finger” DNA binding element. Cortisol also binds the mineralocorticoid receptor with an affinity similar to aldosterone. However, the specificity of mineralocorticoid action is maintained by the production of 11β-hydroxysteroid dehydrogenase, an enzyme that inactivates cortisol to cortisone in the kidney. Glucocorticoids have important functions in intermediary metabolism but also affect connective tissue, bone, immune, cardiovascular, renal, and central nervous systems, as outlined in Table 38-15.

**Sex Steroids.** Adrenal androgens are produced in the zona fasciculata and reticularis from 17-hydroxypregnenolone in response to ACTH stimulation. They include dehydroepiandrosterone (DHEA) and its sulfated counterpart (DHEAS), androstenedione, and small amounts of testosterone and estrogen. Adrenal androgens are weakly bound to plasma albumin. They exert their major effects by peripheral conversion to the more potent testosterone and dihydrotestosterone but also have weak intrinsic androgen activity. Androgen metabolites are conjugated as glucuronides or sulfates and excreted in the urine. During fetal development, adrenal androgens promote the formation of male genitalia. In normal adult males, the contribution of adrenal androgens is minimal; however, they are responsible for the development of secondary sexual characteristics at puberty. Adrenal androgen excess leads to precocious puberty in boys and virilization, acne, and hirsutism in girls and women.

**Catecholamines.** Catecholamine hormones (epinephrine, norepinephrine, and dopamine) are produced not only in the central and sympathetic nervous system but also the adrenal medulla. The substrate, tyrosine, is converted to catecholamines via a series of steps shown in Fig. 38-41A. Phenylethanolamine N-methyltransferase, which converts norepinephrine to epinephrine, is only present in the adrenal medulla and the organ of Zuckerkandl. Therefore, the primary catecholamine produced may be used to distinguish adrenal medullary tumors from those situated at extra-adrenal sites. Catecholamines are stored in granules in combination with other neuropeptides, ATP, calcium, magnesium, and water-soluble proteins called chromogranins. Hormonal secretion is stimulated by various stress stimuli and mediated by the release of acetylcholine at the preganglionic nerve terminals. In the circulation, these proteins are bound to albumin and other proteins. Catecholamines are cleared by several mechanisms including reuptake by sympathetic nerve endings, peripheral inactivation by catechol O-methyltransferase and monoamine oxidase, and direct excretion by the kidneys. Metabolism of catecholamines takes place...
Table 38-15

Functions of glucocorticoid hormones

<table>
<thead>
<tr>
<th>FUNCTION/SYSTEM</th>
<th>EFFECTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose metabolism</td>
<td>Increased hepatic glycogen deposition, gluconeogenesis, decreased muscle glucose uptake and metabolism</td>
</tr>
<tr>
<td>Protein metabolism</td>
<td>Decreased muscle protein synthesis, increased catabolism</td>
</tr>
<tr>
<td>Fat metabolism</td>
<td>Increased lipolysis in adipose tissue</td>
</tr>
<tr>
<td>Connective tissue</td>
<td>Inhibition of fibroblasts, loss of collagen, thinning of skin, striae formation</td>
</tr>
<tr>
<td>Skeletal system</td>
<td>Inhibition of bone formation, increased osteoclast activity, potentiate the action of PTH</td>
</tr>
<tr>
<td>Immune system</td>
<td>Increases circulation of polymorphonuclear cells; decreases numbers of lymphocytes, monocytes, and eosinophils; reduces migration of inflammatory cells to sites of injury</td>
</tr>
<tr>
<td>Cardiovascular system</td>
<td>Increases cardiac output and peripheral vascular tone</td>
</tr>
<tr>
<td>Renal system</td>
<td>Sodium retention, hypokalemia, hypertension via mineralocorticoid effect, increased glomerular filtration via glucocorticoid effects</td>
</tr>
<tr>
<td>Endocrine system</td>
<td>Inhibits TSH synthesis and release, decreased TBG levels, decreased conversion of T₄ to T₃</td>
</tr>
</tbody>
</table>

PTH = parathyroid hormone; T₂ = 3,5'-3-triiodothyronine; T₄ = thyroxine; TBG = thyroxine-binding globulin; TSH = thyroid-stimulating hormone.

primarily in the liver and kidneys and leads to the formation of metabolites such as metanephrines, normetanephrines, and VMA, which may undergo further glucuronidation or sulfation before being excreted in the urine (Fig. 38-41B).

Adrenergic receptors are transmembrane-spanning molecules that are coupled to G proteins. They may be subdivided into α and β subtypes, which are localized in different tissues, have varying affinity to various catecholamines, and mediate distinct biologic effects (Table 38-16). The receptor affinities for α receptors are—epinephrine > norepinephrine >> isoproterenol; β₁ receptors—iso- proterenol > epinephrine = norepinephrine; and β₂ receptors—iso- proterenol > epinephrine >> norepinephrine.

Disorders of the Adrenal Cortex

Hyperaldosteronism. Hyperaldosteronism may be secondary to stimulation of the renin-angiotensin system from renal artery stenosis and to low-flow states such as congestive heart failure and cirrhosis. Hyperaldosteronism resulting from these conditions is reversible by treatment of the underlying cause. Primary hyperaldosteronism results from autonomous aldosterone secretion, which, in turn, leads to suppression of renin secretion. Primary aldosteronism usually occurs in individuals between the ages of 30 to 50 years old and accounts for 1% of hypertension cases. It is associated with hypokalemia; however, more patients with Conn’s syndrome are being diagnosed with normal potassium levels. Most cases result from a solitary functioning adrenal adenoma (~70%) and idiopathic bilateral hyperplasia (30%). Adrenocortical carcinoma and glucocorticoid-suppressible hyperaldosteronism are rare, each accounting for <1% of cases. Glucocorticoid-suppressible hyperaldosteronism is an autosomal dominant form of hypertension in which aldosterone secretion is abnormally regulated by ACTH. This condition is caused by recombinations between linked genes encoding closely related isozymes, 11b-hydroxylase (CYP11B1), and aldosterone synthase (CYP11B2) generating a dysregulated chimeric gene with aldosterone synthase activity. This entity is now designated familial hyperaldosteronism type I (FH-I).

Initially, FH-III referred to patients with massive adrenal hyperplasia refractory to glucocorticoid administration. However, the term is now more commonly used to describe patients with primary hyperaldosteronism due to germline KCNJ5 mutations. This gene encodes an inward rectifier potassium channel, and the mutations affected amino acids in or close to the channel’s selectivity filter. Somatic gain of function mutations are found in up to 40% of aldosterone-producing adenomas (APA). Familial hyperaldosteronism type II refers to families in which two first-degree relatives have been diagnosed with primary hyperaldosteronism (adrena or hyperplasia) and in whom types I and III have been excluded. Other genes mutated in (APA) include CACNA1D (encodes the plasma membrane calcium ATPase), ATP2B3 (encodes the plasma membrane Ca²⁺ ATPase), CACNA1H (encodes the α₁ subunit of the T-type voltage calcium channel), and CTNNB1 (β-catenin).

Symptoms and Signs Patients typically present with hypertension, which is long-standing, moderate to severe, and may be difficult to control despite multiple-drug therapy. Other symptoms include muscle weakness, polydipsia, polyuria, nocturia, headaches, and fatigue. Weakness and fatigue are related to the presence of hypokalemia.

Diagnostic Studies

Laboratory Studies. Hypokalemia is a common finding, and hyperaldosteronism must be suspected in any hypertensive patient who presents with coexisting spontaneous hypokalemia (K <3.2 mmol/L) or hypokalemia (<3 mmol/L) while on diuretic therapy, despite potassium replacements. However, it is important to note that up to 40% of patients with a confirmed aldosteronoma were normokalemic preoperatively. Once the diagnosis is suspected, further tests are necessary to confirm the diagnosis. Before testing, patients must receive adequate sodium and potassium. Antihypertensive medications should be held, if possible, and spironolactone, β-blockers, ACE inhibitors, and angiotensin II receptor blockers should be avoided. Patients with primary hyperaldosteronism have an elevated plasma aldosterone concentration level with a suppressed plasma renin activity; a plasma aldosterone concentration-to-plasma renin activity ratio of 1:25 to 30 is strongly suggestive of the diagnosis. False-positive results can occur, particularly in patients with chronic renal failure. Patients with primary hyperaldosteronism also fail to suppress aldosterone levels with sodium loading. This test can be performed by performing a 24-hour urine collection for cortisol, sodium, and aldosterone after 5 days of a high-sodium diet or alternatively giving the...
**Figure 38-41.** A. Synthesis of catecholamines. B. Metabolism of catecholamine hormones.
In this procedure, the adrenal veins are cannulated, and blood samples for aldosterone and cortisol are obtained from both adrenal veins and the vena cava after ACTH administration. Measurement of cortisol levels is necessary to confirm proper placement of the catheters in the adrenal veins. A greater than fourfold difference in the aldosterone-to-cortisol ratios between the adrenal veins indicates the presence of a unilateral tumor. Some investigators use this study routinely, but it is invasive, requires an experienced interventional radiologist, and can lead to adrenal vein rupture in approximately 1% of cases. Therefore, most groups advocate use of this modality selectively in ambiguous cases, when the tumor cannot be localized and in patients with bilateral adrenal enlargement to determine whether there is unilateral or bilateral increased secretion of aldosterone. Additional indications for forgoing AVS include patients who are suspected of having adrenocortical carcinoma, those with comorbid conditions precluding surgery and those with proven familial hyperaldosteronism type I or III. Scintigraphy with \(^{131}\text{I}-6\beta\)-iodomethyl noriodocholesterol (NP-59) also may be used for the same purpose. Like cholesterol, this compound is taken up by the adrenal cortex, but unlike cholesterol, it remains in the gland without undergoing further metabolism. Adrenal adenomas appear as “hot” nodules with suppressed contralateral uptake, whereas hyperplastic glands show bilaterally increased uptake. This test, however, is not widely available. Newer isotopes such as \(^{11}\text{C}\)-metomidate in conjunction with PET-CT have also shown promise in the localization of aldosteronomas.
**Treatment** Preoperatively, control of hypertension and adequate potassium supplementation (to keep K >3.5 mmol/L) are important. Patients generally are treated with spironolactone (an aldosterone antagonist), amiloride (a potassium-sparing diuretic that blocks sodium channels in the distal nephron), nifedipine (a calcium channel blocker), or captopril (an ACE inhibitor). Unilateral tumors producing aldosterone are best managed by adrenalectomy, either by a laparoscopic approach (preferred) or via a posterior open approach. If a carcinoma is suspected because of the large size of the adrenal lesion or mixed hormone secretion, an anterior transabdominal approach is preferred to permit adequate determination of local invasion and distal metastases. Only 20% to 30% of patients with hyperaldosteronism secondary to bilateral adrenal hyperplasia benefit from surgery, and as described, selective venous catheterization is useful to predict which patients will respond. For the other patients, medical therapy with spironolactone, amiloride, or triamterene is the mainstay of management. Glucocorticoid-suppressible hyperaldosteronism is treated by administering exogenous dexamethasone at doses of 0.5 to 1 mg daily. Treatment with spironolactone may help decrease glucocorticoid requirements in this condition and avoid symptoms of Cushing’s syndrome. Postoperatively, some patients experience transient hypoaldosteronism requiring mineralocorticoids for up to 3 months. Recent studies suggest that postresection hyperkalemia may be more common and last longer than previously appreciated; therefore, it should be screened for in patients who are older and who have had a longer duration of hypertension, impaired kidney function, and higher preoperative aldosterone levels, making them a high-risk group. Rarely, acute Addison’s disease may occur 2 to 3 days after adrenalectomy. Adrenalectomy is >90% successful in improving hypokalemia and about 70% successful in correcting hypertension. Patients who respond to spironolactone therapy and those with a shorter duration of hypertension with minimal renal damage are more likely to achieve improvement in hypertension, whereas male patients, those >50 years old, and those with multiple adrenal nodules, are least likely to benefit from adrenalectomy.

**Cushing’s Syndrome.** Cushing described patients with a peculiar fat deposition, amenorrhea, impotence (men), hirsutism, purple striae, hypertension, diabetes, and other features that constitute the syndrome (Fig. 38-43). He also recognized that several of these patients had basophilic tumors of the pituitary gland and concluded that these tumors produced hormones that caused adrenocortical hyperplasia, thus resulting in the manifestations of the syndrome. Today, the term Cushing’s syndrome refers to a complex of symptoms and signs resulting from hypersecretion of cortisol regardless of etiology. In contrast, Cushing’s disease refers to a pituitary tumor, usually an adenoma, which leads to bilateral adrenal hyperplasia and hypercortisolism. Cushing’s syndrome (endogenous) is a rare disease, affecting 10 in 1 million individuals. It is more common in adults but may occur in children. Women are more commonly affected (male-to-female ratio is 1:8). Although most individuals have sporadic disease, Cushing’s syndrome may be found in MEN1 families and can result from ACTH-secreting pituitary tumors, primary adrenal neoplasms, or an ectopic ACTH-secreting carcinoid tumor (more common in men) or bronchial adenoma (more common in women).

Cushing’s syndrome may be classified as ACTH-dependent or ACTH-independent (Table 38-17). The most common cause of hypercortisolism is exogenous administration of steroids. However, approximately 70% of cases of endogenous Cushing’s syndrome are caused by an ACTH-producing pituitary tumor. Primary adrenal sources (adenoma, hyperplasia, and

---

**Table 38-17**

**Etiology of Cushing’s syndrome**

<table>
<thead>
<tr>
<th>Classification</th>
<th>Etiology</th>
</tr>
</thead>
</table>
| ACTH-dependent (70%) | • Pituitary adenoma or Cushing’s disease (~70%)  
• Ectopic ACTH production (~10%)  
• Ectopic CRH production (<1%) |
| ACTH-independent (20–30%) | • Adrenal adenoma (10–15%)  
• Adrenal carcinoma (5–10%)  
• Adrenal hyperplasia—pigmented micronodular cortical hyperplasia or gastric inhibitory peptide-sensitive macronodular hyperplasia (5%) |
| Other | • Pseudo-Cushing’s syndrome  
• Iatrogenic—exogenous administration of steroids |

*From small cell lung tumors, pancreatic islet cell tumors, medullary thyroid cancers, pheochromocytomas, and carcinoid tumors of the lung, thymus, gut, pancreas, and ovary.

ACTH = adrenocorticotropic hormone; CRH = corticotrophin-releasing hormone.
cancer) account for about 20% of cases, and ectopic ACTH-secreting tumors account for <10% of cases. CRH also may be secreted ectopically in bronchial carcinoid tumors, pheochromocytomas, and other tumors. These patients are difficult to distinguish from those with ectopic ACTH production, but can be diagnosed by determining CRH levels. Patients with major hypertension, psychosis, pregnancy, chronic renal failure, or stress also may have elevated cortisol levels and symptoms of hypercortisolism. However, these manifestations resolve with treatment of the underlying disorder, and these patients are deemed to have pseudo-Cushing’s syndrome.

Primary adrenal hyperplasia may be micronodular, macronodular, or massively macronodular. Adrenal hyperplasia resulting from ACTH stimulation usually is macronodular (3-cm nodules). Primary pigmented nodular adrenocortical disease is a rare cause of ACTH-independent Cushing’s syndrome, which is characterized by the presence of small (<5 mm), black adrenal nodules. Primary pigmented nodular adrenocortical disease may be associated with Carney complex (atrial myxomas, schwannomas, and pigmented nevi) and is thought to be immune related.

**Symptoms and Signs** The classical features of Cushing’s syndrome are listed in Table 38-18. Early diagnosis of this disease requires a thorough knowledge of these manifestations, coupled with a high clinical suspicion. In some patients, symptoms are less pronounced and may be more difficult to recognize, particularly given their diversity and the absence of a single defining symptom or sign. Progressive truncal obesity is the most common symptom, occurring in up to 95% of patients. This pattern results from the lipogenic action of excessive corticosteroids centrally and catabolic effects peripherally, along with peripheral muscle wasting. Fat deposition also occurs in unusual sites, such as the supraclavicular space and posterior neck region, leading to the so-called buffalo hump. Purple striae are often visible on the protuberant abdomen. Rounding of the face leads to moon facies, and thinning of subcutaneous tissues leads to plethora. There is an increase in fine hair growth on the face, upper back, and arms, although true virilization is more commonly seen with adrenocortical cancers. Endocrine abnormalities include glucose intolerance, amenorrhea, and decreased libido or impotence. In children, Cushing’s syndrome is characterized by obesity and stunted growth. Patients with Cushing’s disease also may present with headaches, visual field defects, and panhypopituitarism. Hyperpigmentation of the skin, if present, suggests an ectopic ACTH-producing tumor with high levels of circulating ACTH.

**Diagnostic Tests** The aims of diagnostic tests in the evaluation of patients suspected of having Cushing’s syndrome are twofold: to confirm the presence of Cushing’s syndrome and to determine its etiology (Fig. 38-44).

**Laboratory Studies.** Cushing’s syndrome is characterized by elevated glucocorticoid levels that are not suppressible by exogenous hormone administration and loss of diurnal variation. This phenomenon is used to screen patients using the overnight low-dose dexamethasone suppression test. In this test, 1 mg of a synthetic glucocorticoid (dexamethasone) is given at 11 p.m. and plasma cortisol levels are measured at 8 a.m. the following morning. Physiologically normal adults suppress cortisol levels to <3 μg/dL, whereas most patients with Cushing’s syndrome do not. False-negative results may be obtained in patients with mild disease; therefore, some authors consider the test positive only if cortisol levels are suppressed to <1.8 μg/dL. False-positive results can occur in up to 3% of patients with chronic renal failure, depression, or those taking medications such as phenytoin, which enhance dexamethasone metabolism. In patients with a negative test but a high clinical suspicion, the classic low-dose dexamethasone (0.5 mg every 6 hours for eight doses, or 2 mg over 48 hours) suppression test or urinary cortisol measurement should be performed. Measurement of elevated 24-hour urinary cortisol levels is a very sensitive (95–100%) and specific (98%) modality of diagnosing Cushing’s syndrome and is particularly useful for identifying patients with pseudo-Cushing’s syndrome. A urinary cortisol-free excretion of less than 100 μg/dL (in most laboratories) rules out hypercortisolism. Recently, salivary cortisol measurements using commercially available kits also have demonstrated superior sensitivity in diagnosing Cushing’s syndrome and are being increasingly used. Overall, 24-hour urinary tests for free cortisol and the overnight dexamethasone suppression test at the 5 μg/dL cutoff have the highest specificity for the diagnosis of Cushing’s syndrome. Once a diagnosis of hypercortisolism is established, further testing is aimed at determining whether it is ACTH-dependent or ACTH-independent Cushing’s syndrome. This is best accomplished by measurement of plasma ACTH levels (normal 10–100 pg/mL). Elevated ACTH levels are found in patients with adrenal hyperplasia due to Cushing’s disease (15–500 pg/mL) and those with CRH-secreting tumors, but the highest levels are found in patients with ectopic sources of ACTH (>1000 pg/mL). In contrast, ACTH levels are characteristically suppressed (<5 pg/mL) in patients with primary cortisol-secreting adrenal tumors. The high-dose dexamethasone suppression test is used to distinguish between the causes of ACTH-dependent Cushing’s syndrome (pituitary vs. ectopic). The standard test (2 mg dexamethasone every 6 hours for 2 days) or the overnight test (8 mg) may be used, with 24-hour urine collections for cortisol and 17-hydroxy steroids performed over the second day. Failure to suppress urinary cortisol by 50% confirms the diagnosis of an ectopic ACTH-producing tumor. Patients suspected of having ectopic tumors should also undergo testing for MTC and pheochromocytoma. Bilateral petrosal vein sampling also is helpful for determining whether the patient has Cushing’s disease or ectopic Cushing’s syndrome.

### Table 38-18

<table>
<thead>
<tr>
<th>SYSTEM</th>
<th>MANIFESTATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>General</td>
<td>Weight gain—central obesity, buffalo hump, supraclavicular fat pads</td>
</tr>
<tr>
<td>Integumentary</td>
<td>Hirsutism, plethora, purple striae, acne, ecchymosis</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>Hypertension</td>
</tr>
<tr>
<td>Musculoskeletal</td>
<td>Generalized weakness, osteopenia</td>
</tr>
<tr>
<td>Neuropsychiatric</td>
<td>Emotional lability, psychosis, depression</td>
</tr>
<tr>
<td>Metabolic</td>
<td>Diabetes or glucose intolerance, hyperlipidemia</td>
</tr>
<tr>
<td>Renal</td>
<td>Polyuria, renal stones</td>
</tr>
<tr>
<td>Gonadal</td>
<td>Impotence, decreased libido, menstrual irregularities</td>
</tr>
</tbody>
</table>
The CRH test also is helpful in determining the etiology of Cushing’s syndrome. Ovine CRH (1 μg/kg) is administered intravenously, followed by serial measurements of ACTH and cortisol at 15-minute intervals for 1 hour. Patients with a primary adrenal hypercortisolism exhibit a blunted response (ACTH peak <10 pg/mL), whereas those with ACTH-dependent Cushing’s syndrome demonstrate a higher elevation of ACTH (>30 pg/mL). CRH stimulation also can enhance the usefulness of petrosal vein sampling. Patients with pituitary tumors also have a higher peak ACTH than those with ectopic ACTH-producing tumors.

**Radiologic Studies.** CT and MRI scans of the abdomen can identify adrenal tumors with 95% sensitivity. They also are helpful in distinguishing adrenal adenomas from carcinomas, as discussed in the subsequent section “Adrenocortical Cancer.” MRI scans have the added advantage of allowing assessment of vascular anatomy. Adrenal adenomas appear darker than the liver on T2-weighted imaging. Radioscintigraphic imaging of the adrenals using NP-59 also can be used to distinguish adenoma from hyperplasia. Reports suggest that “cold” adrenal nodules are more likely to be cancerous, although this distinction is not absolute. NP-59 scanning is most useful in identifying patients with an adrenal source of hypercortisolism and primary pigmented micronodular hyperplasia.

Thin-section head CT scans are 22% sensitive, and contrast-enhanced brain MRI scans are 33% to 67% sensitive at identifying pituitary tumors. Inferior petrosal sinus sampling for ACTH before and after CRH injection has been helpful in this regard and has a sensitivity approaching 100%. In this study, catheters are placed in both internal jugular veins and a peripheral vein. A ratio of petrosal to peripheral vein ACTH level of >2 in the basal state and >3 after CRH stimulation is diagnostic.

---

**Figure 38-44.** Diagnosis of Cushing’s syndrome. ACTH = adrenocorticotropic hormone; CT = computed tomography; DST = dexamethasone suppression test.
of a pituitary tumor. In patients suspected of having ectopic ACTH production, CT or MRI scans of the chest and anterior mediastinum are performed first, followed by imaging of the neck, abdomen, and pelvis if the initial studies are negative.

**Treatment** Laparoscopic adrenalectomy is the treatment of choice for patients with adrenal adenomas. Open adrenalectomy is reserved for large tumors (≥6 cm) or those suspected to be adrenocortical cancers. Bilateral adrenalectomy is curative for primary adrenal hyperplasia.

The treatment of choice in Cushing’s disease is transsphenoidal excision of the pituitary adenoma, which is successful in 80% of patients. Pituitary irradiation has been used for patients with persistent or recurrent disease after surgery. However, it is associated with a high rate of panhypopituitarism, and some patients develop visual deficits. This has led to increased use of stereotactic radiosurgery, which uses CT guidance to deliver high doses of radiotherapy to the tumor (photon or gamma knife) and also bilateral laparoscopic adrenalectomy. Patients who fail to respond to either treatment are candidates for pharmacologic therapy with adrenal inhibitors (medical adrenalectomy) such as ketoconazole, metyrapone, or aminoglutethimide.

Patients with ectopic ACTH production are best managed by treating the primary tumor, including recurrences, if possible. Medical or bilateral laparoscopic adrenalectomy has been used to palliate patients with unresectable disease and those whose ectopic ACTH-secreting tumor cannot be localized.

Patients undergoing surgery for a primary adrenal adenoma secreting glucocorticoids require preoperative and postoperative steroids due to suppression of the contralateral adrenal gland. These patients are also at increased predisposition for infectious and thromboembolic complications, the latter due to a hypercoagulable state resulting from an increase in clotting factors including factor VIII and von Willebrand’s factor complex, and by impaired fibrinolysis. Duration of steroid therapy is determined by the ACTH stimulation test. Exogenous steroids may be needed for up to 2 years but are needed indefinitely in patients who have undergone bilateral adrenalectomy. This latter group of patients also may require mineralocorticoid replacement therapy. Typical replacement doses include hydrocortisone (10–20 mg every morning and 5–10 mg every evening) and fludrocortisone (0.05–0.1 mg/d every morning).

**Adrenocortical Cancer.** Adrenal carcinomas are rare neoplasms with a worldwide incidence of two per 1 million. These tumors have a bimodal age distribution, with an increased incidence in children and adults in the fourth and fifth decades of life. The majority are sporadic, but adrenocortical carcinomas also occur in association with germline mutations of p53 (Li-Fraumeni syndrome) and MEN1 (multiple endocrine neoplasia type 1) genes. Loci on 11p (Beckwith-Wiedemann syndrome), 2p (Carney complex), and 9q also have been implicated. Somatic p53 mutations are present in up to 33% of tumors, and LOH at the p53 locus has been reported in >85% of adrenocortical carcinomas. In addition, insulin-like growth factor II is overexpressed in 90% of tumors, and approximately 30% harbor somatic activating mutations in the β-catenin gene. Recently identified genes mutated in adrenal cancers include ZNRF3 (an E3 ubiquitin ligase) and others involved in chromatin modeling and several microRNAs. Patterns of mutations may also help better define prognosis in adrenocortical cancers.104

**Symptoms and Signs** Approximately 50% of adrenocortical cancers are nonfunctioning. The remaining secrete cortisol (30%), androgens (20%), estrogens (10%), aldosterone (2%), or multiple hormones (35%). Patients with functioning tumors often present with the rapid onset of Cushing’s syndrome accompanied by virilizing features. Nonfunctioning tumors more commonly present with an enlarging abdominal mass and abdominal or back pain. Rarely, weight loss, anorexia, and nausea may be present.

**Diagnostic Tests** Diagnostic evaluation of these patients begins with measurement of serum electrolyte levels to rule out hypokalemia, urinary catecholamines to rule out pheochromocytomas, an overnight 1-mg dexamethasone suppression test, and a 24-hour urine collection for cortisol, and 17-ketosteroids to rule out Cushing’s syndrome.

CT and MRI scans are useful to image these tumors (Fig. 38-45). The size of the adrenal mass on imaging studies is the single most important criterion to help diagnose malignancy. In the series reported by Copeland, 92% of adrenal cancers were >6 cm in diameter.103 The sensitivity, specificity, and likelihood ratio of tumor size in predicting malignancy (based on Surveillance, Epidemiology, and End Results program data) were reported as 96%, 51%, and 2 for tumors ≥4 cm, and 90%, 78%, and 4.1 for tumors ≥6 cm.106 Other CT imaging characteristics suggesting malignancy include tumor heterogeneity, irregular margins, and the presence of hemorrhage and adjacent lymphadenopathy or liver metastases. Moderately bright signal intensity on T2-weighted images (adrenal mass–to–liver ratio 1.2:2.8), significant lesion enhancement, and slow washout after injection of gadolinium contrast also indicate malignancy, as does evidence of local invasion into adjacent structures such as the liver, blood vessels (IVC), and distant metastases. FDG-PET or PET-CT scans may have some utility in distinguishing benign from malignant lesions, as discussed in the section on incidentalomas. Once adrenal cancer is diagnosed, CT scans of the chest and pelvis or FDG-PET or PET-CT scans are performed for staging. The tumor-node-metastasis (TNM) staging system for adrenocortical carcinoma is depicted in Table 38-19. Up to 70% of patients present with stage III or IV disease.

**Pathology** Most adrenocortical cancers are large, weighing between 100 and 1000 g. On gross examination, areas of
hazard and necrosis often are evident. Microscopically, cells are hyperchromatic and typically have large nuclei and prominent nucleoli. It is very difficult to distinguish benign adrenal adenomas from carcinomas by histologic examination alone. Capsular or vascular invasion is the most reliable sign of cancer. Weiss and associates studied a combination of nine criteria for their usefulness in distinguishing malignant from benign adrenal tumors: nuclear grade III or IV; mitotic rate greater than 5 per 50 high-power fields; atypical mitoses; clear cells comprising 25% or less of the tumor; a diffuse architecture; microscopic necrosis; and invasion of venous, sinusoidal, and capsular structure. Tumors with four or more of these criteria were likely to metastasize and/or recur.107 Rarely, the diagnosis of malignancy of a completely resected adrenal tumor is often made in retrospect by the finding of metastatic disease. However, the therapeutic effectiveness is significantly increased recurrence-free survival in the treatment group, even in long-term follow-up.108,109 However, a study of several centers in the United States failed to show similar results.110 The routine use of this medication awaits evaluation in randomized, controlled trials. Determination of blood mitotane levels is helpful to ascertain whether therapeutic and nontoxic levels are present. Adrenocortical tumors commonly metastasize to the liver, lung, and bone.

Surgical debulking is recommended for isolated, recurrent disease and has been demonstrated to prolong survival. Systemic chemotherapeutic agents used in this tumor include etoposide, cisplatin, doxorubicin, and, more recently, paclitaxel, but consistent responses are rare, possibly due to the expression of the multidrug resistance gene (MDR-1) in tumor cells. In vitro data indicate that mitotane may be able to reverse this resistance when combined with various chemotherapeutic agents. Results from the First International Randomized Trial in Advanced or Metastatic Adrenocortical Carcinoma Treatment (FIRM-ACT) showed that patients receiving etoposide, doxorubicin, cisplatin, and mitotane had better response rates and progression-free survival rates than patients receiving streptozotocin and mitotane.111 There has been recent interest in the use of suramin, a growth factor inhibitor, as therapy for adrenocortical carcinoma; however, this requires further study, particularly because this drug may be associated with significant neurotoxicity. Gossypol, a naturally occurring insecticide (from the cotton plant Gossypium species), also appears to inhibit the growth of adrenocortical cancer cell lines and tumors in vivo. However, poor response rates combined with high death rates in limited clinical studies have reduced enthusiasm for this agent. Adrenocortical cancers also are relatively insensitive to conventional external-beam radiation therapy. However, this modality is used in the adjuvant setting in patients with incomplete resections and palliation of bony metastases. Ketoconazole, metyrapone, or aminglutethimide may also be useful in controlling steroid hypersecretion. Targeted molecular therapies such as VEGF/EGF-receptor inhibitors, tyrosine-kinase inhibitors, and IGF-2 inhibitors have had disappointing results in patients with ACC.

**Sex Steroid Excess.** Adrenal adenomas or carcinomas that secrete adrenal androgens lead to virilizing syndromes. Although women with virilizing tumors develop hirsutism, amenorrhea, infertility, and other signs of masculinization, such as increased muscle mass, deepened voice, and temporal balding, men with these tumors are more difficult to diagnose and, hence, usually present with disease in advanced stages. Children with virilizing tumors have accelerated growth, premature development of facial and pubic hair, acne, genital enlargement, and deepening of their voice. Feminizing adrenal tumors are less common and occur in men in the third to fifth decades of life. These tumors lead to gynecomastia, impotence, and testicular atrophy. Women with these tumors develop irregular menses or dysfunctional uterine bleeding. Vaginal bleeding may occur in postmenopausal women. Girls with these tumors experience precocious puberty with breast enlargement and early menarche.

**Diagnostic Tests** Virilizing tumors produce excessive amounts of the androgen precursor, DHEA, which can be measured

<table>
<thead>
<tr>
<th>TUMOR</th>
<th>NODE</th>
<th>METASTASIS</th>
<th>STAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>N0</td>
<td>M0</td>
<td>I</td>
</tr>
<tr>
<td>T2</td>
<td>N0</td>
<td>M0</td>
<td>II</td>
</tr>
<tr>
<td>T1</td>
<td>N1</td>
<td>M0</td>
<td>III</td>
</tr>
<tr>
<td>T2</td>
<td>N1</td>
<td>M0</td>
<td>III</td>
</tr>
<tr>
<td>T3</td>
<td>Any N</td>
<td>M0</td>
<td>III</td>
</tr>
<tr>
<td>T4</td>
<td>Any N</td>
<td>M0</td>
<td>III</td>
</tr>
<tr>
<td>Any T</td>
<td>Any N</td>
<td>M1</td>
<td>IV</td>
</tr>
</tbody>
</table>

Primary tumor (T): T1, size ≤5 cm without local invasion; T2, size >5 cm without local invasion; T3, any size with local invasion but no involvement of adjacent organs; T4, any size with involvement of adjacent organs.

Nodes (N): N0, no involvement of regional nodes; N1, positive regional lymph nodes.

Metastasis (M): M0, no known distal metastases; M1, distant metastases present.

in plasma or urine as 17-ketosteroids. Patients with feminizing tumors also have elevated urinary 17-ketosteroids in addition to increased estrogen levels. Androgen-producing tumors often are associated with production of other hormones such as glucocorticoids.

**Treatment** Virilizing and feminizing tumors are treated by adrenalectomy. Malignancy is difficult to diagnose histologically but is suggested by the presence of local invasion, recurrence, or distal metastases. Adrenolytic drugs such as mitotane, aminoglutethimide, and ketoconazole may be useful in controlling symptoms in patients with metastatic disease.

**Congenital Adrenal Hyperplasia.** CAH refers to a group of disorders that result from deficiencies or complete absence of enzymes involved in adrenal steroidogenesis. 21-Hydroxylase (CYP21A2) deficiency is the most common enzymatic defect, accounting for >90% of cases of CAH. This deficiency prevents the production of 11-deoxycortisol and 11-DOC from progesterone precursors. Deficiency of glucocorticoids and aldosterone leads to elevated ACTH levels and overproduction of adrenal androgens and corticosteroid precursors such as 17-hydroxyprogesterone and Δ4-androstenedione. These compounds are converted to testosterone in the peripheral tissues, thereby leading to virilization. Complete deficiency of 21-hydroxylase presents at birth with virilization, diarrhea, hypovolemia, hyponatremia, hyperkalemia, and hyperpigmentation. Partial enzyme deficiency may present at birth or later with virilizing features. These patients are less prone to the salt wasting that characterizes complete enzyme deficiency. 11β-Hydroxylase deficiency is the second most common form of CAH and leads to hypertension (from 11-DOC accumulation), virilization, and hyperpigmentation. Other enzyme deficiencies include 3β-hydroxydehydrogenase and 17-hydroxylase deficiency. Congenital adrenal lipoid hyperplasia is the most severe form of CAH, which is caused by cholesterol desmolase deficiency. It leads to the disruption of all steroid biosynthetic pathways, thus resulting in a fatal salt-wasting syndrome in phenotypic female patients.

**Diagnostic Tests** The particular enzyme deficiency can be diagnosed by karyotype analysis and measurement of plasma and urinary steroids. The most common enzyme deficiency, absence of 21-hydroxylase, leads to increased plasma 17-hydroxyprogesterone and progesterone levels because these compounds cannot be converted to 11-deoxycortisol and 11-DOC, respectively. 11β-Hydroxylase deficiency is the next most common disorder and results in elevated plasma 11-DOC and 11-deoxycortisol. Urinary 17-hydroxyprogesterone, androgens, and 17-ketosteroids also are elevated. The dexamethasone suppression test (2–4 mg divided four times a day for 7 days) can be used to distinguish adrenal hyperplasia from neoplasia. CT, MRI, and iodocholesterol scans generally are used to localize the tumors.

**Treatment** Patients with CAH traditionally have been managed medically, with cortisol and mineralocorticoid replacement to suppress the hypothalamic-pituitary-adrenal axis. However, the doses of steroids required often are supraphysiologic and lead to iatrogenic hypercortisolism. More recently, bilateral laparoscopic adrenalectomy has been proposed as an alternative treatment for this disease and has been successfully performed in a limited number of patients for various forms of CAH.

---

**Disorders of the Adrenal Medulla**

**Pheochromocytomas.** Pheochromocytomas are rare tumors with prevalence rates ranging from 0.3% to 0.95% in autopsy series and approximately 1.9% in series using biochemical screening. They can occur at any age, with a peak incidence in the fourth and fifth decades of life, and have no gender predilection. Extra-adrenal tumors, also called functional paragangliomas, may be found at sites of sympathetic ganglia in the organ of Zuckerkandl, neck, mediastinum, abdomen, and pelvis. Pheochromocytomas often are called the "10 percent tumor" because 10% are bilateral, 10% are malignant, 10% occur in pediatric patients, 10% are extra-adrenal, and 10% are familial.

Pheochromocytomas occur in families with MEN2A and MEN2B in approximately 50% of patients. Both syndromes are inherited in an autosomal dominant fashion and are caused by germline mutations in the RET proto-oncogene. Another syndrome with an increased risk of pheochromocytomas is von Hippel-Lindau (VHL) disease, which also is inherited in an autosomal dominant manner. This syndrome also includes retinal angioma, hemangioblastomas of the central nervous system, renal cysts and carcinomas, pancreatic cysts, and epilidymal cystadenomas. The incidence of pheochromocytomas in the syndrome is approximately 14%. The gene causing VHL has been mapped to chromosome 3p and is a tumor suppressor gene. Pheochromocytomas also are included within the tumor spectrum of neurofibromatosis type 1 (NF1 gene) and other neuroectodermal disorders (Sturge-Weber syndrome and tuberous sclerosis), Carney’s syndrome (gastric epithelioid leiomyosarcoma, pulmonary chondroma, and extra-adrenal paraganglioma), MEN1 syndrome, and the familial paraganglioma and pheochromocytoma syndrome are caused by mutations in the succinyl dehydrogenase family of genes (SDHB, SDHC, and SDHD), which comprise portions of the mitochondrial complex II. More recently, mutations in SDHA and SDHB have also been identified. Additional susceptibility loci include TMEM127 (involved in the mTORC1-signaling pathway) and MAX (myc-associated factor X).

**Symptoms and Signs** Headache, palpitations, and diaphoresis constitute the “classic triad” of pheochromocytomas. Symptoms such as anxiety, tremulousness, paresthesias, flushing, chest pain, shortness of breath, abdominal pain, nausea, vomiting, and others are nonspecific and may be episodic in nature. Cardiovascular complications such as myocardial infarction and cerebrovascular accidents may ensue. These symptoms can be incited by a range of stimuli including exercise, micturition, and defecation. The most common clinical sign is hypertension. Pheochromocytomas are one of the few curable causes of hypertension and are found in 0.1% to 0.2% of hypertensive patients. Hypertension related to this tumor may be paroxysmal with intervening normotension, sustained with paroxysms or sustained hypertension alone. Sudden death may occur in patients with undiagnosed tumors who undergo other surgeries or biopsy.

**Diagnostic Tests**

**Biochemical Studies.** Pheochromocytomas are diagnosed by testing 24-hour urine samples for catecholamines and their metabolites as well as by determining plasma metanephrine levels. Urinary metanephrines are 98% sensitive and also about 98% specific for pheochromocytomas, whereas VMA measurements are slightly less sensitive and specific. False-positive VMA tests may result from ingestion of caffeine, raw fruits,
or medications (α-methyldopa). Fractionated urinary catecholamines (norepinephrine, epinephrine, and dopamine) also are very sensitive but less specific for pheochromocytomas. Because extra-adrenal sites lack phenylethanolamine N-methyltransferase, these tumors secrete norepinephrine, whereas epinephrine is the main hormone secreted from adrenal pheochromocytomas. Many physiologic and pathologic states can alter the levels of plasma catecholamines. Hence, they often are thought to be less accurate than urinary tests. Both epinephrine and norepinephrine should be measured, as tumors often secrete one or the other hormone. Sensitivities of 85% and specificities of 95% have been reported using cutoff values of 2000 pg/mL for norepinephrine and 200 pg/mL for epinephrine. Clonidine is an agent that suppresses neurogenically mediated catecholamine excess but not secretion from pheochromocytomas. A normal clonidine suppression test is defined by a decrease of basal catecholamine levels to <500 pg/mL within 2 to 3 hours after an oral dose of 0.3 mg of clonidine. Chromogranin A is a monomeric, acidic protein, which is stored in the adrenal medulla and other neuroendocrine tumors and released along with catecholamine hormones. It has been reported to have a sensitivity of 83% and a specificity of 96% and is useful in conjunction with catecholamine measurement for diagnosing pheochromocytomas. Some studies have shown that plasma metanephrines should be the first-line test to identify pheochromocytomas, as the predictive value of a negative test is very high and normal levels exclude pheochromocytoma in patients with preclinical disease or dopamine secreting tumors. Although sensitivities of 96% to 100% have been reported, specificity is lower at 85% to 89% and may be much lower at 77% in elderly patients. Although attractive because of the simplicity of a blood test, measurement of plasma metanephrines is generally reserved for cases for which there is a high index of suspicion. 

Radiologic Studies. Radiologic studies are useful to localize tumors and to assess the extent of spread once the diagnosis has been made with biochemical tests. CT scans are 85% to 95% sensitive and 70% to 100% specific for pheochromocytomas (Fig. 38-46A). The scans should be performed without contrast to minimize the risk of precipitating a hypertensive crisis, although some recent studies suggest that intravenous contrast may be used. Images should include the region from the diaphragm to the aortic bifurcation so as to include the organ of Zuckerkandl. CT scans do not provide functional information and cannot definitively diagnose pheochromocytomas. MRI scans are 95% sensitive and almost 100% specific for pheochromocytomas because these tumors have a characteristic appearance on T2-weighted images or after gadolinium. MRI is also the study of choice in pregnant women as there is no risk of radiation exposure. Metaiodobenzylguanidine (MIBG) is taken up and concentrated by vesicles in the adrenal medullary cells because its structure is similar to norepinephrine. Normal adrenal medullary tissue does not take up appreciable MIBG. 131I-radiolabeled MIBG is, therefore, useful for localizing pheochromocytomas (Fig. 38-46B), especially those in ectopic positions. This test has a reported sensitivity of 77% to 89% and a specificity ranging from 88% to 100%.

Treatment. The medical management of pheochromocytomas is aimed chiefly at blood pressure control and volume repletion. Irreversible, long-acting α-blockers such as phenoxycyzbenzamine are started 1 to 3 weeks before surgery at doses of 10 mg twice daily, which may be increased to 300 to 400 mg/d with rehydration. Patients should be warned about orthostatic hypotension. β-Blockers such as propranolol at doses of 10 to 40 mg every 6 to 8 hours often need to be added preoperatively in patients who have persistent tachycardia and arrhythmias. β-Blockers should only be instituted after adequate α-blockade and hydration to avoid the effects of unopposed α-stimulation (i.e., hypertensive crisis and congestive heart failure) and are typically initiated 3 to 4 days preoperatively. Patients also should be volume repleted preoperatively to avoid postoperative hypotension, which ensues with the loss of vasoconstriction after tumor removal. Other α-blockers such as prazosin, terazosin, and doxazosin, which are selective α1-adrenergic blockers, have a better side effect profile and are preferable to phenoxybenzamine when long-term

**Figure 38-46.** A left-sided pheochromocytoma (arrows) imaged by a computed tomography scan of the abdomen (A) and a metaiodobenzylguanidine scan viewed posteriorly (B).
pharmacologic therapy is needed, as in patients with metastatic pheochromocytoma. Nicardipine is the most commonly used calcium channel blocker and inhibits norepinephrine-mediated calcium transport into vascular smooth muscle. When used as the primary mode of treatment, it appears to be just as effective as α- and β-blockade preoperatively and for intraoperative hemodynamics.[113] In some patients, catecholamine-synthesis inhibitors such as α-methyl-p-l-tyrosine (methyrosine) may need to be added if standard α- and β-blockade is poorly tolerated or is ineffective in reaching target blood pressure and when moderate intraoperative tumor manipulation is anticipated.

Adrenalectomy is the treatment of choice for patients with pheochromocytoma. The chief goal of surgery is to resect the tumor completely with minimal tumor manipulation or rupture of the tumor capsule. Surgery should be performed with both noninvasive and invasive monitors, including an arterial line and central venous lines. In patients with congestive heart failure or underlying coronary artery disease, Swan-Ganz catheters may be necessary. Stress must be avoided during anesthesia induction, and use of inhaled agents like isoflurane and enflurane are preferred because they have minimal cardiac depressant effects. Fentanyl, ketamine, and morphine should be avoided as they can potentially stimulate catecholamine release from the tumor. The common medications used for intraoperative blood pressure control include nitroprusside, nitroglycerin, phentolamine, and nicardipine. Intraoperative arrhythmias are best managed by short-acting β-blockers such as esmolol. Adrenalectomy usually was performed via an open anterior approach to facilitate detection of bilateral tumors, extra-adrenal lesions, or metastatic lesions. However, most pheochromocytomas <5 cm in diameter can be safely resected laparoscopically. Postoperatively, these patients are prone to hypotension due to loss of adrenergic stimulation and consequent vasodilatation and therefore need large volume resuscitation.

Hereditary Pheochromocytomas. Inherited pheochromocytomas tend to be multiple and bilateral. Generally, unilateral adrenalectomy is recommended in the absence of obvious lesions in the contralateral adrenal gland because of Addison’s disease, requiring lifelong steroid replacement in patients undergoing bilateral adrenalectomy. For patients with tumors in both adrenal glands, cortical-sparing subtotal adrenalectomy may preserve adrenocortical function and avoid the morbidity of bilateral total adrenalectomy. Laparoscopic subtotal adrenalectomy has been shown to provide short-term clinical results comparable to total adrenalectomy, with reduced surgical morbidity and may be done as a unilateral or bilateral subtotal procedure.[114] However, these patients remain at risk for recurrent pheochromocytoma, which has been reported in 20% of patients with VHL disease a median of 40 months after partial adrenalectomy, and in 33% of MEN2 patients followed for 54 to 88 months after surgery. Autotransplantation of adrenocortical tissue after total adrenalectomy may be another option for these patients. However, the transplanted cortical tissue rarely provides full function, and steroid replacement usually is required.

Malignant Pheochromocytomas. Approximately 12% to 29% of pheochromocytomas are malignant, and these tumors are associated with decreased survival. There are no definitive histologic criteria defining malignant pheochromocytomas. In fact, pleomorphism, nuclear atypia, and abundant mitotic figures are seen in benign tumors. Capsular and vascular invasion may be seen in benign lesions as well. Malignancy usually is diagnosed when there is evidence of invasion into surrounding structures or distant metastases. The most common sites for metastatic disease are bone, liver, regional lymph nodes, lung, and peritoneum, although the brain, pleura, skin, and muscles may also occasionally be involved. Some studies also suggest that older patient age and larger tumors are associated with a higher risk of malignancy. Although risk of malignancy increases with size for all pheochromocytomas, size does not seem to reliably predict malignancy in pheochromocytomas with local disease only.[115] Given this difficulty of defining malignancy clinically (in the absence of metastatic disease), a number of other features such as DNA ploidy, tumor size, and necrosis, neuropeptide Y mRNA expression, and serum neuron-specific enolase expression have been studied. Malignant pheochromocytomas are more likely to express p53 and bcl-2 and have activated telomerase. Recent data suggest that flow cytometry and molecular markers such as expression of Ki-67, tissue inhibitor of metalloproteinase, and COX-2 also have shown some use in determining malignancy. When pheochromocytomas develop in the MEN syndromes, they rarely are malignant. In contrast, patients with germline SDHB mutations appear to have a higher propensity for extra-adrenal and malignant tumors. In general, soft tissue lesions are treated with resection if feasible. External-beam radiation can be used for unresectable lesions or symptomatic skeletal metastases. Therapeutic 131I-MIBG irradiation may be useful in patients with diffuse disease showing 123I-MIBG uptake on a diagnostic scan. Chemotherapy regimens typically use cyclophosphamide, vincristine, and dacarbazine with variable response rates. However, molecular targeted therapies such as sunitinib have shown some promising results.

The Adrenal Incidentaloma
Adrenal lesions discovered during imaging performed for unrelated reasons are referred to as incidentalomas. This definition excludes tumors discovered on imaging studies performed for evaluating symptoms of hormone hypersecretion or staging patients with known cancer. The incidence of these lesions identified by CT scans ranges from 0.4% to 4.4%.

Differential Diagnosis. The differential diagnosis of adrenal incidentalomas is shown in Table 38-20. Nonfunctional cortical adenomas account for the majority (36–94%) of adrenal incidentalomas in patients without a history of cancer. In a series of patients from the Mayo Clinic, no nonfunctional

| Table 38-20 |
| Differential diagnosis of adrenal incidentaloma |
| **FUNCTIONING LESIONS** | **NONFUNCTIONING LESIONS** |
| Benign | Benign |
| Aldosteronoma | Cortical adenoma |
| Cortisol-producing adenoma | Myelolipoma |
| Sex steroid–producing adenoma | Cyst |
| Pheochromocytoma | Ganglioneuroma |
| Hemorrhage | |
| | Malignant |
| Adrenocortical cancer | Metastasis |
| Malignant pheochromocytoma | Adrenocortical cancer |
lesion progressed to cause clinical or biochemical abnormalities. However, other studies indicate that 5% to 20% of patients with apparently nonfunctioning cortical adenomas have underlying, subtle abnormalities of glucocorticoid secretion, and a rare benign-appearing incidentaloma is a cancer.

By definition, patients with incidentalomas do not have clinically overt Cushing’s syndrome, but subclinical Cushing’s syndrome is estimated to occur in approximately 8% of patients. This disorder is characterized by subtle features of cortisol excess, such as weight gain, skin atrophy, facial fullness, diabetes, and hypertension, accompanied by loss of normal diurnal variation in cortisol secretion, autonomous cortisol secretion, and resistance to suppression by dexamethasone. Total cortisol produced and 24-hour urinary cortisol levels may be normal. Examination of the natural history of subclinical Cushing’s syndrome indicates that, although most patients remain asymptomatic, some do progress to clinically evident Cushing’s syndrome. Furthermore, cases of postoperative adrenal crisis from unrecognized suppression of the contralateral adrenal have been reported, making preoperative identification of this condition imperative, particularly in the era of early discharge following laparoscopic adrenalectomy.

The adrenal is a common site of metastases of lung and breast tumors, melanoma, renal cell cancer, and lymphoma. In patients with a history of nonadrenal cancer and a unilateral adrenal mass, the incidence of metastatic disease has been reported to range from 32% to 73%. Myelolipomas are benign, biochemically nonfunctioning lesions composed of elements of hematopoietic and mature adipose tissue, which are rare causes of adrenal incidentaloma. Other less commonly encountered lesions include adrenal cysts, ganglioneuromas, and hemorrhage.

**Diagnostic Investigations.** The diagnostic workup of an adrenal incidentaloma is aimed at identifying patients who would benefit from adrenalectomy (i.e., patients with functioning tumors and tumors at increased risk of being malignant). It is not necessary for asymptomatic patients whose imaging studies are consistent with obvious cysts, hemorrhage, myelolipomas, or diffuse metastatic disease to undergo additional investigations. All other patients should be tested for underlying hormonally active tumors using (a) a low-dose (1 mg) overnight dexamethasone suppression test to rule out subclinical Cushing’s syndrome and 17-ketosteroids (if sex steroid excess is suspected); (b) a 24-hour urine collection for catecholamines, metanephrines, VMA, or plasma metanephrine to rule out pheochromocytoma; and (c) in hypertensive patients, serum electrolytes, plasma aldosterone, and plasma renin to rule out an aldosteronoma. In patients with a high index of suspicion for subclinical Cushing’s (those with hypertension, obesity, or diabetes), three tests (i.e., dexamethasone suppression test, salivary cortisol, and 24-hour urine free cortisol) may be used. Confirmatory tests can be performed based on the results of the initial screening studies.

Determination of the malignant potential of an incidentaloma is more difficult. The risk of malignancy in an adrenal lesion is related to its size. Lesions >6 cm in diameter have an approximate risk of malignancy of about 25%. However, this size cutoff is not absolute because adrenal carcinomas also have been reported in lesions <6 cm. Carcinomas account for 2% of lesions <4 cm and 6% of lesions 4.1 to 6 cm in size. This has led to increased use of the imaging characteristics of incidentalomas to predict malignancy. Benign adrenal adenomas tend to be homogeneous, well encapsulated, and have smooth and regular margins. They also tend to be hypoattenuating lesions (<10 Hounsfield units) on CT scanning. In contrast, adrenal cancers tend to be hyperattenuating (>18 Hounsfield units) and inhomogeneous, have irregular borders, and may show evidence of local invasion or adjacent lymphadenopathy. On MRI T2-weighted imaging, adenomas demonstrate low signal intensity when compared to the liver (adrenal mass-to-liver ratio <1.4), whereas carcinomas and metastases have moderate intensity (mass-to-liver ratio 1.2:2.8). Pheochromocytomas are extremely bright, with mass-to-liver ratios >3. Unfortunately, the ranges overlap, and signal intensity is not 100% reliable for determining the nature of the lesion. Radionuclide imaging with NP-59 also has been used to distinguish between various adrenal lesions, with some investigators suggesting that uptake of NP-59 was 100% predictive of a benign lesion (adenoma), whereas absence of imaging was 100% predictive of a nonadenomatous lesion. However, the technique has not gained widespread acceptance because patients need to be given cold iodine 1 week before the study to prevent thyroid uptake, imaging needs to be delayed by 5 to 7 days after administration of the contrast, and false-positive and false-negative results occur. FDG-PET or PET-CT scans may have some utility in distinguishing potentially malignant from benign lesions in cases of inconclusive CT densitometry. However, caution must be exercised for false-positive (some adenomas and pheochromocytomas) and false-negative results (small lesions or those with hemorrhage or necrosis). FNAB cannot be used to distinguish adrenal adenomas from carcinomas. This being said, FNAB is useful in the setting of a patient with a history of cancer and a solitary adrenal mass. The positive predictive value of FNAB in this situation has been shown to be almost 100%, although false-negative rates of up to 33% have been reported. Biopsies usually are performed under CT guidance, and appropriate testing to rule out pheochromocytomas should be undertaken before the procedure to avoid precipitating a hypertensive crisis.

**Management.** An algorithm for the management of patients with incidentalomas is shown in Fig. 38-47. The AACE and American Association of Endocrine Surgeons (AAES) have published management guidelines for patients with adrenal incidentalomas. Patients with functional tumors or obviously malignant lesions should undergo adrenalectomy. The optimal management of patients with subclinical Cushing’s syndrome is controversial, especially due to the paucity of data from high-quality prospective trials. In general, operative intervention is advised in patients with subclinical Cushing’s syndrome with suppressed plasma ACTH levels and elevated urinary cortisol levels because these patients are at high risk for progression to overt Cushing’s syndrome. The adrenal incidentaloma guidelines also recommend adrenalectomy in patients with worsening hypertension, abnormal glucose tolerance, or osteoporosis.

For nonfunctional lesions, the risk of malignancy needs to be balanced with operative morbidity and mortality. The AACE/AAES guidelines recommend that lesions with suspicious features on imaging studies such as heterogeneity, irregular capsule, or adjacent nodes should be treated by adrenalectomy. Nonoperative therapy, with close periodic follow-up, is advised for lesions <4 cm in diameter with benign imaging characteristics, whereas adrenalectomy is recommended for lesions ≥4 cm in size due to the increased risk of cancer.
Questions
1) Low dose DST
2) Plasma metanephrines or 24 h urine catecholamines, VMA, metanephrines
3) Plasma aldosterone, renin activity, electrolytes

Adrenalectomy
Yes
No
Past history of cancer?
Yes
Solitary metastasis
Consider adrenalectomy
No
Systemic therapy

≥4 cm, Indeterminate or suspicious imaging features
Adrenalectomy

<4 cm, Benign imaging features
Repeat imaging in 3–6 months, biochemical evaluation annually

Figure 38-47. Management algorithm for an adrenal incidentaloma. CT = computed tomography; DST = dexamethasone suppression test; VMA = vanillylmandelic acid.

However, several important points must be considered in the management of these patients. First, size criteria for malignancy are not definitive and are derived from a selected series of patients. Second, the actual size of adrenal tumors can be underestimated by at least 1 cm by modalities such as CT and MRI scans because tumors are larger in a cephalocaudal axis. Third, the natural history of incidentalomas is variable and depends on the underlying diagnosis, age of the study population, and the size of the mass. Older patients are more likely to have nonfunctioning adenomas. Existing data in terms of the long-term behavior of these nonfunctional lesions, although limited, indicate that malignant transformation is uncommon. Furthermore, tumors that increase in size by at least 1 cm over a 2-year follow-up period and those with subtle hormonal abnormalities appear to be more likely to enlarge. Overt hormone overproduction is more likely in tumors >3 cm and those with increased NP-59 uptake. Surgeons are more likely to operate on a 40-year-old patient with a 4-cm lesion, while electing to follow an 80-year-old patient with a similar lesion but multiple concurrent comorbidities. Based on the above considerations, some surgeons use a size threshold for adrenalectomy with a nonfunctioning homogeneous tumor of 3 to 4 cm in young patients with no comorbidities and 5 cm in older patients with significant comorbidity.

Lesions that grow during follow-up also are treated by adrenalectomy. Myelolipomas generally do not warrant adrenalectomy unless there is concern regarding malignancy, which is rare, or bleeding into the lesion, which is more likely in myelolipomas >4 cm in size. These tumors, even when large, can be removed laparoscopically. Resection of solitary adrenal metastases in patients with a history of nonadrenal cancer has been demonstrated to lead to prolonged patient survival. Suspected adrenal metastases also may be resected for diagnosis.

There is no consensus regarding the follow-up of patients with adrenal incidentaloma. The AACE/AAES guidelines recommend repeating hormonal screening with a 1-mg dexamethasone suppression test and urinary catecholamines and metabolites yearly for 5 years as the risk of hypersecretion appears to plateau after this time period. It also recommends repeat imaging at 3 to 6 months and then annually for 1 to 2 years. Less frequent imaging is reasonable or small (<2 cm), uniform, hypodense cortical nodules in patients without a history of malignant disease. Adrenalectomy is recommended for lesions that grow ≥1 cm or if autonomous hormone secretion develops during follow-up.

Adrenal Insufficiency
Adrenal insufficiency may be primary, resulting from adrenal disease, or secondary, due to a deficiency of ACTH (Table 38-21). The most commonly encountered causes of primary adrenal insufficiency are autoimmune disease, infections, and metastatic deposits. Spontaneous adrenal hemorrhage can occur in patients with fulminant meningococcal septicemia (Waterhouse-Friderichsen syndrome). Bilateral adrenal hemorrhage also can occur secondary to trauma, severe stress, infection, and coagulopathies and, if unrecognized, is lethal. Exogenous glucocorticoid therapy with suppression of the adrenal glands is the most common cause of secondary adrenal insufficiency.

Symptoms and Signs. Acute adrenal insufficiency should be suspected in stressed patients with any of the relevant risk factors. It may mimic sepsis, myocardial infarction, or pulmonary embolus and presents with fever, weakness, confusion, nausea, vomiting, lethargy, abdominal pain, or severe hypotension. Chronic adrenal insufficiency, such as that occurring in patients with metastatic tumors, may be more subtle. Symptoms include fatigue, salt craving, weight loss, nausea, vomiting, and
abdominal pain. These patients may appear hyperpigmented from increased secretion of CRH and ACTH, with increased α-melanocyte-stimulating hormone side-products.

**Diagnostic Studies.** Characteristic laboratory findings include hyponatremia, hyperkalemia, eosinophilia, mild azotemia, and fasting or reactive hypoglycemia. The peripheral blood smear may demonstrate eosinophilia in approximately 20% of patients. Adrenal insufficiency is diagnosed by the ACTH stimulation test. ACTH (250 μg) is infused intravenously, and cortisol levels are measured at 0, 30, and 60 minutes. Peak cortisol levels <20 μg/dL suggest adrenal insufficiency. ACTH levels also allow primary insufficiency to be distinguished from secondary causes. High ACTH levels with low plasma cortisol levels are diagnostic of primary adrenal insufficiency.

**Treatment.** Treatment must be initiated based on clinical suspicion alone, even before test results are obtained, or the patient is unlikely to survive. Management includes volume resuscitation with at least 2 to 3 L of a 0.9% saline solution or 5% dextrose in saline solution. Blood should be obtained for electrolyte (decreased Na+ and increased K+), glucose (low), and cortisol (low) levels; ACTH (increased in primary and decreased in secondary); and quantitative eosinophilic count. Dexamethasone (4 mg) should be administered intravenously. Hydrocortisone (100 mg intravenously every 8 hours) also may be used, but it interferes with testing of cortisol levels. Once the patient has been stabilized, underlying conditions such as infection should be sought, identified, and treated. The ACTH stimulation test should be performed to confirm the diagnosis. Glucocorticoids can then be tapered to maintenance doses (oral hydrocortisone 15–20 mg in the morning and 10 mg in the evening). Mineralocorticoids (fludrocortisone 0.05–0.1 mg daily) may be required once the saline infusions are discontinued.

**Adrenal Surgery**

**Choice of Procedure.** Adrenalectomy may be performed via a laparoscopic or open approach. In either approach, the gland may be approached anteriorly, laterally, or posteriorly via the retroperitoneum. The choice of approach depends on the size and nature of the lesion and expertise of the surgeon. Laparoscopic adrenalectomy has rapidly become the standard procedure of choice for the excision of most benign-appearing adrenal lesions <6 cm in diameter. The role of laparoscopic adrenalectomy in the management of adrenocortical cancers is controversial. The data with respect to local tumor recurrence and intra-abdominal carcinomatosis from laparoscopic adrenalectomy for malignant adrenal tumors that were not appreciated as such, preoperatively or intraoperatively, are conflicting. Although laparoscopic adrenalectomy appears to be feasible and safe for solitary adrenal metastasis116 (provided there is no local invasion and the tumor can be resected intact), open adrenalectomy or laparoscopic-assisted open adrenalectomy is the safest option for suspected or known adrenocortical cancers and malignant pheochromocytomas. Technical considerations and surgeon experience, rather than absolute tumor size, usually determine the size threshold for laparoscopic resection. Hand-assisted laparoscopic adrenalectomy may provide a bridge between laparoscopic adrenalectomy and conversion to an open procedure. There have been no randomized trials directly comparing open vs. laparoscopic adrenalectomy. However, studies have shown that laparoscopic adrenalectomy is associated with decreased blood loss, postoperative pain, and narcotic use; reduced length of hospital stay; and faster return to work.

**Laparoscopic Adrenalectomy.** The procedure is performed under general anesthesia. Arterial lines are used routinely, and central lines are necessary for patients in whom massive fluid shifts are anticipated (e.g., those with large, active pheochromocytomas). A nasogastric tube and Foley catheter are recommended. Routine preoperative antibiotics are not needed, except in patients with Cushing’s syndrome. The adrenals can be removed laparoscopically via a transabdominal (anterior or lateral) or retroperitoneal (lateral or posterior) approach. The lateral approach is preferred by most laparoscopic surgeons and uses gravity to aid retraction of surrounding organs. Patients, however, need to be repositioned for a bilateral procedure. The anterior transabdominal approach offers the advantage of a conventional view of the abdominal cavity and allows a bilateral adrenalectomy to be performed without the necessity of repositioning the patient. The posterior retroperitoneal approach has also been gaining popularity in recent years, particularly in patients with previous anterior abdominal surgery and peritoneal adhesions. In addition, several centers have successfully utilized robotic approaches for both lateral transabdominal and retroperitoneal laparoscopic adrenal surgery. Single incision laparoscopic adrenalectomy is another option. While these latter approaches are feasible,119 their widespread use awaits analysis of long-term outcomes data and cost analyses. The lateral transabdominal approach is widely used and described in detail in the following section.
Lateral Transabdominal Approach  The patient is placed in the lateral decubitus position, and the operating table is flexed at the waist to open the space between the lower rib cage and the iliac crest (Fig. 38-48). The surgeon and assistant both stand on the same side, facing the front of the patient. Pneumoperitoneum is created using a Veress needle or insufflation via a Hasson port. In general, four 10-mm trocars are placed between the midclavicular line medially and anterior axillary line laterally, one to two fingerbreadths below the costal margin (see Fig. 38-48), although additional ports may be placed, if needed. A 30° laparoscope is inserted through the second or midclavicular port. Most of the dissection is carried out via the two most lateral ports. However, the instruments and ports may be changed to provide optimum exposure, as needed.

For a right adrenalectomy, a fan retractor is inserted through the most medial port to retract the liver. An atraumatic grasper and an L-hook cautery are inserted via the two lateral ports for the dissection. The right triangular ligament is divided, and the liver is rotated medially (Fig. 38-49A). Rarely, the hepatic flexure of the colon may need mobilization during a right adrenalectomy. The right kidney is identified visually and by palpation with an atraumatic grasper. The adrenal gland is identified on the superomedial aspect of the kidney. Gerota’s fascia is incised with the hook cautery. Dissection of the adrenal is started superomedially and then proceeds inferiorly, dissecting around the adrenal in a clockwise manner. The periadrenal tissues are grasped or moved with a blunt grasper to facilitate circumferential dissection. The right adrenal vein is identified at its junction with the IVC, ligated with clips, and divided using endoscopic scissors. Alternatively, a vascular stapler may be used to divide the vein endoscopically. There may be a second adrenal vein on the right. Generally, two clips are left on the vena cava side. Although early identification of the adrenal vein is helpful to facilitate mobilization and prevent injury, it can be dissected whenever it is safe to do so. Early ligation of the adrenal vein makes it easier to mobilize the gland but may make subsequent dissection more difficult due to venous congestion. The arterial branches to the adrenal gland can be electrocoagulated if small or clipped and divided.

For a left adrenalectomy, the fan retractor is used to retract the spleen. The splenic flexure is mobilized early, and the lateral attachments to the spleen and the tail of the pancreas are divided using the electrocautery (Fig. 38-49B). Gravity allows the spleen and the pancreatic tail to fall medially. The remainder of the dissection proceeds similarly to that described for the right adrenal. In addition to the adrenal vein, the inferior phrenic vein, which joins the left adrenal vein medially, also needs to be
dissected, doubly clipped, and divided. As with the right adrenal vein, the left-sided veins also can be divided with a vascular stapler. Once the dissection is complete, the area of the adrenal bed can be irrigated and suctioned. A drain is rarely necessary. The gland is placed in a nylon specimen bag, which is brought out via one of the ports after morcellation, if necessary.

**Posterior Retroperitoneal Approach** The retroperitoneal approach provides a more direct access to the adrenal gland and avoids abdominal adhesions in patients who have had previous abdominal surgery. Furthermore, bilateral adrenalectomy can be performed without repositioning the patient. Intraoperative ultrasound is helpful for identifying the adrenal, but the dissection and exposure are more difficult because the working space is limited. This makes vascular control difficult and also renders it unsuitable for large (>5 cm) lesions. This technique is being increasingly used for small adenomas causing hyperaldosteronism.

The patient is placed in the prone-jackknife position, and the operating table is flexed at the waist to open the space between the posterior costal margin and the pelvis. Palpation is used to identify the position of the twelfth rib. Percutaneous ultrasound is performed to determine the outline of the underlying kidney and adrenal. When done laparoscopically, the surgeon stands on the side of the adrenal to be removed, and the assistant stands on the opposite side. A 1.5-cm incision is placed 2 cm inferior and parallel to the twelfth rib, laterally at the level of the inferior pole of the kidney. Gerota’s space is entered under direct vision using a 12-mm direct viewing trocar with a 0° laparoscope through the muscle layers of the posterior abdominal wall. Alternatively, blunt dissection with the surgeon’s finger also can identify the space behind Gerota’s fascia. The trocar is then replaced by a dissecting balloon, which is manually inflated using a hand pump under direct vision through the laparoscope. A 12-mm trocar is then reinserted into this space, and CO₂ is insufflated to 12 to 15 mmHg pressure. The 0° laparoscope is replaced by a 45° laparoscope. Two additional 5- or 10-mm trocars are placed, one each on either side of the first port. Laparoscopic ultrasound then is used to help locate the adrenal gland and vessels. The adrenal dissection is begun at the superior pole and then proceeds to the lateral and inferior aspect. The medial dissection usually is performed last, and the vessels are identified and divided as described in the earlier “Lateral Transabdominal Approach” section.

**Open Adrenalectomy.** Open adrenalectomy may be performed via four approaches, each with specific advantages and disadvantages. The anterior approach allows examination of the abdominal cavity and resection of bilateral tumors via a single incision. The posterior approach avoids the morbidity of a laparotomy incision, especially in patients with cardiopulmonary disease and those prone to wound complications (Cushing’s syndrome) and avoids abdominal adhesions in patients who have undergone previous abdominal surgery. Recovery time is also quicker and hospitalization shorter. However, the retroperitoneal exposure is difficult, particularly in obese patients, and the small working space makes it unsuitable for tumors >6 cm in diameter. The lateral approach is best for obese patients and for large tumors because it provides a bigger working space. The thoracoabdominal approach is most useful for en bloc resection of large (>10 cm), malignant lesions. However, it is associated with significant morbidity and should be used selectively.

**Anterior Approach** The adrenals may be removed via a midline incision or bilateral subcostal incision (Fig. 38-50). The former allows adequate infraumbilical exposure for examination of extra-adrenal tumors, whereas the latter provides better superior and lateral exposure. For the right side, the hepatic flexure of the colon is mobilized inferiorly, and the triangular ligament is incised to retract the liver medially and superiorly. A generous Kocher maneuver is used to mobilize the duodenum anteriorly and expose the retroperitoneal fat and the IVC (Fig. 38-51A). Gerota’s fascia is incised, and the gland is freed of surrounding fibro-fatty tissue and the kidney inferiorly. The lateral and superior surfaces usually are mobilized first. Then, the short, right adrenal vein is dissected, ligated, and divided, taking care not to injure the hepatic veins and IVC. On the left side, the adrenal is located cephalad to the pancreatic tail and just lateral to the aorta. For large tumors, the adrenal is best approached by medial visceral rotation to mobilize the spleen, colon, and pancreas toward the midline (Fig. 38-51B). An alternative approach is to enter the lesser sac by division of the gastrocolic ligament. The pancreas is mobilized superiority by incision of its inferior
peritoneal attachments, thus exposing the left kidney and adrenal. The gland is then mobilized as on the right side.

**Posterior Approach** The patient is placed prone on the operating table, similar to the laparoscopic approach. A hockey stick or curvilinear incision may be used, and extended through the latissimus dorsi and sacrospinalis fascia. The twelfth rib generally is excised at its base, and the eleventh rib is retracted superiorly to reveal the pleura and the lateral arcuate ligament of the liver on the right side. The pleura also is mobilized cephalad, and the adrenal and kidney are identified. The superior aspect of the gland is dissected first, and the superior vessels are identified and ligated. This prevents superior retraction of the adrenal gland. The remainder of the gland is then dissected and the adrenal gland and tumor removed. The resulting space generally is filled with perinephric fat and closed in layers. A chest X-ray is obtained postoperatively to rule out a pneumothorax.

**Lateral Approach** The patient is placed in a lateral position with the table flexed, and an incision is made between the eleventh and twelfth ribs or subcostally. The dissection then is performed as indicated previously in “Anterior Approach.”

**Complications of Adrenal Surgery.** Patients with Cushing’s syndrome are more prone to infectious (incisional and intra-abdominal abscess) and thrombotic complications. Creation of pneumoperitoneum may result in injury to various organs from Veress needle and trocar insertion, subcutaneous emphysema, pneumothorax, and hemodynamic compromise. Excessive retraction and dissection may lead to bleeding from injury to the IVC and renal vessels, or from injury to surrounding organs such as the liver, pancreas, spleen, and stomach. Postoperative hemodynamic instability may be evident in patients with pheochromocytomas, and patients are at risk of adrenal insufficiency after bilateral adrenalectomy and sometimes after unilateral adrenalectomy (unrecognized Cushing’s syndrome or, very rarely, Conn’s syndrome). Long-term morbidity results mainly from injury to nerve roots during trocar insertion, which can lead to chronic pain syndromes or muscle weakness, although this is more of an issue in case of open procedures.

Approximately 30% of patients who undergo bilateral adrenalectomy for Cushing’s disease are at risk of developing Nelson’s syndrome from progressive growth of the preexisting pituitary tumor. This leads to increased ACTH levels, hyperpigmentation, visual field defects, headaches, and extraocular muscle palsies. Transsphenoidal pituitary resection is the initial mode of therapy, and external-beam radiotherapy is used in patients with residual tumor or extrasellar invasion.

**REFERENCES**

Entries highlighted in bright blue are key references.


INTRODUCTION

In his 1953 classic textbook entitled The Surgery of Infancy and Childhood, Dr. Robert E. Gross summarized the essential challenge of pediatric surgery: “Those who daily operate upon adults, even with the greatest of skill, are sometimes appalled—or certainly are not at their best—when called upon to operate upon and care for a tiny patient. Something more than diminutive instruments or scaled-down operative manipulations are necessary to do the job in a suitable manner.” To this day, surgical residents and other trainees often approach the pediatric surgical patient with the same mix of fear, trepidation, and anxiety. These same trainees often complete their pediatric surgical rotations with a profound respect for the resilience of young children to undergo complex operations and an appreciation for the precision required from their caregivers, both in the operating room and during the perioperative period. Over the decades, the specialty of pediatric surgery has evolved considerably in its care for the smallest of surgical patients, such that in utero surgery is now an option in an increasing number of circumstances. Similarly, our understanding of the pathophysiology of the diseases that pediatric surgeons face has increased to the point that some pediatric surgical diseases are now understood at the level of molecular or cellular signaling pathways. Pediatric surgery provides the opportunity to intervene in a wide array of diseases and to exert a long-lasting impact on the lives of children and their grateful parents. The scope of diseases encountered in the standard practice of pediatric surgery is immense, with patients...
Key Points

1. In infants with Bochdalek-type congenital diaphragmatic hernia, the severity of pulmonary hypoplasia and the resultant pulmonary hypertension are key determinants of survival. Barotrauma and hypoxia should be avoided.

2. During initial management of an infant with esophageal atresia and distal tracheoesophageal fistula, every effort should be made to avoid distending the gastrointestinal tract, especially when using mechanical ventilation. The patient should be evaluated for components of the VACTERL (vertebral, anorectal, cardiac, tracheoesophageal, renal, radial limb) anomalies. Timing and extent of surgery are dictated by the stability of the patient.

3. Although malrotation with midgut volvulus occurs most commonly within the first few weeks of life, it should always be considered in the differential diagnosis in a child with bilious emesis. Volvulus is a surgical emergency; therefore, in a critically ill child, prompt surgical intervention should not be delayed for any reason.

4. When evaluating a newborn infant for vomiting, it is critical to distinguish between proximal and distal causes of intestinal obstruction using both prenatal and postnatal history, physical examination, and abdominal radiographs.

5. Risk factors for necrotizing enterocolitis (NEC) include prematurity, formula feeding, bacterial infection, and intestinal ischemia. Critical to the management of infants with advanced (Bell stage III) or perforated NEC is timely and adequate source control of peritoneal contamination. Early sequelae of NEC include perforation, sepsis, and death. Later sequelae include short bowel syndrome and stricture.

6. In patients with intestinal obstruction secondary to Hirschsprung’s disease, a leveling ostomy or endorectal pull-through should be performed using ganglionated bowel, proximal to the transition zone between ganglionic and aganglionic intestine.

7. Prognosis of infants with biliary atresia is directly related to age at diagnosis and timing of portoenterostomy. Infants with advanced age at the time of diagnosis or infants who fail to demonstrate evidence of bile drainage after portoenterostomy usually require liver transplantation.

8. Infants with omphaloceles have greater associated morbidity and mortality than infants with gastroschisis due to a higher incidence of congenital anomalies and pulmonary hypoplasia. Gastroschisis can be associated with intestinal atresia, but not with other congenital anomalies. An intact omphalocele can be repaired electively, whereas gastroschisis requires urgent intervention to protect the exposed intestine.

9. Prognosis for children with Wilms’ tumor is defined by the stage of disease at the time of diagnosis and the histologic type (favorable vs. unfavorable). Preoperative chemotherapy is indicated for bilateral involvement, a solitary kidney, or tumor in the inferior vena cava above the hepatic veins. Gross tumor rupture during surgery automatically changes the stage to 3 (at a minimum).

10. Injury is the leading cause of death in children older than 1 year of age. Blunt mechanisms account for the majority of pediatric injuries. The central nervous system is the most commonly injured organ system and the leading cause of death in injured children.

PEDIATRIC SURGICAL THEMES: PITFALLS AND PEARLS

This chapter focuses on the unique considerations regarding the diagnosis and management of surgical diseases in the pediatric population. Many surgical trainees approach the surgical care of children with some degree of fear and trepidation. As any pediatric caregiver will attest to, the surgical management of infants and children requires delicate, careful, and professional interactions with their parents. The stress that the parents of sick children experience in the hospital setting can, at times, be overwhelming. It is due, in part, to the uncertainty regarding a particular prognosis, the feeling of helplessness that evolves when one is unable to care for one’s own child, and in certain cases, the guilt or remorse that one feels for not seeking medical care earlier, or for consenting to a particular procedure. Management of the sick child and his or her family requires not only a certain set of skills but also a unique knowledge base. This section is included to summarize some important general principles in accomplishing this task.

1. Children are not little adults, but they are little people. In practical terms, this often-heard refrain implies that children have unique fluid, electrolyte, and medication needs. Thus, the dosage of medications and the administration of IV fluids should at all times be based on their weight. The corollary of this point is that infants and young children are extremely sensitive to perturbations in their normal physiology and may be easily tipped into fluid overload or dehydration.

2. Sick children whisper before they shout. Children with surgical diseases can deteriorate very quickly. But before they deteriorate, they often manifest subtle physical findings. These findings—referred to as “whispers”—may include signs such as tachycardia, bradycardia, hypothermia, fever, recurrent emesis, or feeding intolerance. Meticulous attention to these subtle findings may unmask the development of potentially serious, life-threatening physiological disturbances.

3. Always listen to the mother and the father. Surgical diseases in children can be very difficult to diagnose because children are often minimally communicative, and information that they communicate may be confusing, conflicting, or both. In all cases, it is wise to listen to the child’s parents, who have closely observed their child and know him or her best. Most importantly, the child’s parents know with certainty...
whether or not the child is sick or not, despite not always knowing the precise diagnosis.

4. Pediatric tissue must be handled delicately and with profound respect.


6. Pay particular attention to the postoperative pediatric patient whose pain cannot be soothed by the administration of standard amounts of analgesic agents. Ask yourself whether a significant yet unrecognized postoperative complication exists.

**GENERAL CONSIDERATIONS**

**Fluid and Electrolyte Balance**

In managing the pediatric surgical patient, an understanding of fluid and electrolyte balance is critical as the margin between dehydration and fluid overload is small. This is particularly true in infants, who have little reserve at baseline and even less when ill. Failure to pay meticulous attention to their hydration status can result in significant fluid overload or dehydration. Several surgical diagnoses such as gastrochisis or short-gut syndrome are characterized by a predisposition to fluid loss. Others require judicious restoration of intravascular volume in order to prevent cardiac failure as is the case in patients with congenital diaphragmatic hernia and associated pulmonary hypertension.

The infant’s physiologic day is approximately eight hours in duration. Accordingly, careful assessment of the individual patient’s fluid balance, including fluid intake and output for the previous eight hours, is essential to prevent dehydration or fluid overload. Clinical signs of dehydration include tachycardia, decreased urine output, reduced skin turgor, depressed fontanelle, absent tears, lethargy, and poor feeding. Fluid overload is often manifested by the onset of a new oxygen requirement, respiratory distress, tachypnea, and tachycardia. The physical assessment of the fluid status of each child must include a complete head-to-toe evaluation, with emphasis on determining whether perturbations in normal physiology are present.

At 12 weeks’ gestation, the total body water of a fetus is approximately 94 cc/kg. By the time the fetus reaches full term, the total body water has decreased to approximately 80 cc/kg. Total body water drops an additional 5% within the first week of life, and by 1 year of life, total body water approaches adult levels, around 60 to 65 cc/kg. Parallel to the drop in total body water is the reduction in extracellular fluid. These changes are accelerated in the preterm infant who may face additional fluid losses due to coexisting congenital anomalies or surgery. Normal daily maintenance fluids for most children can be estimated using the following formula:

\[
100 \text{ mL/kg for the first } 10 \text{ kg, plus } 50 \text{ mL/kg for each additional kilogram of body weight thereafter.}
\]

Because IV (I.V.) fluid orders are written as milliliters per hour, this can be conveniently converted to:

\[
4 \text{ mL/kg/h up to } 10 \text{ kg, add } 2 \text{ mL/kg/h for each additional kilogram body weight thereafter.}
\]

For example, a 26-kg child has an estimated maintenance fluid requirement of \((10 \times 4) + (10 \times 2) + (6 \times 1) = 66 \text{ mL/h in the absence of massive fluid losses or shock. A newborn infant with gastrochisis will manifest significant evaporative losses from the exposed bowel such that fluid requirements can be on the order of 150 to 180 cc/kg/day.}

Precise management of a neonate’s fluid status requires an understanding of changes in the glomerular filtration rate (GFR) and tubular function of the kidney. The term newborn’s GFR is approximately 21 mL/min/1.73 m² compared to 70 mL/min/1.73 m² in an adult. Within the first 2 weeks of life GFR increases to approximately 60, and by 2 years of age it is essentially at adult levels. The capacity to concentrate urine is very limited in preterm and term infants. In comparison to an adult who can concentrate urine to 1200 mosm/kg, infants can concentrate urine at best to 600 mosm/kg. While infants are capable of secreting antidiuretic hormone, ADH, the aquaporin water channel-mediated osmotic water permeability of the infant’s collecting tubules is severely limited compared to that of adults, leading to an insensitivity to ADH.

Sodium requirements range from 2 mEq/kg per day in term infants up to 5 mEq/kg per day in critically ill preterm infants as a consequence of salt wasting. Potassium requirements are on the order of 1 to 2 mEq/kg per day. Calcium and magnesium supplementation of IV fluids is essential to prevent laryngospasm, dysrhythmias, and tetany.

**Acid-Base Equilibrium**

Acute metabolic acidosis usually implies inadequate tissue perfusion and is a serious disorder in children. Potentially life-threatening causes that are specific for the pediatric population must be sought; they include intestinal ischemia from necrotizing enterocolitis (in the neonate), midgut volvulus, or incarcerated hernia. Other causes include chronic bicarbonate loss from the gastrointestinal tract or acid accumulation as in chronic renal failure. Respiratory acidosis implies hypoventilation, the cause of which should be apparent. Treatment of acute metabolic acidosis should be aimed at restoring tissue perfusion by addressing the underlying abnormality first. For severe metabolic acidemia where the serum pH is less than 7.25, sodium bicarbonate should be administered using the following guidelines: base deficit × weight in kilograms × 0.5 (in newborns). The last factor in the equation should be 0.4 for smaller children and 0.3 for older children. The dose should be diluted to a concentration of 0.5 mEq/mL because full-strength sodium bicarbonate is hyperosmolar. One-half the corrective dose is given, and the serum pH is measured again. During cardiopulmonary resuscitation (CPR), one-half the corrective dose can be given as an intravenous bolus and the other half given slowly intravenously.

Respiratory alkalosis is usually caused by hyperventilation, which is readily correctable. Metabolic alkalosis most commonly implies gastric acid loss, as in the child with pyloric stenosis, or aggressive diuretic therapy. In the child with gastric fluid loss, IV fluids of 5% dextrose, 0.5% normal saline, and 20 mEq KCl/L usually correct the alkalosis.

**Blood Volume and Blood Replacement**

Criteria for blood transfusion in infants and children remain poorly defined. The decision to transfuse a critically ill pediatric patient may depend on a number of clinical features that include the patient’s age, primary diagnosis, the presence of ongoing bleeding, coagulopathy, hypoxia, hemodynamic compromise, lactic acidosis, cyanotic heart disease, and overall severity of illness. A recent survey of transfusion practices among pediatric intensivists showed that the baseline hemoglobin levels that would prompt them to recommend RBC transfusion ranged from 7 to 13 g/dL. Patients with cyanotic heart disease are often transfused to
higher hemoglobin values, although the threshold for transfusion in this population remains to be defined. In general terms, there is a trend towards an avoidance of the use of RBC products whenever possible as current studies suggest that lower hemoglobin concentrations are well tolerated by many groups of patients and that administration of RBCs may have unintended negative consequences, including perhaps an increase in predisposition to the development of necrotizing enterocolitis, although this finding is controversial. In addition, there is increasing evidence that PRBC transfusion may have adverse effects on the host immune in both children and adults. These effects are poorly understood but may include effects due to RBC storage and due to factors that are particular to the individual RBC donor. The TRIPICU randomized controlled trial by Lacroix et al in 2007, which was performed in stable critically ill children, determined that a restrictive Hb transfusion trigger (70 g/L) was as safe as a liberal Hb trigger (95 g/L) and was associated with reduced blood use. It remains uncertain whether this can be extrapolated to unstable patients. Expert opinion now generally favors an Hb transfusion trigger of 70 g/L in stable critically ill children, which is the same as the recommendation for adult patients (see Chapter 7).

A higher threshold should be considered if the child has symptomatic anemia or impaired cardiorespiratory function. A useful guideline for estimating blood volume for the newborn infant is approximately 80 mL/kg of body weight. When packed red blood cells are required, the transfusion requirement is usually administered in 10 mL/kg increments, which is roughly equivalent to a 500-mL transfusion for a 70-kg adult. The following formula may be used to determine the volume (ml) of PRBC to be transfused:

\[(\text{Target hematocrit} - \text{Current Hematocrit}) \times \text{weight (kg)} \times \frac{80}{65}\] (65 represents the estimated hematocrit of a unit of PRBC)

As a general rule, blood is recommended for replacement of volume loss if the child’s perfusion is inadequate despite administration of 2 to 3 boluses of 20 mL/kg of isotonic crystalloid. Consideration should be given for the administration of 10 mL/kg of packed red blood cells as soon as possible. Type O blood can be administered without a cross-match and is relatively safe; type-specific blood can be obtained quite quickly; however, unlike fully cross-matched blood, incompatibilities other than ABO and Rh may exist.

In the child, coagulation deficiencies may rapidly assume clinical significance after extensive blood transfusion. It is advisable to have fresh frozen plasma and platelets available if more than 30 mL/kg have been transfused. Plasma is given in a dose of 10 to 20 mL/kg, and platelets are given in a dose of 1 unit/5 kg. Each unit of platelets consists of 40 to 60 mL of fluid (plasma plus platelets). Following transfusion of PRBCs to neonates with tenuous fluid balance, a single dose of a diuretic (such as furosemide 1 mg/kg) may help to facilitate excretion of the extra fluid load. Many clinicians prefer to administer fresh products to minimize the deleterious effects of red cell storage.

In pediatric patients who have lost greater than 30 mL/kg with ongoing bleeding, consideration should be given to initiation of a massive transfusion protocol. Such a protocol involves transfusion, based on weight, of 1:1:1 transfusion of RBCs, plasma, and platelets.

Parenteral Alimentation and Nutrition
The nutritional requirements of the surgical neonate must be met in order for the child to grow and to heal surgical wounds. If inadequate protein and carbohydrate calories are given, the child may not only fail to recover from surgery but may also exhibit growth failure and impaired development of the central nervous system. In general terms, the adequacy of growth must be assessed frequently by determining both total body weight as well as head circumference. Neonates that are particularly predisposed to protein-calorie malnutrition include those with gastroschisis, intestinal atresia, or intestinal insufficiency from other causes, such as necrotizing enterocolitis. The protein and caloric requirements for the surgical neonate are shown in Table 39-1.

Nutrition can be provided via either the enteral or parenteral routes. Whenever possible, the enteral route is preferred because it not only promotes the growth and function of the gastrointestinal system, it also ensures that the infant learns how to feed. There are various enteral feeding preparations available; these are outlined in Table 39-2. The choice of formula is based upon the individual clinical state of the child. Pediatric surgeons are often faced with situations where oral feeding is not possible. This problem can be seen in the extremely premature infant who has not yet developed the feeding skills, or in the infant with concomitant craniofacial anomalies that impair sucking, for example. In these instances, enteral feeds can be administered either a nasojejunal or a gastrostomy tube.

When the gastrointestinal tract cannot be used because of mechanical, ischemic, inflammatory, or functional disorders, parenteral alimentation must be given. Prolonged parenteral nutrition is delivered via a central venous catheter. Peripheral IV alimentation can be given, utilizing less concentrated but greater volumes of solutions. Long-term parenteral nutrition should include supplemental copper, zinc, and iron to prevent the development of trace metal deficiencies. A major complication of long-term total parenteral nutrition (TPN) is the development of parenteral nutrition–associated cholestasis, which can eventually progress to liver failure. To prevent this major complication, concomitant enteral feedings should be instituted, and the gastrointestinal tract should be used as soon as possible. When proximal stomas are in place, gastrointestinal continuity should be restored as soon as possible. Where intestinal insufficiency is associated with dilation of the small intestine, tapering or intestinal lengthening procedures may be beneficial.

### Table 39-1

**Nutritional requirements for the pediatric surgical patient**

<table>
<thead>
<tr>
<th>AGE</th>
<th>CALORIES (kcal/kg/d)</th>
<th>PROTEIN (gram/kg/d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–6 months</td>
<td>100–120</td>
<td>2</td>
</tr>
<tr>
<td>6 months–1 year</td>
<td>100</td>
<td>1.5</td>
</tr>
<tr>
<td>1–3 years</td>
<td>100</td>
<td>1.2</td>
</tr>
<tr>
<td>4–6 years</td>
<td>90</td>
<td>1</td>
</tr>
<tr>
<td>7–10 years</td>
<td>70</td>
<td>1</td>
</tr>
<tr>
<td>11–14 years</td>
<td>55</td>
<td>1</td>
</tr>
<tr>
<td>15–18 years</td>
<td>45</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 39-2
Formulas for pediatric surgical neonates

<table>
<thead>
<tr>
<th>FORMULA</th>
<th>kcal/mL</th>
<th>PROTEIN (g/mL)</th>
<th>FAT (g/mL)</th>
<th>CARBOHYDRATE (g/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human milk</td>
<td>0.67</td>
<td>0.011</td>
<td>0.04</td>
<td>0.07</td>
</tr>
<tr>
<td>Milk-based formula</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enfamil 20</td>
<td>0.67</td>
<td>0.015</td>
<td>0.038</td>
<td>0.069</td>
</tr>
<tr>
<td>Similac 20</td>
<td>0.67</td>
<td>0.015</td>
<td>0.036</td>
<td>0.072</td>
</tr>
<tr>
<td>Soy-based formula</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prosopee</td>
<td>0.67</td>
<td>0.02</td>
<td>0.036</td>
<td>0.07</td>
</tr>
<tr>
<td>Isomil</td>
<td>0.67</td>
<td>0.018</td>
<td>0.037</td>
<td>0.068</td>
</tr>
<tr>
<td>Special formula</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregestimil</td>
<td>0.67</td>
<td>0.019</td>
<td>0.028</td>
<td>0.091</td>
</tr>
<tr>
<td>Alimentum</td>
<td>0.67</td>
<td>0.019</td>
<td>0.038</td>
<td>0.068</td>
</tr>
<tr>
<td>Preterm</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enfamil Premature</td>
<td>.80</td>
<td>0.024</td>
<td>0.041</td>
<td>0.089</td>
</tr>
</tbody>
</table>

Other strategies to minimize the development of TPN-related liver disease include meticulous catheter care to avoid infection, which increases cholestatic symptoms, aggressive treatment of any infection, and early cycling of parenteral nutrition in older children who can tolerate not receiving continuous dextrose solution for a limited period. Evidence suggests that cholestasis eventually resolves in most cases after parenteral nutrition is discontinued, as measured by levels of total bilirubin. Preliminary evidence suggests that substituting omega-3 fish oil lipid emulsion in parenteral nutrition for the standard soybean-based emulsions may prevent the development of TPN-related cholestasis and reverse the effects of established liver disease. A phase 2 trial to determine whether parenteral nutrition–associated liver disease can be reversed or its progression halted by using a parenteral fat emulsion prepared from fish oil as measured by normalization of serum levels of hepatic enzymes and bilirubin is ongoing (ClinicalTrials.gov, identifier NCT00826020).

**Venous Access**

Obtaining reliable vascular access in an infant or child is an important task that often becomes the responsibility of the pediatric surgeon. The goal should always be to place the catheter in the least invasive, least risky, and least painful manner, and in a location that is most accessible and allows for use of the catheter without complications for as long as it is needed. In infants, central venous access may be established using a cutdown approach, either in the antecubital fossa, external jugular vein, facial vein, or proximal saphenous vein. If the internal jugular vein is used, care is taken to prevent venous occlusion. In infants over 3 kg and in older children, percutaneous access of the subclavian, internal jugular, or femoral veins is possible in most cases, and central access is achieved using the Seldinger technique. The use of ultrasound (US) is considered standard of care for placement of central lines in this population for the internal jugular vein and femoral veins, and it significantly improves the safety of the insertion procedure. The catheters are tunneled to an exit site separate from the venotomy site. Where available, PICC lines (peripherally inserted central catheters) may be placed, typically via the antecubital fossa. Regardless of whether the catheter is placed by a cutdown approach or percutaneously, a chest X-ray to confirm central location of the catheter tip and to exclude the presence of a pneumothorax or hemothorax is mandatory. When discussing the placement of central venous catheters with parents, it is important to note that the complication rate for central venous lines in children can be high. The incidence of catheter-related sepsis or infection remains a problem, yet should be less than 1% with meticulous attention to catheter insertion care and exit site management. Superior or inferior vena cava occlusion is a significant risk after the placement of multiple lines, particularly in the smallest premature patients.

**Thermoregulation**

Careful regulation of the ambient environment of infants and children is crucial as these patients are extremely thermolabile. Premature infants are particularly susceptible to changes in environmental temperature. Because they are unable to shiver and lack stores of fat, their potential for thermogenesis is impaired. The innate inability to regulate temperature is compounded by the administration of anesthetic and paralyzing agents. Since these patients lack adaptive mechanisms to cope with the environment, the environment must be carefully regulated. Attention to heat conservation during transport of the infant to and from the operating room is essential. Transport systems incorporating heating units are necessary for premature infants. In the operating room, the infant is kept warm by the use of overhead heating lamps, a heating blanket, warming of inspired gases, and coverage of the extremities and head with occlusive materials. During abdominal surgery, extreme care is taken to avoid wet and cold drapes. All fluids used to irrigate the chest or abdomen must be warmed to body temperature. Laparoscopic approaches for abdominal operations may result in more stable thermoregulation due to decreased heat loss from the smaller wound size. Constant monitoring of the child’s temperature is critical in a lengthy procedure, and the surgeon should continuously communicate with the anesthesiologist regarding the temperature of the patient. The development of hypothermia in infants and children can result in cardiac arrhythmias or coagulopathy. These potentially life-threatening complications can be avoided by careful attention to thermoregulation.
Pain Control

All children including neonates experience pain; the careful recognition and management of pediatric pain represents an important component of the perioperative management of all pediatric surgical patients. There is a range of pain management options that can improve the child’s well-being, as well as the parents’ sense of comfort. Given that morphine and fentanyl have an acceptable safety margin, they should be administered to neonates and children when indicated, bearing in mind that withholding analgesia poses a significant risk, as does administration of excessive analgesic agents. A recent randomized trial of neonates on ventilators showed that the use of a morphine infusion decreased the incidence of intraventricular hemorrhage by 50%. Additional analgesic modalities include the use of topical anesthetic ointment (EMLA cream) and the use of regional anesthesia, such as caudal blocks for hernias and epidural or incisional catheter infusions (On-Q) for large abdominal or thoracic incisions. In surgical neonates that have been administered large concentrations of narcotics over a prolonged period, transient physical dependence should not only be expected but also anticipated. When narcotics are discontinued, symptoms of narcotic withdrawal may develop, including irritability, restlessness, and episodes of hypertension and tachycardia. Early recognition of these signs is essential, as is timely treatment using nalaxone and other agents. It is important to administer pain control in concert with a well-qualified and collaborative pediatric pain-management team, which typically includes anesthesiologists with expertise in pain management, as well as advance practice nurses who can respond rapidly when the pain control is inadequate or excessive. By ensuring that the pediatric surgical patient has adequate analgesia, the surgeon ensures that the patient receives the most humane and thorough treatment and provides important reassurance to all other members of the healthcare team and to the family that pain control is a very high priority.

NECK MASSES

The management of neck masses in children is determined by their location and the length of time that they have been present. Neck lesions are found either in the midline or lateral compartments. Midline masses include thyroglossal duct remnants, thyroid masses, thymic cysts, or dermoid cysts. Lateral lesions include branchial cleft remnants, cystic hygromas, vascular malformations, salivary gland tumors, torticolli, and lipoblastoma (a rare benign mesenchymal tumor of embryonal fat occurring in infants and young children). Enlarged lymph nodes and rare malignancies such as rhabdomyosarcoma can occur either in the midline or laterally.

Lymphadenopathy

The most common cause of a neck mass in a child is an enlarged lymph node, which typically can be found laterally or in the midline. The patient is usually referred to the pediatric surgeon for evaluation after the mass has been present for several weeks. A detailed history and physical examination often helps determine the likely etiology of the lymph node and the need for excisional biopsy. Enlarged tender lymph nodes are usually the result of a bacterial infection (Staphylococcus or Streptococcus). Treatment of the primary cause (e.g., otitis media or pharyngitis) with antibiotics often is all that is necessary. However, when the involved nodes become fluctuant, incision and drainage are indicated. In many North American institutions, there has been an increasing prevalence of methicillin-resistant Staphylococcus aureus infection of the skin and soft tissues, leading to increased staphylococcal lymphadenitis in children. More chronic forms of lymphadenitis, including infections with atypical mycobacteria, as well as cat-scratch fever, are diagnosed based on serologic findings or excisional biopsy. The lymphadenopathy associated with infectious mononucleosis can be diagnosed based on serology. When the neck nodes are firm, fixed, and others are also present in the axillae or groin, or the history suggests lymphoma, excisional biopsy is indicated. In these cases, it is essential to obtain a chest radiograph to look for the presence of a mediastinal mass. Significant mediastinal load portends cardiopulmonary collapse due to loss of venous return and compression of the tracheobronchial tree with general anesthesia.

Thyroglossal Duct Remnants

Pathology and Clinical Manifestations. The thyroid gland buds off the foregut diverticulum at the base of the tongue in the region of the future foramen cecum at 3 weeks of embryonic life. As the fetal neck develops, the thyroid tissue becomes more anterior and caudal until it rests in its normal position. The “descent” of the thyroid is intimately connected with the development of the hyoid bone. Residual thyroid tissue left behind during the migration may persist and subsequently present in the midline of the neck as a thyroglossal duct cyst. The mass is most commonly appreciated in the 2- to 4-year-old child when the baby fat disappears and irregularities in the neck become more readily apparent. Usually the cyst is encountered in the midline at or below the level of the hyoid bone and moves up and down with swallowing or with protrusion of the tongue. Occasionally it presents as an intrathyroidal mass. Most thyroglossal duct cysts are asymptomatic. If the duct retains its connection with the pharynx, infection may occur, and the resulting abscess will necessitate incision and drainage, occasionally resulting in a salivary fistula. Submental lymphadenopathy and midline dermoid cysts can be confused with a thyroglossal duct cyst. Rarely, midline ectopic thyroid tissue masquerades as a thyroglossal duct cyst and may represent the patient’s only thyroid tissue. Therefore, if there is any question regarding the diagnosis or if the thyroid gland cannot be palpated in its normal anatomic position, it is advisable to obtain a nuclear scan to confirm the presence of a normal thyroid gland. Although rarely the case in children, in adults the thyroglossal duct may contain thyroid tissue that can undergo malignant degeneration. The presence of malignancy in a thyroglossal cyst should be suspected when the cyst grows rapidly or when US demonstrates a complex anechoic pattern or the presence of calcification.

Treatment. If the thyroglossal duct cyst presents with an abscess, treatment should first consist of drainage and antibiotics. Following resolution of the inflammation, resection of the cyst in continuity with the central portion of the hyoid bone and the tract connecting to the pharynx in addition to ligation at the foramen cecum (the Sistrunk operation), is curative in over 90% of patients. Lesser operations result in unacceptably high recurrence rates, and recurrence is more frequent following infection. According to a recent review, factors predictive of recurrence included more than two infections prior to surgery, age under 2 years, and inadequate initial operation.
Paired branchial clefts and arches develop early in the fourth gestational week. The first cleft and the first, second, third, and fourth pouches give rise to adult organs. The embryologic communication between the pharynx and the external surface may persist as a fistula. A fistula is seen most commonly with the second branchial cleft, which normally disappears, and extends from the anterior border of the sternocleidomastoid muscle superiorly, inward through the bifurcation of the carotid artery, and enters the posterolateral pharynx just below the tonsillar fossa. In contrast, a third branchial cleft fistula passes posterior to the carotid bifurcation. The branchial cleft remnants may contain small pieces of cartilage and cysts, but internal fistulas are rare. A second branchial cleft sinus is suspected when clear fluid is noted draining from the external opening of the tract at the anterior border of the lower third of the sternomastoid muscle. Rarely, branchial cleft anomalies occur in association with biliary atresia and congenital cardiac anomalies, an association that is referred to as Goldenhar’s complex.

**Treatment.** Complete excision of the cyst and sinus tract is necessary for cure. Dissection of the sinus tract is facilitated with passage of a fine lacrimal duct probe through the external opening into the tract and utilizing it as a guide for dissection. Injection of a small amount of methylene blue dye into the tract also may be useful. A series of two or sometimes three small transverse incisions in a “stepladder” fashion is preferred to a long oblique incision in the neck, which is cosmetically undesirable. Branchial cleft cysts can present as abscesses. In these cases, initial treatment includes incision and drainage with a course of antibiotics to cover *Staphylococcus* and *Streptococcus* species, followed by excision of the cyst after the infection resolves.

**Lymphatic Malformation**

**Etiology and Pathology.** Lymphatic malformation (cystic hygroma or lymphangioma) occurs as a result of sequestration or obstruction of developing lymph vessels in approximately 1 in 12,000 births. Although the lesion can occur anywhere, the most common sites are in the posterior triangle of the neck, axilla, groin, and mediastinum. The cysts are lined by endothelium and filled with lymph. Occasionally unilocular cysts occur, but more often there are multiple cysts “infiltrating” the surrounding structures and distorting the local anatomy. A particularly troublesome variant of lymphatic malformation is that which involves the tongue, floor of the mouth, and structures deep in the neck. Adjacent connective tissue may show extensive lymphocytic infiltration. The mass may be apparent at birth or may appear and enlarge rapidly in the early weeks or months of life as lymph accumulates; most present by age 2 years (Fig. 39-1A). Extension of the lesion into the axilla or mediastinum occurs about 10% of the time and can be demonstrated preoperatively by chest X-ray, US, or computed tomographic (CT) scan, although magnetic resonance imaging (MRI) is preferable. Occasionally lymphatic malformations contain nests of vascular tissue. These poorly supported vessels may bleed and produce rapid enlargement and discoloration of the lesion. Infection within the lymphatic malformations, usually caused by *Streptococcus* or *Staphylococcus*, may occur. In the neck, this can cause rapid enlargement, which may result in airway compromise. Rarely, it may be necessary to carry out percutaneous aspiration of a cyst to relieve respiratory distress.

The diagnosis of lymphatic malformation by prenatal US, before 30 weeks’ gestation, has detected a “hidden mortality” as well as a high incidence of associated anomalies, including abnormal karyotypes and hydrops fetalis. Occasionally, very large lesions can cause obstruction of the fetal airway. Such obstruction can result in the development of polyhydramnios by impairing the ability of the fetus to swallow amniotic fluid. In these circumstances, the airway is usually markedly distorted, which can result in immediate airway obstruction unless the airway is secured at the time of delivery. Orotracheal intubation or emergency tracheostomy while the infant remains attached to the placenta, the so-called EXIT procedure (ex utero intrapartum technique) may be necessary to secure the airway.

**Treatment.** The modern management of most lymphatic malformations includes image-guided sclerotherapy as first-line therapy, which often involves multiple injections. Cyst excision may be used in cases where injection is inadequate.

---

**Figure 39-1.** A. Left cervical cystic hygroma in a 2-day old baby. B. Intraoperative photograph showing a vessel loop around the spinal accessory nerve.
Total removal of all gross disease is often not possible because of the extent of the lymphatic malformation and its proximity to, and intimate relationship with, adjacent nerves, muscles, and blood vessels (Fig. 39-1B). Radical ablative surgery is not indicated for these lesions, which are always benign. Conservative excision and unroofing of remaining cysts is advised, with repeated partial excision of residual cysts and sclerotherapy if necessary, preserving all adjacent crucial structures. In cases in which surgical excision is performed, closed-suction drainage is recommended. Nevertheless, fluid may accumulate beneath the surgically created flaps in the area from which the lymphatic malformation was excised, requiring multiple needle aspirations. A combined sclerotherapy/resectional approach is particularly useful for masses that extend to the base of the tongue or the floor of the mouth.

**Torticollis**

The presence of a lateral neck mass in infancy in association with rotation of the head towards the opposite side of the mass indicates the presence of congenital torticollis. This lesion results from fibrosis of the sternocleidomastoid muscle. The mass may be palpated in the affected muscle in approximately two-thirds of cases, or it may be diagnosed by US. Histologically, the lesion is characterized by the deposition of collagen and fibroblasts around atrophied muscle cells. In the vast majority of cases, physical therapy based on passive stretching of the affected muscle is of benefit. Rarely, surgical transection of the sternocleidomastoid may be indicated.

**RESPIRATORY SYSTEM**

### Congenital Diaphragmatic Hernia (Bochdalek)

**Pathology.** The septum transversum extends to divide the pleural and coelomic cavities during fetal development. This precursor of the diaphragm normally completes separation of these two cavities at the posterolateral aspects of this mesenchymally derived structure. The most common variant of a congenital diaphragmatic hernia is a posterolateral defect, also known as a Bochdalek hernia. Diaphragmatic defects allow abdominal viscera to fill the chest cavity. The abdominal cavity is small and underdeveloped and remains seaploid after birth. Both lungs are hypoplastic, with decreased bronchial and pulmonary artery branching. Lung weight, lung volume, and DNA content are also decreased, and these findings are more striking on the ipsilateral side. This anomaly is encountered more commonly on the left (80–90%). Linkage analyses have recently implicated genetic mutations in syndromic variants of congenital diaphragmatic hernias. In many instances, there is a surfactant deficiency, which compounds the degree of respiratory insufficiency. Amniocentesis with karyotype may identify chromosomal defects, especially trisomy 18 and 21. Associated anomalies, once thought to be uncommon, were identified in 65 of 166 patients in one study, predominately of the heart, followed by abdominal wall defects, chromosomal changes, and other defects.

Prenatal ultrasonography is successful in making the diagnosis of congenital diaphragmatic hernia (CDH) as early as 15 weeks’ gestation, and early antenatal diagnosis is associated with worse outcomes. US findings include herniated abdominal viscera in the chest that may also look like a mass or lung anomaly, changes in liver position, and mediastinal shift away from the herniated viscera (Fig. 39-2). Accurate prenatal prediction of outcome for fetuses who have CDH remains a challenge. One index of severity for patients with left CDH is the lung-to-head ratio (LHR), which is the product of the length and the width of the right lung at the level of the cardiac atria divided by the head circumference (all measurements in millimeters). An LHR value of less than 1.0 is associated with a very poor prognosis, whereas an LHR greater than 1.4 predicts a more favorable outcome. The utility of the LHR in predicting outcome in patients with CDH has recently been questioned because of the tremendous interobserver variability in calculating this ratio for a particular patient, as well as the lack of reliable measures to determine postnatal disease severity. Because the LHR is not gestational age independent, Jani and colleagues proposed the introduction of a new measurement: the observed (o/e) LHR, to correct for gestational age. The observed LHR may be expressed as a percentage of the expected mean for gestational age of the observed/expected lung-to-head ratio (o/e LHR), which is considered extreme if <15%, severe at 15% to 25%, moderate at 26% to 35%, and mild at 36% to 45%. The most reliable prenatal predictor of postnatal survival is absence of liver herniation, where in 710 fetuses, there was significantly higher survival rate in fetuses without herniation (74% without herniation vs. 45% with herniation).

Following delivery, the diagnosis of CDH is made by CXR (Fig. 39-3). The differential diagnosis includes bronchopulmonary foregut malformations, in which the intrathoracic loops of bowel may be confused for lung or foregut pathology. The vast majority of infants with CDH develop immediate respiratory distress, which is due to the combined effects of three factors. First, the air-filled bowel in the chest compresses the mobile mediastinum, which shifts to the opposite side of the chest, compromising air exchange in the contralateral lung. Second, pulmonary hypertension develops. This phenomenon results in persistent fetal circulation with resultant decreased pulmonary perfusion and impaired gas exchange. Finally, the lung on the affected side is often hypoplastic, such that it is essentially nonfunctional. Varying degrees of pulmonary hypoplasia on the opposite side may compound these effects. The second and third factors are thought to be the most important. Neonates with CDH are usually in respiratory distress requiring
ventilation and intensive care, and the overall mortality in most series is around 50%.

**Treatment.** CDH care has been improved through effective use of improved methods of ventilation and timely cannulation for extracorporeal membrane oxygenation (ECMO). Many infants are symptomatic at birth due to hypoxia, hypercarbia, and metabolic acidosis. Prompt cardiorespiratory stabilization is mandatory. It is noteworthy that the first 24 to 48 hours after birth are often characterized by a period of relative stability with high levels of PaO₂ and relatively good perfusion. This has been termed the “ honeymoon period” and is often followed by progressive cardiorespiratory deterioration. In the past, correction of the hernia was believed to be a surgical emergency, and patients underwent surgery shortly after birth. It is now accepted that the presence of persistent pulmonary hypertension that results in right-to-left shunting across the open foramen ovale or the ductus arteriosus, and the degree of pulmonary hypoplasia, are the leading causes of cardiorespiratory insufficiency. Current management is directed toward managing the pulmonary hypertension, and minimizing barotrauma while optimizing oxygen delivery. To achieve this goal, infants are placed on mechanical ventilation using relatively low or “gentle” settings that prevent overinflation of the noninvolved lung. Levels of PaCO₂ in the range of 50 to 60 mmHg or higher are accepted as long as the pH remains ≥7.25. If these objectives cannot be achieved using conventional ventilation, high frequency oscillatory ventilation (HFOV) may be employed to avoid the injurious effects of conventional tidal volume ventilation. Echocardiography will assess the degree of pulmonary hypertension and identify the presence of any coexisting cardiac anomaly. ICU goals include minimal sedation, meticulous attention to endotracheal tube secretions, and gradual changes to ventilator settings to avoid inducing pulmonary hypertension via hypoxia. To minimize the degree of pulmonary hypertension, inhaled nitric oxide may be administered, and in some patients, this improves pulmonary perfusion. Nitric oxide is administered into the ventilation circuit and is used in concentrations up to 40 parts per million. Correction of acidosis using bicarbonate solution may minimize the degree of pulmonary hypertension. As the degree of pulmonary hypertension becomes hemodynamically significant, right-sided heart failure develops, and systemic perfusion is impaired. Administration of excess IV fluid will compound the degree of cardiac failure and lead to marked peripheral edema. Inotropic support using epinephrine, dopamine, and milrinone alone or in combination may be useful in optimizing cardiac contractility and maintaining mean arterial pressure.

Infants with CDH who remain severely hypoxic despite maximal ventilatory care may be candidates for treatment of their respiratory failure ECMO, with access via venovenous (VV) or venoarterial (VA) routes. VV bypass is established with a single cannula through the right internal jugular vein, with blood removed from and infused into the right atrium by separate ports. VA bypass provides additional cardiac support, whereas VV bypass requires a well-functioning heart and relies on the lungs for some oxygenation as well. In VA ECMO, the right atrium is cannulated by means of the internal jugular vein and the aortic arch through the right common carotid artery. As much of the cardiac output is directed through the membrane oxygenator as is necessary to provide oxygenated blood to the infant and remove carbon dioxide. The infant is maintained on bypass until the pulmonary hypertension is resolved and lung function, as measured by compliance and the ability to oxygenate and ventilate, is improved. This is usually seen within 7 to 10 days, but in some infants, it may take up several weeks to occur. Complications associated with ECMO increase after 14 days and include cannula malposition, bleeding in multiple locations, and infection. The use of ECMO is associated with significant risk. Because patients require systemic anticoagulation, bleeding complications are the most significant. They may occur intracranially or at the site of cannula insertion, and they can be life-threatening. Systemic sepsis is a significant problem and may necessitate decannulation. Criteria for placing infants on ECMO include the presence of normal cardiac anatomy by echocardiography, the absence of fatal chromosome anomalies, and the expectation that the infant would die without ECMO. Traditionally, a threshold of weight greater than 2 kg and gestational age greater than 34 weeks has been applied, although success has been achieved at weights as low as 1.8 kg. Upon decannulation, some centers repair the carotid artery. In instances in which the child is cannulated for a brief period (5 days or less) this may be feasible. A recent study failed to show any benefit from repairing the carotid artery, although this finding remains to be studied further.

A strategy that does not involve the use of ECMO but instead emphasizes the use of permissive hypercapnia and the avoidance of barotrauma may provide equal overall outcome in patients with CDH. This likely reflects the fact that mortality is related to the degree of pulmonary hypoplasia and the presence of congenital anomalies, neither of which are correctable by ECMO.
The timing of diaphragmatic hernia repair still varies from center to center, particularly when the infant is on ECMO. In patients that are not on ECMO, repair should be performed once the hemodynamic status has been optimized. In neonates that are on ECMO, some surgeons perform early repair on bypass; others wait until the infant’s lungs are improved and the pulmonary hypertension has subsided and then repair the diaphragm and discontinue bypass within hours of surgery. Still others repair the diaphragm only after the infant is off bypass. Operative repair of the diaphragmatic hernia may be accomplished either by an abdominal or transthoracic approach and can be performed either via open or minimally invasive techniques. Through a subcostal incision the abdominal viscera are withdrawn from the chest, exposing the defect in the diaphragm. Care must be taken when reducing the spleen and liver, as bleeding from these structures can be fatal. The anterior margin is often apparent, while the posterior muscular rim is attenuated. If the infant is heparinized on bypass, minimal dissection of the muscular margins is performed. Electrocautery is used liberally to minimize postoperative bleeding. Most infants who require ECMO support prior to hernia repair have large defects, often lacking the medial and posterior margins. About three-fourths of infants repaired on bypass require prosthetic material to patch the defect, suturing it to the diaphragmatic remnant or around ribs or costal cartilages for the large defects. If there is adequate muscle for closure, a single layer of nonabsorbable horizontal mattress suture, pledgeted or not, closes the defect. Just before the repair is complete, a chest tube may be positioned in the thoracic cavity but is not mandatory. Patients repaired on ECMO are at risk for developing a hemotherax, which can significantly impair ventilation. Anatomic closure of the abdominal wall may be impossible after reduction of the viscera. Occasionally, a prosthetic patch or acellular material may be sutured to the fascia to facilitate closure. The patch can be removed at a later time, and the ventral hernia can be closed at that time or subsequently. In patients who are deemed to be candidates for a minimally invasive approach (stable patients, >2 kg, no pulmonary hypertension), a thorascoscopic repair may be safely performed although concerns have been raised about possible effects of the longer operative time for thorascoscopic repair and higher recurrence rates. If the diaphragm has been repaired on ECMO, weaning and decannulation are accomplished as soon as possible. All infants are ventilated postoperatively to maintain preductal arterial oxygenation of 80 to 100 torr. Very slow weaning from the ventilator is necessary to avoid recurrent pulmonary hypertension.

Fetal tracheal occlusion is an experimental prenatal therapy for the treatment of severe congenital diaphragmatic hernia that reverses lung hypoplasia. The rationale for this approach is that the occlusion of the fetal trachea leads to net accumulation of lung liquid under pressure, which results in the development of large fluid-filled lungs. The balloon may be placed into the trachea under laparoscopic guidance, then removed prior to delivery when maximal lung growth has been achieved. The use of fetal tracheal occlusion remains investigational, although early reports are promising.

**Congenital Lobar Emphysema**

Congenital lobar emphysema (CLE) is a condition manifested during the first few months of life as a progressive hyperexpansion of one or more lobes of the lung. It can be life-threatening in the newborn period if extensive lung tissue is involved, but in the older infant and in cases in which the lesion is less severely distended it causes less respiratory distress. Air entering during inspiration is trapped in the lobe; on expiration, the lobe cannot deflate and progressively overexpands, causing atelectasis of the adjacent lobe or lobes. This hyperexpansion eventually shifts the mediastinum to the opposite side and compromises the other lung. CLE usually occurs in the upper lobes of the lung (left greater than right), followed next in frequency by the right middle lobe, but it also can occur in the lower lobes. It is caused by intrinsic bronchial obstruction from poor bronchial cartilage development or extrinsic compression. Approximately 14% of children with this condition have cardiac defects, with an enlarged left atrium or a major vessel causing compression of the ipsilateral bronchus.

Symptoms range from mild respiratory distress to full-blown respiratory failure with tachypnea, dyspnea, cough, and late cyanosis. These symptoms may be stationary or they may progress rapidly or result in recurrent pneumonia. Occasionally, infants with CLE present with failure to thrive, which likely reflects the increased work associated with the overexpanded lung. A hyperexpanded hemithorax on the ipsilateral side is pathognomonic for CLE. Diagnosis is typically confirmed by chest X-ray that shows a hyperlucent affected lobe with adjacent lobar compression and atelectasis. The mediastinum may be shifted as a consequence of mass effect to the contralateral side causing compression and atelectasis of the contralateral lung (Fig. 39-4). Although chest radiograph is usually sufficient, it is sometimes important to obtain at CT scan of the chest to clearly establish the diagnosis of CLE. This should be done only in the stable patient. Unless foreign body or mucous plugging is suspected as a cause of hyperinflation, bronchoscopy is not advisable because it can lead to more air trapping and cause life-threatening respiratory distress in a stable infant. Treatment is resection of the affected lobe, which can be safely performed using either an open or thorascoscopic approach. Unless symptoms necessitate earlier surgery, resection can usually be performed after the infant is several months of age. The prognosis is excellent.

**Figure 39-4.** Congenital lobar emphysema of the left upper lobe in a 2-week-old boy. Mediastinal shift is present.
Bronchopulmonary Foregut Malformations

Bronchopulmonary foregut malformations include foregut duplication cysts, congenital pulmonary airway malformations, and pulmonary sequestrations as discussed in the following sections.

Congenital Pulmonary Airway Malformations. Previously denoted as congenital cystic adenomatous malformation, (CCAM), congenital pulmonary airway malformations (CPAM) exhibits cystic proliferation of the terminal airway, producing cyst lined by mucus-producing respiratory epithelium, and elastic tissue in the cyst walls without cartilage formation. There may be a single cyst with a wall of connective tissue containing smooth muscle. Cysts may be large and multiple (type I), smaller and more numerous (type II), or they may resemble fetal lung without macroscopic cysts (type III). CPAMs frequently occur in the left lower lobe. However, this lesion can occur in any location and may occur in more than one lobe on more than one side, although this is rare. Clinical symptoms range from none to severe respiratory failure at birth. Over time, these malformations can be subject to repeated infections and produce fever and cough in older infants and children. The diagnosis is usually confirmed by CT for surgical planning and characteristic features that might delineate other bronchopulmonary foregut malformations (Fig. 39-5). Prenatal US may suggest the diagnosis. Resection is curative and may need to be performed urgently in the infant with severe respiratory distress. Long term, there is a risk of malignant degeneration in unresected CPAMs, but this risk occurs over decades and has not been fully defined. As a result, resection of the affected lobe is usually performed (Fig. 39-6). Antenatal resection may be rarely indicated in those instances in which fetal development is complicated by hydrops as a result of the mechanical and vascular effects of the lung lesion.

Pulmonary Sequestration. Pulmonary sequestration is uncommon and consists of a mass of lung tissue, usually in the left lower chest, occurring without the usual connections to the pulmonary artery or tracheobronchial tree, yet with a systemic blood supply from the aorta. There are two kinds of sequestration. Extralobar sequestration is usually a small area of nonaerated lung separated from the main lung mass, with a systemic blood supply, located immediately above the left diaphragm. It is commonly found in cases of CDH. Intralobar sequestration more commonly occurs within the parenchyma of the left lower lobe but can occur on the right. There is no major connection to the tracheobronchial tree, but a secondary connection may be established, perhaps through infection or via adjacent intrapulmonary shunts. The blood supply frequently originates from the aorta below the diaphragm; multiple vessels may be present (Fig. 39-7). Venous drainage of both types can be systemic or pulmonary. The cause of sequestration is unknown but most probably involves an abnormal budding of the developing lung that picks up a systemic blood supply and never becomes connected with the bronchus or pulmonary vessels. Sequestrations may, in some cases, exhibit mixed pathology with components consistent with CCAMs. Extralobar sequestration is asymptomatic and is usually discovered incidentally on chest X-ray. If the diagnosis can be confirmed, e.g., by CT scan, resection is not necessary. Diagnosis of intralobar sequestration may be made prenatally and confirmed on postnatal CT scan. Alternatively, the diagnosis of intralobar sequestration may be established after repeated infections manifested by cough, fever, and consolidation in the posterior basal segment of the left lower lobe. Increasingly the diagnosis is being made in the early months of life by US, and color Doppler often can be helpful in delineating the systemic arterial supply. Removal of the entire left lower lobe is usually necessary since the diagnosis often is made late after multiple infections. Occasionally segmental resection

Figure 39-5. Computed tomography scan of the chest showing a congenital cystic adenomatoid malformation of the left lower lobe.

Figure 39-6. Intraoperative photograph showing left lower lobe congenital cystic adenomatoid malformation seen in Fig. 39-5.

Figure 39-7. Arteriogram showing large systemic artery supply to intralobar sequestration of the left lower lobe.
of the sequestered part of the lung can be performed using an open, or ideally, a thoracoscopic approach. If an open approach is used, it is important to open the chest through a low intercostal space (sixth or seventh) to gain access to the vascular attachments to the aorta. These attachments may insert into the aorta below the diaphragm; in these cases, division of the vessels as they traverse the thoracic cavity is essential. Prognosis is generally excellent. However, failure to obtain adequate control of these vessels may result in their retraction into the abdomen and result in uncontrollable hemorrhage. It is also possible to perform a combined thoracoscopic and open approach, wherein the vessels are clipped and divided thoracoscopically and then the lesion safely removed through a limited thoracotomy.

**Bronchogenic Cyst.** Bronchogenic cysts are duplication cysts originating from the airway, regardless of the identity of the lining epithelial identity. They can occur anywhere along the respiratory tract and can present at any age, although typically they present after accumulation of intraluminal contents and not within the newborn period. Histologically, they are hamartomatous and usually consist of a single cyst lined with an epithelium; the mesenchyme contains cartilage and smooth muscle. They are probably embryonic rests of foregut origin that have been pinched off from the main portion of the developing tracheobronchial tree and are closely associated in causation with other foregut duplication cysts such as those arising from the esophagus. Bronchogenic cysts may be seen on prenatal US but are discovered most often incidentally on postnatal chest X-ray. Although they may be completely asymptomatic, bronchogenic cysts may produce symptoms, usually compressive, depending on the anatomic location and size, which increases over time if there is no egress for building luminal contents. In the para-tracheal region of the neck they can produce airway compression and respiratory distress. In the lung parenchyma, they may become infected and present with fever and cough. In addition, they may cause obstruction of the bronchial lumen with distal atelectasis and infection, or they may cause mediastinal compression. Rarely, rupture of the cyst can occur. Chest X-ray usually shows a dense mass, and CT scan or MRI delineates the precise anatomic location of the lesion. Treatment consists of resection of the cyst, which may need to be undertaken in emergency circumstances for airway or cardiac compression. Resection can be performed either as an open procedure, or more commonly using a thoracoscopic approach. If resection of a common wall will result in injury to the airway, resection of the inner epithelial cyst lining after marsupialization is acceptable.

**Bronchiectasis**

Bronchiectasis is an abnormal and irreversible dilatation of the bronchi and bronchioles associated with chronic suppurative disease of the airways. Usually patients have an underlying congenital pulmonary anomaly, cystic fibrosis, or immunologic deficiency. Bronchiectasis can also result from chronic infection secondary to a neglected bronchial foreign body. The symptoms include a chronic cough, often productive of purulent secretions, recurrent pulmonary infection, and hemoptysis. The diagnosis is suggested by a chest X-ray that shows increased bronchovascular markings in the affected lobe. Chest CT delineates bronchiectasis with excellent resolution. The preferred treatment for bronchiectasis is medical, consisting of antibiotics, postural drainage, and bronchodilator therapy because many children with the disease show signs of airflow obstruction and bronchial hyperresponsiveness. Lobectomy or segmental resection is indicated for localized disease that has not responded appropriately to medical therapy. In severe cases, lung transplantation may be required to replace the terminally damaged, septic lung.

**Foreign Bodies**

The inherent curiosity of children and their innate propensity to place new objects into their mouths to fully explore them place them at great risk for aspiration. Aspirated objects can be found either in the airway or in the esophagus; in both cases the results can be life-threatening.

**Airway Ingestion.** Aspiration of foreign bodies most commonly occurs in the toddler age group. Peanuts are the most common object that is aspirated, although other materials (popcorn, for instance) may also be involved. A solid foreign body often will cause air trapping, with hyperlucency of the affected lobe or lung seen especially on expiration. Oil from the peanut is very irritating and may cause pneumonia. Delay in diagnosis can lead to atelectasis and infection. The most common anatomic location for a foreign body is the right main stem bronchus or the right lower lobe. The child usually will cough or choke while eating but may then become asymptomatic. Total respiratory obstruction with tracheal foreign body may occur; however, respiratory distress is usually mild if present at all. A unilateral wheeze is often heard on auscultation. This wheeze often leads to an inappropriate diagnosis of “asthma” and may delay the correct diagnosis for some time. Chest X-ray will show a radiopaque foreign body, but in the case of nuts, seeds, or plastic toy parts, the only clue may be hyperexpansion of the affected lobe on an expiratory film or fluoroscopy. Bronchoscopy confirms the diagnosis and allows removal of the foreign body. It can be a very simple procedure or it may be extremely difficult, especially with a smooth foreign body that cannot be grasped easily or one that has been retained for some time. The rigid bronchoscope should be used in all cases, and utilization of the optical forcesps facilitates grasping the inhaled object. Epinephrine may be injected into the mucosa when the object has been present for a long period of time, which minimizes bleeding. Bronchietasis may be seen as an extremely late phenomenon after repeated infections of the poorly aerated lung and may require partial or total resection of the affected lobe. The differential diagnosis of a bronchial foreign body includes an intraluminal tumor (i.e., carcinoid, hemangioma, or neurofibroma).

**Foreign Bodies and Esophageal Injury.** The most common foreign body in the esophagus is a coin, followed by small toy parts. Toddlers are most commonly affected. The coin is retained in the esophagus at one of three locations: the cricopharyngeus, the area of the aortic arch, or the gastroesophageal junction, all of which are areas of normal anatomic narrowing. Symptoms are variable depending on the anatomic position of the foreign body and the degree of obstruction. There is often a relatively asymptomatic period after ingestion. The initial symptoms are gastrointestinal, and include dysphagia, drooling, and dehydration. The longer the foreign body remains in the esophagus with oral secretions unable to transit the esophagus, the greater the incidence of respiratory symptoms including cough, stridor, and wheezing. These findings may be interpreted as signs of upper respiratory infections. Objects that are present for a long period of time—particularly in children who have underlying neurological impairment—may manifest as chronic dysphagia. The chest X-ray is diagnostic in the case of a coin. A contrast swallow, or preferably an esophagoscopy, may be required for nonradiopaque foreign bodies. Coins lodged within the upper
esophagus for less than 24 hours may be removed using Magill forceps during direct laryngoscopy. For all other situations, the treatment is by esophagoscopy, rigid or flexible, and removal of the foreign body. In the case of sharp foreign bodies such as open safety pins, extreme care is required on extraction to avoid injury to the esophagus. Rarely, esophagotomy is required for removal, particularly of sharp objects. Diligent follow-up is required after removal of foreign bodies, especially batteries, which can cause strictures, and sharp objects, which can injure the underlying esophagus. In the case of a retained battery, this case should be handled as a surgical emergency, as the negative pole of the battery directly damages the surrounding tissue, and tracheoesophageal fistula, aortic exsanguination, and mediastinitis have all been described after local tissue necrosis at the site where the battery has lodged.

**ESOPHAGUS**

**Esophageal Atresia and Tracheoesophageal Fistula**

The management of esophageal atresia (EA) and tracheoesophageal fistula (TEF) is one of the most gratifying pediatric surgical conditions to treat. In the not so distant past, nearly all infants born with EA and TEF died. In 1939 Ladd and Leven achieved the first success repair by ligating the fistula, placing a gastrostomy, and reconstructing the esophagus at a later time. Subsequently, Dr. Cameron Haight, in Ann Arbor, Michigan, performed the first successful primary anastomosis for esophageal atresia, which remains the current approach for treatment of this condition. Despite the fact that there are several common varieties of this anomaly and the underlying cause remains obscure, a careful approach consisting of meticulous perioperative care and attention to the technical detail of the operation can result in an excellent prognosis in most cases.

**Anatomic Varieties.** The five major varieties of EA and TEF are shown in Fig. 39-8. The most commonly seen variety is esophageal atresia with distal tracheoesophageal fistula (type C), which occurs in approximately 85% of the cases in most series. The next most frequent is pure esophageal atresia (type A), occurring in 8% to 10% of patients, followed by tracheoesophageal fistula without esophageal atresia (type E). This occurs in 8% of cases and is also referred to as an H-type fistula, based upon the anatomic similarity to that letter (Fig. 39-9). Esophageal atresia with fistula between both proximal and distal ends of the esophagus and trachea (type D) is seen in approximately 2% of cases, and type B, esophageal atresia with tracheoesophageal fistula between distal esophagus and trachea, is seen in approximately 1% of all cases.

**Etiology and Pathologic Presentation.** The esophagus and trachea share a common embryologic origin. At approximately 4 weeks’ gestation, a diverticulum forms off the anterior aspect of the proximal foregut in the region of the primitive pharynx. This diverticulum extends caudally with progressive formation of the laryngo-tracheal groove, thus, creating a separate trachea and esophagus. Successful development of these structures is the consequence of extremely intricate interplay of growth and transcription factors necessary for rostral-caudal and anterior-posterior specification. The variations in clinically observed EA and TEF that must result in failure of successful formation of these structures are depicted in Fig. 39-8. While definitive genetic mutations have been difficult to identify in isolated EA-TEF, mutations in N-myc, Sox2, and CHD7 have been characterized in syndromic EA-TEF with associated anomalies. Other congenital anomalies commonly occur in association with EA-TEF. For instance, VACTERL syndrome is associated with vertebral anomalies (absent vertebrae or hemi-vertebrae) and anorectal anomalies (imperforate anus), cardiac

---

**Figure 39-8.** The five varieties of esophageal atresia and tracheoesophageal fistula. A. Isolated esophageal atresia. B. Esophageal atresia with tracheoesophageal fistula between proximal segment of esophagus and trachea. C. Esophageal atresia with tracheoesophageal fistula between distal esophagus and trachea. D. Esophageal atresia with fistula between both proximal and distal ends of esophagus and trachea. E. Tracheoesophageal fistula without esophageal atresia (H-type fistula).

**Figure 39-9.** Barium esophagram showing H-type tracheoesophageal fistula (arrow).
defects, tracheoesophageal fistula, renal anomalies (renal agenesis, renal anomalies), and radial limb hyperplasia. In nearly 20% of the infants born with esophageal atresia, some variant of congenital heart disease occurs.

Clinical Presentation of Infants With Esophageal Atresia and Tracheoesophageal Fistula. The anatomic variant of infants with EA-TEF predicts the clinical presentation. When the esophagus ends either as a blind pouch or as a fistula into the trachea (as in types A, B, C, or D), infants present with excessive drooling, followed by choking or coughing immediately after feeding is initiated as a result of aspiration through the fistula tract. As the neonate coughs and cries, air is transmitted through the fistula into the stomach, resulting in abdominal distention. As the abdomen distends, it becomes increasingly more difficult for the infant to breathe. This leads to further atelectasis, which compounds the pulmonary dysfunction. In patients with type C and D varieties, the regurgitated gastric juice passes through the fistula where it collects in the trachea and lungs and leads to a chemical pneumonitis, which further exacerbates the pulmonary status. In many instances, the diagnosis is actually made by the nursing staff who attempt to feed the baby and notice the accumulation of oral secretions.

The diagnosis of esophageal atresia is confirmed by the inability to pass an orogastric tube into the stomach (Fig. 39-10). The dilated upper pouch may be occasionally seen on a plain chest radiograph. If a soft feeding tube is used, the tube will coil in the upper pouch, which provides further diagnostic certainty. An important alternative diagnosis that must be considered when an orogastric tube does not enter the stomach is that of an esophageal perforation. This problem can occur in infants after traumatic insertion of a nasogastric or orogastric tube. In this instance, the perforation classically occurs at the level of the piriform sinus, and a false passage is created, which prevents the tube from entering the stomach. Whenever there is any diagnostic uncertainty, a contrast study will confirm the diagnosis of EA and occasionally document the TEF. The presence of a tracheoesophageal fistula can be demonstrated clinically by finding air in the gastrointestinal tract. This can be proven at the bedside by percussion of the abdomen and confirmed by obtaining a plain abdominal radiograph. Occasionally, a diagnosis of EA-TEF can be suspected prenatally on US evaluation. Typical features include failure to visualize the stomach and the presence of polyhydramnios. These findings reflect the absence of efficient swallowing by the fetus.

In a child with esophageal atresia, it is important to identify whether coexisting anomalies are present. These include cardiac defects in 38%, skeletal defects in 19%, neurologic defects in 15%, renal defects in 15%, anorectal defects in 8%, and other abnormalities in 13%. Examination of the heart and great vessels with echocardiography is important to exclude cardiac defects, as these are often the most important predictors of survival in these infants. The echocardiogram also demonstrates whether the aortic arch is left sided or right sided, which may influence the approach to surgical repair. Vertebral anomalies are assessed by plain radiography, and a spinal US is obtained if any are detected. A patent anus should be confirmed clinically. The kidneys in a newborn may be assessed clinically by palpation. A US of the abdomen will demonstrate the presence of renal anomalies, which should be suspected in the child who fails to make urine. The presence of extremity anomalies is suspected when there are missing digits and confirmed by plain radiographs of the hands, feet, forearms, and legs. Rib anomalies may also be present. These may include the presence of a 13th rib.

Initial Management. The initial treatment of infants with EA-TEF includes attention to the respiratory status, decompression of the upper pouch, and appropriate timing of surgery. Because the major determinant of poor survival is the presence of other severe anomalies, a search for other defects including congenital cardiac disease is undertaken in a timely fashion. The initial strategy after the diagnosis is confirmed is to place the neonate in an infant warmer with the head elevated at least 30°. A sump catheter is placed in the upper pouch on continuous suction. Both of these strategies are designed to minimize the degree of aspiration from the esophageal pouch. When saliva accumulates in the upper pouch and is aspirated into the lungs, coughing, bronchospasm, and desaturation episodes can occur, which may be minimized by ensuring the patency of the sump catheter. IV antibiotic therapy is initiated, and warmed electrolyte solution is administered. Where possible, the right upper extremity is avoided as a site to start an IV line, as this location may interfere with positioning of the patient during the surgical repair. Some surgeons place a central line in all patients to facilitate the administration of antibiotics and total parenteral nutrition as needed.

The timing of repair is influenced by the stability of the patient. Definitive repair of the EA-TEF is rarely a surgical emergency. If the child is hemodynamically stable and is oxygenating well, definitive repair may be performed within 1 to 2 days after birth. This allows for a careful determination of the presence of coexisting anomalies and for selection of an experienced anesthetic team.

Management of Esophageal Atresia and Tracheoesophageal Fistula in the Preterm Infant. The ventilated, premature neonate with EA-TEF and associated hyaline membrane disease represents a patient who may develop severe, progressive, cardiopulmonary dysfunction. The tracheoesophageal fistula can worsen the fragile pulmonary status as a result of recurrent aspiration through the fistula, and as a result of increased abdominal distention, which impairs lung expansion. Moreover, the elevated airway pressure that is required to ventilate these patients can worsen the clinical course by forcing air through the fistula into the stomach, thereby exacerbating the
degree of abdominal distention and compromising lung expansion. In this situation, the first priority is to minimize the degree of positive pressure needed to adequately ventilate the child. This can be accomplished using high frequency oscillatory ventilation (HFOV). If the gastric distention becomes severe, a gastrostomy tube should be placed. This procedure can be performed at the bedside under local anesthetic, if necessary. The dilated, air-filled stomach can easily be accessed through an incision in the left-upper quadrant of the abdomen. Once the gastrostomy tube is placed and the abdominal pressure is relieved, the pulmonary status can paradoxically worsen. This is because the ventilated gas may pass preferentially through the fistula, which is the path of least resistance, and bypass the lungs thereby worsening the hypoxemia. To correct this problem, the gastrostomy tube may be placed under water seal, elevated, or intermittently clamped. If these maneuvers are to no avail, ligation of the fistula may be required. This procedure can be performed in the neonatal intensive care unit if the infant is too unstable to be transported to the operating room. These interventions allow for the infant’s underlying hyaline membrane disease to improve, for the pulmonary secretions to clear, and for the infant to reach a period of stability so that definitive repair can be performed.

**Primary Surgical Correction.** In a stable infant, definitive repair is achieved through performance of a primary esophagegosphagostomy. There are two approaches to this operation: open thoracotomy or thoracoscopy. In the open approach, the infant is brought to the operating room, intubated, and placed in the lateral decubitus position with the right side up in preparation for right posterolateral thoracotomy. If a right-sided arch was determined previously by echocardiography, consideration is given to performing the repair through the left chest, although most surgeons believe that the repair can be performed safely from the right side as well. Bronchoscopy may be performed to exclude the presence of additional, upper-pouch fistulae in cases of esophageal atresia (i.e., differentiation of types B, C, and D variants) and identification of a laryngeo-tracheoesophageal cleft.

The operative technique for primary repair is as follows (Fig. 39-11). A retropleural approach is generally used as this technique prevents widespread contamination of the thorax if a postoperative anastomotic leak occurs. The sequence of steps is as follows: (a) mobilization of the pleura to expose the structures in the posterior mediastinum; (b) division of the fistula and closure of the tracheal opening; (c) mobilization of the upper esophagus sufficiently to permit an anastomosis without tension and to determine whether a fistula is present between the upper esophagus and the trachea (forward pressure by the anesthesia staff on the sump drain in the pouch can greatly facilitate dissection at this stage of the operation; care must be taken when dissecting posteriorly to avoid violation of either the lumen of trachea and esophagus); (d) mobilization of the distal esophagus (this needs to be performed judiciously to avoid

---

**Figure 39-11.** Primary repair of type C tracheosophageal fistula. A. Right thoracotomy incision. B. Azygos vein transected, proximal and distal esophagus demonstrated, and fistula identified. C. Tracheoesophageal fistula transected and defect in trachea closed. D. End-to-end anastomosis between proximal and distal esophagus (posterior row). E. Completed anastomosis.
devascularization since the blood supply to the distal esophagus is segmental from the aorta; most of the esophageal length is obtained from mobilizing the upper pouch since the blood supply travels via the submucosa from above; (e) performing a primary esophagosophageal anastomosis (most surgeons perform this procedure in a single layer using 5-0 sutures; if there is excess tension, the muscle of the upper pouch can be circumferentially incised without compromising blood supply to increase its length; many surgeons place a transanastomotic feeding tube in order to institute feeds in the early postoperative period); and (f) placement of a retropleural drain and closure of the incision in layers.

When a minimally invasive approach is selected, the patient is prepared for right-sided, transthoracic thorascoscopic repair. The same steps as described earlier for the open repair are undertaken, and the magnification and superb optics that are provided by the thorascoscopic approach provide for superb visualization. Identification of the fistula is performed as a first step; this can be readily ligated and divided between thorascopically placed sutures. The anastomosis is performed in a single layer. The thorascopically performed TEF repair requires clear and ongoing communication between the operating surgeons and the anesthesiologist; visualization can be significantly reduced with sudden changes in lung inflation, potentially leading to the need to convert to open repair. Although clear guidelines for patient selection for a thorascoscopic repair as opposed to an open repair remain lacking, reasonable selection criteria include patients over 2.5 kg who are hemodynamically stable and without comorbidities.

Postoperative Course. The postoperative management strategy of patients with EA-TEF is influenced to a great degree by the preference of the individual surgeon and the institutional culture. Many surgeons prefer not to leave the infants intubated postoperatively to avoid the effects of positive pressure on the site of tracheal closure. However, early extubation may not be possible in babies with preoperative lung disease either from prematurity or pneumonia or when there is any vocal cord edema. When a transanastomotic tube is placed, feeds are begun slowly in the postoperative period. Some surgeons institute parenteral nutrition for several days, using a central line. The retropleural drain is assessed daily for the presence of saliva, indicating an anastomotic leak. Many surgeons obtain a contrast swallow 1 week after repair to assess the caliber of the anastomosis and to determine whether a leak is present. If there is no leak, feedings are started. The principal benefit of the thorascoscopic approach is that postoperative pain is significantly reduced, as is the requirement for postoperative narcotic analgesia.

Complications of Surgery. Anastomotic leak occurs in 10% to 15% of patients and may be seen either in the immediate postoperative period or after several days. Early leakage (i.e., within the first 24 to 48 hours) is manifested by a new pleural effusion, pneumothorax, and sepsis and requires immediate exploration. In these circumstances, the anastomosis may be completely disrupted, possibly due to excessive tension. Revision of the anastomosis may be possible. If not, cervical esophagostomy and gastrostomy placement is required, with a subsequent procedure to reestablish esophageal continuity. Anastomotic leakage that is detected after several days usually heals without intervention, particularly if a retropleural approach is used. Under these circumstances, broad spectrum antibiotics, pulmonary toilet, and optimization of nutrition are important. After approximately a week or so, a repeat esophagram should be performed, at which time the leakage may have resolved.

Strictures at the anastomosis are not infrequent (10–20%), particularly if a leak has occurred. A stricture may become apparent at any time, from the early postoperative period to months or years later. It may present as choking, gagging, or failure to thrive, but it often becomes clinically apparent with the transition to eating solid food. A contrast swallow or esophagoscopy is confirmatory, and simple dilatation is usually corrective. Occasionally, repeated dilatations are required. These may be performed in a retrograde fashion, during which a silk suture is placed into the oropharynx and delivered from the esophagus through a gastrostomy tube. Tucker dilators are then tied to the suture and passed in a retrograde fashion from the gastrostomy tube and delivered out of the oropharynx. Increasing sizes are used, and the silk is replaced at the end of the procedure where it is taped to the side of the face at one end, and to the gastrostomy tube at the other. Alternatively, image-guided balloon dilation over a guide wire may be performed, using intraoperative contrast radiography to determine the precise location of the stricture and to assess the immediate response to the dilation.

"Recurrent" tracheosophageal fistula may represent a missed upper pouch fistula or a true recurrence. This may occur after an anastomotic disruption, during which the recurrent fistula may heal spontaneously. Otherwise, reoperation may be required. Recently, the use of fibrin glue has been successful in treating recurrent fistulas, although long-term follow-up is lacking.

Gastroesophageal reflux commonly occurs after repair of EA-TEF, potentially due to alterations in esophageal motility and the anatomy of the gastroesophageal junction. The clinical manifestations of such reflux are similar to those seen in other infants with primary gastroesophageal reflux disease (GERD). A loose antireflux procedure, such as a Nissen fundoplication, is used to prevent further reflux, but the child may have feeding problems after antireflux surgery as a result of the intrinsic dysmotility of the distal esophagus. The fundoplication may be safely performed laparoscopically in experienced hands, although care should be taken to ensure that the wrap is not excessively tight.

Special Circumstances. Patients with type E tracheosophageal fistulas (also called H-type) most commonly present beyond the newborn period. Presenting symptoms include recurrent chest infections, bronchospasm, and failure to thrive. The diagnosis is suspected using barium esophagography and confirmed by endoscopic visualization of the fistula. Surgical correction is generally possible through a cervical approach with concurrent placement of a balloon catheter across the fistula and requires mobilization and division of the fistula. Outcome is usually excellent.

Patients with duodenal atresia and EA-TEF may require urgent treatment due to the presence of a closed obstruction of the stomach and proximal duodenum. In stable patients, treatment consists of repair of the esophageal anomaly and correction of the duodenal atresia if the infant is stable during surgery. If not, a staged approach should be utilized consisting of ligation of the fistula and placement of a gastrostomy tube. Definitive repair can then be performed at a later point in time.

Primary esophageal atresia (type A) represents a challenging problem, particularly if the upper and lower ends are too far apart for an anastomosis to be created. Under these
circumstances, treatment strategies include placement of a gastrostomy tube and performing serial bougienage to increase the length of the upper pouch. This occasionally allows for primary anastomosis to be performed. Occasionally, when the two ends cannot be brought safely together, esophageal replacement is required using either a gastric pull-up or colon interposition (see the following section).

**Outcome.** Various classification systems have been utilized to predict survival in patients with EA-TEF and to stratify treatment. A system devised by Waterston in 1962 was used to stratify neonates based on birth weight, the presence of pneumonia, and the identification of other congenital anomalies. In response to advances in neonatal care, the surgeons from the Montreal Children’s Hospital proposed a new classification system in 1993. In the Montreal experience only two characteristics independently affected survival: preoperative ventilator dependence and associated major anomalies. Pulmonary disease as defined by ventilator dependence appeared to be more accurate than pneumonia. When the two systems were compared, the Montreal system more accurately identified children at highest risk. Spitz and colleagues analyzed risk factors in infants who died with EA-TEF. Two criteria were found to be important predictors of outcome: birth weight less than 1500 g and the presence of major congenital cardiac disease. A new classification for predicting outcome in esophageal atresia was therefore proposed: group I: birth weight ≥1500 g, without major cardiac disease, survival 97% (283 of 293); group II: birth weight <1500 g, or major cardiac disease, survival 59% (41 of 70); and group III: birth weight <1500 g, and major cardiac disease, survival 22% (2 of 9).

In general, surgical correction of EA-TEF leads to a satisfactory outcome with nearly normal esophageal function in most patients. Overall survival rates of greater than 90% have been achieved in patients classified as stable, in all the various staging systems. Unstable infants have an increased mortality (40–60% survival) because of potentially fatal associated cardiac and chromosomal anomalies or prematurity. However, the use of a staged procedure also has increased survival in even these high-risk infants.

**Corrosive Injury of the Esophagus**

Injury to the esophagus after ingestion of corrosive substances most commonly occurs in the toddler age group. Both strong alkali and strong acids produce injury by liquefaction or coagulation necrosis, and since all corrosive agents are extremely hygroscopic, the caustic substance will cling to the esophageal epithelium. Subsequent strictures occur at the anatomic narrowed areas of the esophagus, cricopharyngeus, midesophagus, and gastroesophageal junction. A child who has swallowed an injurious substance may be symptom-free but usually will be drooling and unable to swallow saliva. The injury may be restricted to the oropharynx and esophagus, or it may extend to include the stomach. There is no effective immediate antidote. Diagnosis is by careful physical examination of the mouth and endoscopy with a flexible or rigid esophagoscope. It is important to endoscope only to the first level of the burn in order to avoid perforation. Early barium swallow may delineate the extent of the mucosal injury. It is important to realize that the esophagus may be burned without evidence of injury to the mouth. Although previously used routinely, steroids have not been shown to alter stricture development or modify the extent of injury and are no longer part of the management of caustic injuries. Antibiotics are administered during the acute period.

The extent of injury is graded endoscopically as either mild, moderate, or severe (grade I, II, or III). Circumferential esophageal injuries with necrosis have an extremely high likelihood of stricture formation. These patients should undergo placement of a gastrostomy tube once clinically stable. A string should be inserted through the esophagus either immediately or during repeat esophagoscopy several weeks later. When established strictures are present (usually 3 to 4 weeks), dilatation is performed. Fluoroscopically guided balloon dilatation of the stricture is effective, which should be performed in association with esophagoscopy, and allows for a precise evaluation of the nature and extent of the stenosis. The procedure should be performed under general anesthesia, and care must be taken to ensure there is no airway injury. Dislodgment of the endotracheal tube can occur during this procedure, and careful communication with the anesthesiologist is critical during the procedure.

In certain circumstances, especially if a gastrostomy tube has been placed, retrograde dilatation may be performed, using graduated dilators brought through the gastrostomy and advanced into the esophagus via the transesophageal string. Management of esophageal perforation during dilatation should include antibiotics, irrigation, and closed drainage of the thoracic cavity to prevent systemic sepsis. When recognition is delayed or if the patient is systemically ill, esophageal diversion may be required with staged reconstruction at a later time.

Although the native esophagus can be preserved in most cases, severe stricture formation that does not respond to dilatation is best managed by esophageal replacement. The most commonly used options for esophageal substitution are the colon (right colon or transverse/left colon) and the stomach (gastric tubes or gastric pull-up). Pedicled or free grafts of the jejunum are rarely used. The right colon is based on a pedicle of the middle colic artery, and the left colon is based on a pedicle of the middle colic or left colic artery. Gastric tubes are fashioned from the greater curvature of the stomach based on the pedicle of the left gastropiploic artery. When the entire stomach is used, as in gastric pull-up, the blood supply is provided by the right gastric artery. The neoesophagus may traverse (a) submucosally; (b) through a transthoracic route; or (c) through the posterior mediastinum to reach the neck. A feeding jejunostomy is placed at the time of surgery and tube feedings are instituted once the postoperative ileus has resolved. Long-term follow-up has shown that all methods of esophageal substitution can support normal growth and development, and the children enjoy reasonably normal eating habits. Because of the potential for late complications such as ulceration and stricture, follow-up into adulthood is mandatory, but complications appear to diminish with time.

**Gastroesophageal Reflux**

Gastroesophageal reflux (GER) occurs to some degree in all children and refers to the passage of gastric contents into the esophagus. By contrast, gastroesophageal reflux disease (GERD) describes the situation where reflux is symptomatic. Typical symptoms include failure to thrive, bleeding, stricture formation, reactive airway disease, aspiration pneumonia, or apnea. Failure to thrive and pulmonary problems are particularly common in infants with GERD, whereas strictures and esophagitis are more common in older children and adolescents. GERD is particularly problematic in neurologically impaired children.

**Clinical Manifestations.** Because all infants experience occasional episodes of GER to some degree, care must be taken
before a child is labeled as having pathologic reflux. A history of repeated episodes of vomiting that interferes with growth and development, or the presence of apparent life-threatening events, are required for the diagnosis of GERD. In older children, esophageal bleeding, stricture formation, severe heartburn, or the development of Barrett’s esophagus unequivocally connotes pathologic reflux or GERD. In neurologically impaired children, vomiting due to GER must be distinguished from chronic retching.

The workup of patients suspected of having GERD includes documentation of the episodes of reflux and evaluation of the anatomy. A barium swallow should be performed as an initial test. This will determine whether there is obstruction of the stomach or duodenum (due to duodenal webs or pyloric stenosis) and will determine whether malrotation is present. The frequency and severity of reflux should be assessed using a 24-hour pH probe study. Although this test is poorly tolerated, it provides the most accurate determination that GERD is present. Esophageal endoscopy with biopsies may identify the presence of esophagitis, and it is useful to determine the length of intrabdominal esophagus and the presence of Barrett’s esophagus. Some surgeons obtain a radioisotope “milk scan” to evaluate gastric emptying, although there is little evidence to show that this test changes management when a diagnosis of GERD has been confirmed using the aforementioned modalities.

**Treatment.** Most patients with GERD are treated initially by conservative means. In the infant, propping and thickening the formula with rice cereal are generally recommended. Some authors prefer a prone, head-up position. In the infant unresponsive to position and formula changes and the older child with severe GERD, medical therapy is based on gastric acid reduction with an H₂-blocking agent and/or a proton pump inhibitor. Medical therapy is successful in most neurologically normal infants and younger children, many of whom will outgrow their need for medications. In certain patients, however, medical treatment does not provide symptomatic relief and surgery is therefore indicated. The least invasive surgical option includes the placement of a nasojejunal or gastrojejunal feeding tube. The tubes often become dislodged, acid reflux still occurs, and bolus feeding is generally not possible. Fundoplication provides definitive treatment for gastroesophageal reflux and is highly effective in most circumstances. The fundus may be wrapped around the distal esophagus either 360° (i.e., Nissen) or to lesser degrees (i.e., Thal or Toupet). At present, the standard approach in most children is to perform these procedures laparoscopically whenever possible. In children with feeding difficulties and in infants under 1 year of age, a gastrostomy tube should be placed at the time of surgery. Early postoperative complications include pneumonia and atelectasis, often due to inadequate pulmonary toilet and pain control with abdominal splinting. Late postoperative complications include wrap breakdown with recurrent reflux, which may require repeat fundoplication, and dysphagia due to a wrap performed too tightly, which generally responds to dilation. These complications are more common in children with neurologic impairment. The keys to successful surgical management of patients with GERD include careful patient selection and meticulous operative technique. There are emerging concerns regarding the long-term use of acid reducing agents, which may increase the frequency with which antireflux procedures are performed in children, especially those with neurological impairment.

**GASTROINTESTINAL TRACT**

**An Approach to the Vomiting Infant**

All infants vomit. Because infant vomiting is so common, it is important to differentiate between normal and abnormal vomiting, which may be indicative of a potentially serious underlying disorder. In order to determine the seriousness of a particular infant’s bouts of emesis, one needs to characterize what the vomit looks like and how sick the baby is. Vomit that looks like feeds and comes up immediately after a feeding is almost always gastroesophageal reflux. This may or may not be of concern, as described earlier. Vomiting that occurs a short while after feeding, or vomiting that projects out of the baby’s mouth may be indicative of pyloric stenosis. By contrast, vomit that has any green color in it is always worrisome. This may be reflective of intestinal volvulus, an underlying infection, or some other cause of intestinal obstruction. A more detailed description of the management of these conditions is provided in the following sections.

**Hypertrophic Pyloric Stenosis**

**Clinical Presentation.** Infants with hypertrophic pyloric stenosis (HPS) typically present with nonbilious vomiting that becomes increasingly projectile, over the course of several days to weeks due to progressive thickening of the pylorus muscle. HPS occurs in approximately 1 in 300 live births and commonly in infants between 3 and 6 weeks of age. Male-to-female ratio is nearly 5:1.

Eventually as the pyloric muscle thickening progresses, the infant develops a complete gastric outlet obstruction and is no longer able to tolerate any feeds. Over time, the infant becomes increasingly hungry, unsuccessfully feeds repeatedly, and becomes increasingly dehydrated. Wet diapers become less frequent, and there may even be a perception of less passage of flatus. HPS may be associated with jaundice due to an indirect hyperbilirubinemia, although the nature of this relation is unclear.

The cause of HPS has not been determined. Studies have shown that HPS is found in several generations of the same family, suggesting a familial link. Recently, a genome-wide significant locus for pyloric stenosis at chromosome 11q23.3 was identified, and the single-nucleotide polymorphism (SNP) with the greatest significance was associated with part of the genome that regulates cholesterol. It is not clear how this links to the development of pyloric stenosis, but it does suggest a potential dietary link.

Infants with HPS develop a hypochloremic, hypokalemic metabolic alkalosis. The urine pH level is high initially, but eventually drops because hydrogen ions are preferentially exchanged for sodium ions in the distal tubule of the kidney as the hypochloremia becomes severe (paradoxical aciduria). While in the past the diagnosis of pyloric stenosis was most often made on physical examination by palpation of the typical “olive” in the right upper quadrant and the presence of visible gastric waves on the abdomen, current standard of care is to perform an US, which can diagnose the condition accurately in 95% of patients. Criteria for US diagnosis include a channel length of over 16 mm and pyloric thickness over 4 mm. It is important to note that younger babies may have lower values...
for pyloric thickness and still be abnormal, and a close clinical correlation with the US result is mandatory. In cases in which the diagnosis remains unclear, upper gastrointestinal evaluation by contrast radiography will reveal delayed passage of contents from the stomach through the pyloric channel and a typical thickened appearance to the pylorus.

**Treatment.** Given frequent fluid and electrolyte abnormalities at time of presentation, pyloric stenosis is never a surgical emergency. Fluid resuscitation with correction of electrolyte abnormalities and metabolic alkalosis is essential prior to induction of general anesthesia for operation. For most infants, fluid containing 5% dextrose and 0.45% saline with added potassium of 2 to 4 mEq/kg over 24 hours at a rate of approximately 150 to 175 mL/kg for 24 hours will correct the underlying deficit. It is important to ensure that the child has an adequate urine output (>2 cc/kg per hour) as further evidence that rehydration has occurred.

After resuscitation, a Fredet-Ramstedt pyloromyotomy is performed (Fig. 39-12). It may be performed using an open or laparoscopic approach. The open pyloromyotomy is performed through either an umbilical or a right upper quadrant transverse abdominal incision. The former route is cosmetically more appealing, although the transverse incision provides easier access to the antrum and pylorus. In recent years, the laparoscopic approach has gained great popularity. Two randomized trials have demonstrated that both the open and laparoscopic approaches may be performed safely with equal incidence of postoperative complications, although the cosmetic result is clearly superior with the laparoscopic approach. Whether done through an open or laparoscopic approach, surgical treatment of pyloric stenosis involves splitting the pyloric muscle while leaving the underlying submucosa intact. The incision extends from just proximal to the pyloric vein of Mayo to the gastric antrum; it typically measures between 1 and 2 cm in length. Postoperatively, IV fluids are continued for several hours, after which Pedialyte is offered, followed by formula or breast milk, which is gradually increased to 24 cc every 3 hours. Most infants can be discharged home within 24 to 48 hours following surgery. Recently, several authors have shown that ad lib feeds are safely tolerated by the neonate and result in a shorter hospital stay.

The complications of pyloromyotomy include perforation of the mucosa (1–3%), bleeding, wound infection, and recurrent symptoms due to inadequate myotomy. When perforation occurs, the mucosa is repaired with a stitch that is placed to tack the mucosa down and reapproximate the serosa in the region of the tear. A nasogastric tube is left in place for 24 hours. The outcome is generally very good.

### Intestinal Obstruction in the Newborn

The cardinal symptom of intestinal obstruction in the newborn is bilious emesis. Prompt recognition and treatment of neonatal intestinal obstruction can truly be lifesaving.

The incidence of neonatal intestinal obstruction is 1 in 2000 live births. The approach to intestinal obstruction in the newborn infant is critical for timely and appropriate intervention. When a neonate develops bilious vomiting, one must consider a surgical etiology. Indeed, the majority of newborns with bilious emesis have a surgical condition. In evaluating a potential intestinal obstruction, it is helpful to determine whether the intestinal obstruction is either proximal or distal to the ligament of Treitz. One must conduct a detailed prenatal and immediate postnatal history and a thorough physical examination. In all cases of intestinal obstruction, it is vital to obtain abdominal films in the supine and upright (or lateral decubitus) views to assess the presence of air-fluid levels or free air as well as how far downstream air has managed to travel. Importantly, one should recognize that it is difficult to determine whether a loop of bowel is part of either the small or large intestine, as neonatal bowel lacks clear features, such as haustra or plica circulares, normally present in older children or adults. As such, contrast imaging may be necessary for diagnosis in some instances.

Proximal intestinal obstructions typically present with bilious emesis and minimal abdominal distention. The normal neonate should have a rounded, soft abdomen; in contrast, a neonate with a proximal intestinal obstruction typically exhibits a flat or scaphoid abdomen. On a series of upright and supine abdominal radiographs, one may see a paucity or absence of bowel gas, which normally should be present throughout the gastrointestinal tract within 24 hours. Of utmost importance is the exclusion of a malrotation with midgut volvulus from all other intestinal obstructions as this is a surgical emergency.

Distal obstructions typically present with bilious emesis and abdominal distention. Passage of black-green meconium should have occurred within the first 24 to 38 hours. Of great

---

**Figure 39-12.** Fredet-Ramstedt pyloromyotomy. A. Pylorus delivered into wound and seromuscular layer incised. B. Seromuscular layer separated down to submucosal base to permit herniation of mucosa through pyloric incision. C. Cross-section demonstrating hypertrophied pylorus, depth of incision, and spreading of muscle to permit mucosa to herniate through incision.
Figure 39-13. Abdominal X-ray showing “double bubble” sign in a newborn infant with duodenal atresia. The two “bubbles” are numbered.
when there is diagnostic uncertainty, or when distal intestinal obstruction is apparent, a barium enema is useful to establish whether a microcolon is present and to diagnose the presence of meconium plugs, small left colon syndrome, Hirschsprung’s disease, or meconium ileus. Judicious use of barium enema is therefore required in order to safely manage neonatal intestinal obstruction, based on an understanding of the expected level of obstruction.

Surgical correction of the small intestinal atresia should be performed relatively urgently, especially when there is a possibility of volvulus. At laparotomy, one of several types of atresia will be encountered. In type 1 there is a mucosal atresia with intact muscularis. In type 2, the atretic ends are connected by a fibrous band. In type 3A, the two ends of the atresia are separated by a V-shaped defect in the mesentery. Type 3B is an “apple-peel” deformity or “Christmas tree” deformity in which the bowel distal to the atresia receives its blood supply in a retrograde fashion from the ileocolic or right colic artery (Fig. 39-15). In type 4 atresia, there are multiple atresias with a “string of sausage” or “string of beads” appearance. Disparity in lumen size between the proximal distended bowel and the small diameter of collapsed bowel distal to the atresia has led to a number of innovative techniques of anastomosis. However, under most circumstances, an anastomosis can be performed using the end-to-back technique in which the distal, compressed loop is “fish-mouthed” along its antimesenteric border. The proximal distended loop can be tapered as previously described. Because the distended proximal bowel rarely has normal motility, the extremely dilated portion should be resected prior to performing the anastomosis.

Occasionally the infant with intestinal atresia will develop ischemia or necrosis of the proximal segment secondary to volvulus of the dilated, bulbous, blind-ending proximal bowel. Under these conditions, primary anastomosis may be performed as described earlier. Alternatively, an end ileostomy and mucus fistula should be created, and the anastomosis should be deferred to another time after the infant stabilizes.

Malrotation and Midgut Volvulus
Embryology. During the sixth week of fetal development, the midgut grows too rapidly to be accommodated in the abdominal cavity and therefore herniates into the umbilical cord. Between the 10th and 12th week, the midgut returns to the abdominal cavity, undergoing a 270° counterclockwise rotation around the superior mesenteric artery. Because the duodenum also rotates caudal to the artery, it acquires a C-loop, which traces this path. The cecum rotates cephalad to the artery, which determines the location of the transverse and ascending colon. Subsequently, the duodenum becomes fixed retroperitoneally in its third portion and at the ligament of Treitz, while the cecum becomes fixed to the lateral abdominal wall by peritoneal bands. The takeoff of the branches of the superior mesenteric artery elongates and becomes fixed along a line extending from its emergence from the aorta to the cecum in the right lower quadrant. Genetic mutations likely disrupt the signaling critical for normal intestinal rotation. For instance, mutations in the gene BCL6 resulting in absence of left-sided expression of its transcript lead to reversed cardiac orientation, defective ocular development, and malrotation. The essential role of the dorsal gut mesentery in mediating normal intestinal rotation and the role of the forkhead box transcription factor FOXF1 in formation of the dorsal mesentery in mice are consistent with the noted association of intestinal malrotation with alveolar capillary dysplasia, caused by mutations in FOXF1. If rotation is incomplete, the cecum remains in the epigastrium, but the bands fixing the duodenum to the retroperitoneum and cecum continue to form. This results in (Ladd’s) bands extending from the cecum to the lateral abdominal wall and crossing the duodenum, which creates the potential for obstruction. The mesenteric takeoff remains confined to the epigastrium, resulting in a narrow pedicle suspending all the branches of the superior mesenteric artery and the entire midgut. A volvulus may therefore occur around the mesentery. This twist not only obstructs the proximal jejunum but also cuts off the blood supply to the midgut. Intestinal obstruction and complete infarction of the midgut occur unless the problem is promptly corrected surgically.

Presentation and Management. Midgut volvulus can occur at any age, though it is seen most often in the first few weeks of life. Bilious vomiting is usually the first sign of volvulus and all infants with bilious vomiting must be evaluated rapidly to ensure that they do not have intestinal malrotation with volvulus. The child with irritability and bilious emesis should raise particular suspicions for this diagnosis. If left untreated, vascular...
compromise of the midgut initially causes bloody stools, but it eventually results in circulatory collapse. Additional clues to the presence of advanced ischemia of the intestine include erythema and edema of the abdominal wall, which progresses to shock and death. It must be reemphasized that the index of suspicion for this condition must be high, since abdominal signs are minimal in the early stages. Abdominal films show a paucity of gas throughout the intestine with a few scattered air-fluid levels (Fig. 39-16). When these findings are present, the patient should undergo immediate fluid resuscitation to ensure adequate perfusion and urine output followed by prompt exploratory laparotomy. In cases where the child is stable, laparoscopy may be considered.

Often the patient will not appear ill, and the plain films may suggest partial duodenal obstruction. Under these conditions, the patient may have malrotation without volvulus. This is best diagnosed by an upper gastrointestinal series that shows incomplete rotation with the duodenojejunal junction displaced to the right. The duodenum may show a corkscrew effect diagnosing volvulus, or complete duodenal obstruction, with the small bowel loops entirely in the right side of the abdomen. Barium enema may show a displaced cecum, but this sign is unreliable, especially in the small infant in whom the cecum is normally in a somewhat higher position than in the older child.

When volvulus is suspected, early surgical intervention is mandatory if the ischemic process is to be avoided or reversed. Volvulus occurs clockwise, and it is therefore untwisted counterclockwise. This can be remembered using the memory aid "turn back the hands of time." Subsequently, a Ladd’s procedure is performed. This operation does not correct the malrotation, but it does broaden the narrow mesenteric pedicle to prevent volvulus from recurring. This procedure is performed as follows (Fig. 39-17). The bands between the cecum and the abdominal wall and between the duodenum and terminal ileum are divided sharply to splay out the superior mesenteric artery and its branches. This maneuver brings the straightened duodenum into the right lower quadrant and the cecum into the left lower quadrant. The appendix is usually removed to avoid diagnostic errors in later life. No attempt is made to suture the cecum or duodenum in place. With advanced ischemia, reduction of the volvulus without the Ladd’s procedure is accomplished, and a “second look” 24 to 36 hours later often may show some vascular recovery. A plastic transparent silo may be placed to facilitate constant evaluation of the intestine and to plan for the timing of reexploration. Clearly necrotic bowel can then be resected conservatively. With early diagnosis and correction, the prognosis is excellent. However, diagnostic delay can lead to mortality or to short-gut syndrome requiring intestinal transplantation.

A subset of patients with malrotation will demonstrate chronic obstructive symptoms. These symptoms may result from Ladd’s bands across the duodenum, or occasionally, from intermittent volvulus. Symptoms include intermittent abdominal pain and intermittent vomiting that may occasionally be bilious. Infants with malrotation may demonstrate failure to thrive, and they may be diagnosed initially as having gastroesophageal reflux disease. Surgical correction using Ladd’s procedure as described earlier can prevent volvulus from occurring and improve symptoms in many instances. In these cases, a laparoscopic approach may be taken, where diagnosis of Ladd’s bands and direct visualization of the relevant anatomy may be achieved.

**Meconium Ileus**

**Pathogenesis and Clinical Presentation.** Infants with cystic fibrosis have characteristic pancreatic enzyme deficiencies and abnormal chloride secretion in the intestine that result in the production of viscous, water-poor meconium. This phenotype is explained by the presence of mutations in the CFTR gene. Meconium ileus occurs when this thick, highly viscous meconium becomes impacted in the ileum and leads to high-grade intestinal obstruction. Recently, additional mutations were identified in genes encoding multiple apical plasma membrane proteins of infants with meconium ileus. Meconium ileus can be either uncomplicated, in which there is no intestinal perforation, or complicated, in which prenatal perforation of the intestine has occurred or vascular compromise of the distended ileum develops. Antenatal US may reveal the presence of intra-abdominal or scrotal calcifications, or distended bowel loops. These infants present shortly after birth with progressive abdominal distention and failure to pass meconium with intermittent bilious emesis. Abdominal radiographs show dilated loops of intestine. Because the enteric contents are so viscous, air-fluid levels do not form, even when obstruction is complete. Small bubbles of gas become entrapped in the inspissated meconium in the distal ileum, where they produce a characteristic “ground glass” appearance.

The diagnosis of meconium ileus is confirmed by a contrast enema that typically demonstrates a microcolon. In patients with uncomplicated meconium ileus, the terminal ileum is filled with pellets of meconium. In patients with complicated meconium ileus, intraperitoneal calcifications form, producing an eggshell pattern on plain abdominal X-ray.

**Management.** The treatment strategy depends on whether the patient has complicated or uncomplicated meconium ileus. Patients with uncomplicated meconium ileus can be
treated nonoperatively. Either dilute water-soluble contrast or N-acetylcysteine (Mucomyst) is infused transanally via catheter under fluoroscopic control into the dilated portion of the ileum. Because these agents act by absorbing fluid from the bowel wall into the intestinal lumen, infants undergoing treatment are at risk of fluid and electrolyte abnormalities so that appropriate resuscitation of the infant during this maneuver is extremely important. The enema may be repeated at 12-hour intervals over several days until all the meconium is evacuated. Inability to reflux the contrast into the dilated portion of the ileum signifies the presence of an associated atresia or complicated meconium ilus, and thus warrants exploratory laparotomy. If surgical intervention is required because of failure of contrast enemas to relieve obstruction, operative irrigation with dilute contrast agent, N-acetylcysteine, or saline through a purse-string suture may be successful. Alternatively, resection of the distended terminal ileum is performed, and the meconium pellets are flushed from the distal small bowel. At this point, an end ileostomy may be created. The distal bowel may be brought up as a mucus fistula or sewn to the side of the ileum as a classic Bishop-Koop anastomosis. An end-to-end anastomosis may also be considered in the appropriate setting (Fig. 39-18).

**Necrotizing Enterocolitis**

**Clinical Features.** Necrotizing enterocolitis (NEC) is the most frequent and lethal gastrointestinal disorder affecting the intestine of the stressed, preterm neonate. The overall mortality ranges between 10% and 50%. Advances in neonatal care such as surfactant therapy as well as improved methods of mechanical ventilation have resulted in increasing numbers of
low-birth-weight infants surviving neonatal hyaline membrane disease. An increasing proportion of survivors of neonatal respiratory distress syndrome will therefore be at risk for developing NEC. Consequently, it is estimated that NEC may eventually surpass respiratory distress syndrome as the principal cause of death in the preterm infant. This is especially relevant, as NEC is a significant risk factor for more severe respiratory distress in premature infants.

Multiple risk factors have been associated with the development of NEC. These include prematurity, initiation of enteral feeding, bacterial infection, intestinal ischemia resulting from birth asphyxia, umbilical artery cannulation, persistence of a patent ductus arteriosus, cyanotic heart disease, and maternal cocaine abuse. Nonetheless, the mechanisms by which these complex interacting etiologies lead to the development of the disease remain undefined. The only consistent epidemiologic precursors for NEC are prematurity and enteral alimentation, representing the commonly encountered clinical situation of a stressed infant who is fed enterally. Of note, there is some debate regarding the type and strategy of enteral alimentation in the pathogenesis of NEC. A prospective randomized study showed no increase in the incidence of NEC despite an aggressive feeding strategy.

The indigenous intestinal microbial flora has been shown to play a central role in the pathogenesis of NEC. The importance of bacteria in the pathogenesis of NEC is further supported by the finding that NEC occurs in epistemic waves that can be abrogated by infection control measures, and the fact that NEC usually develops at least 10 days postnatally, when the GI tract is colonized by coliforms. More recently, outbreaks of NEC have been reported in infants fed formula contaminated with *Enterobacter sakazakii*. Common bacterial isolates from the blood, peritoneal fluid, and stool of infants with advanced NEC include *Escherichia coli*, *Enterobacter*, Klebsiella, and occasionally, coagulase-negative *Staphylococcus* species.

NEC may involve single or multiple segments of the intestine, most commonly the terminal ileum, followed by the colon. The gross findings in NEC include bowel distention with patchy areas of thinning, pneumatosis, gangrene, or frank perforation. The microscopic features include the appearance of a “bland infarct” characterized by full thickness necrosis.
Clinical Manifestations. Infants with NEC present with a spectrum of disease. In general, the infants are premature and may have sustained one or more episodes of stress, such as birth asphyxia, or they may have congenital cardiac disease. The clinical picture of NEC has been characterized as progressing from a period of mild illness to that of severe, life-threatening sepsis by Bell and colleagues. Although not all infants progress through the various “Bell stages,” this classification scheme provides a useful format to describe the clinical picture associated with the development of NEC. In the earliest stage (Bell stage I), infants present with feeding intolerance. This is suggested by vomiting or by the presence of a large residual volume from a previous feeding in the stomach at the time of the next feeding. Following appropriate treatment, which consists of bowel rest and IV antibiotics, many of these infants will not progress to more advanced stages of NEC. These infants are colloquially described as suffering from an “NEC scare” and represent a population of neonates who are at risk of developing more severe NEC if a more prolonged period of stress supervenes.

Infants with Bell stage II have established NEC that is not immediately life-threatening. Clinical findings include abdominal distention and tenderness, bilious nasogastric aspirate, and bloody stools. These findings indicate the development of intestinal ileus and mucosal ischemia, respectively. Abdominal examination may reveal a palpable mass indicating the presence of an inflamed loop of bowel, diffuse abdominal tenderness, cellulitis, and edema of the anterior abdominal wall. The infant may appear systemically ill, with decreased urine output, hypotension, tachycardia, and noncardiac pulmonary edema. Hematologic evaluation reveals either leukocytosis or leukopenia, an increase in the number of bands, and thrombocytopenia. An increase in the blood urea nitrogen and plasma creatinine level may be found, which signify the development of renal dysfunction. The diagnosis of NEC may be confirmed by abdominal radiography. The pathognomonic radiographic finding in NEC is pneumatosis intestinalis, which represents invasion of the ischemic mucosa by gas producing microbes (Fig. 39-19). Other findings include the presence of ileus or portal venous gas. The latter is a transient finding that indicates the presence of severe NEC with intestinal necrosis. A fixed loop of bowel may be seen on serial abdominal radiographs, which suggests the possibility that a diseased loop of bowel, potentially with a localized perforation, is present. Although these infants are at risk of progressing to more severe disease, with timely and appropriate treatment, they often recover.

Infants with Bell stage III have the most advanced form of NEC. Abdominal radiographs often demonstrate the presence of pneumoperitoneum, indicating that intestinal perforation has occurred. These patients may develop a fulminant course with progressive peritonitis, acidosis, sepsis, disseminated intravascular coagulopathy, and death.

Pathogenesis of Necrotizing Enterocolitis. Several theories have been proposed to explain the development of NEC. In general terms, the development of diffuse pneumatosis intestinalis—which is associated with the development of stage II NEC—is thought to be due to the presence of gas within the wall of the intestine from enteric bacteria, suggesting the causative role of bacteria in the pathogenesis of NEC. Furthermore, the development of pneumoperitoneum indicates disease progression with severe disruption of the intestinal barrier (intestinal perforation).

Finally, systemic sepsis with diffuse multisystem organ dysfunction suggests the role for circulating proinflammatory cytokines in the pathogenesis of NEC. It has also been demonstrated that the premature intestine responds in an exaggerated fashion to bacterial products, rendering the host susceptible to barrier dysfunction and the development of NEC. Various groups have shown that NEC pathogenesis requires activation of the bacterial receptor—Toll-like receptor 4 (TLR4)—in the intestinal epithelium. The expression of TLR4 is significantly elevated in the premature infant intestine as compared with the full-term infant intestine, a consequence of the role that TLR4 plays in normal intestinal development. When the infant is born prematurely and TLR4 expression levels are elevated, subsequent activation of TLR4 by colonizing bacteria in the neonatal intensive care unit leads to the induction of a severe proinflammatory response and the development of NEC. It is noteworthy that breast milk—long known to be protective against NEC—is able to suppress TLR4 signaling and that synthetic TLR4 antagonists are known to prevent NEC in preclinical models, suggesting the possibility of preventive approaches for this disease.

Treatment. In all infants suspected of having NEC, feedings are discontinued, a nasogastric tube is placed, and broad-spectrum parenteral antibiotics are given. The infant is resuscitated, and inotropes are administered to maintain perfusion as needed. Intubation and mechanical ventilation may be required to maintain oxygenation. Total parenteral nutrition is started. Subsequent treatment may be influenced by the particular stage of NEC that is present. Patients with Bell stage I are closely monitored and generally remain NPO and on IV antibiotics for 7 to 10 days, prior to reintroducing enteral nutrition. If the infant fully recovers, feedings may be reinitiated.

Patients with Bell stage II disease merit close observation. Serial physical examinations are performed looking for the development of diffuse peritonitis, a fixed mass, progressive abdominal wall cellulitis or systemic sepsis. If infants fail to improve after several days of treatment, consideration should be given to exploratory laparotomy. Paracentesis may be performed, and if the Gram stain demonstrates multiple organisms and leukocytes, perforation of the bowel should be suspected, and patients should undergo laparotomy.

Figure 39-19. Abdominal radiograph of infant with necrotizing enterocolitis. Arrows point to area of pneumatosis intestinalis.
In the most severe form of NEC (Bell stage III), patients have definite intestinal perforation or have not responded to nonoperative therapy. Two schools of thought direct further management. One group favors exploratory laparotomy. At laparotomy, frankly gangrenous or perforated bowel is resected, and the intestinal ends are brought out as stomas. When there is massive intestinal involvement, marginally viable bowel is retained and a “second-look” procedure is carried out after the infant stabilizes (24–48 hours). Patients with extensive necrosis at the second look may be managed by placing a proximal diverting stoma, resecting bowel that is definitely not viable, and leaving questionably viable bowel behind, distal to the diverted segment. When the intestine is viable except for a localized perforation without diffuse peritonitis and if the infant’s clinical condition permits, intestinal anastomosis may be performed. In cases where the diseased, perforated segment cannot be safely resected, drainage catheters may be left in the region of the diseased bowel, and the infant is allowed to stabilize.

An alternative approach to the management of infants with perforated NEC involves drainage of the peritoneal cavity. This may be performed under local anesthesia at the bedside, and it can be an effective means of stabilizing the desperately ill infant by relieving increased intra-abdominal pressure and allowing ventilation. When successful, this method also allows for drainage of perforated bowel by establishing a controlled fistula. Approximately one-third of infants treated with drainage alone survive without requiring additional operations. Infants that do not respond to peritoneal drainage alone after 48 to 72 hours should undergo laparotomy. This procedure allows for the resection of frankly necrotic bowel diversion of the fecal stream and facilitates more effective drainage. It is noteworthy that a recent randomized controlled trial demonstrated that outcomes were similar in infants with NEC that were treated either with primary peritoneal drainage or laparotomy, although this study was criticized for the large number of patients who were excluded from randomization. There was also concern that a number of patients who were thought to have NEC may actually have had spontaneous intestinal perforation, given their lack of pneumatosis and relatively early onset of presentation; these patients would be anticipated to improve after peritoneal drainage due to the more local nature of their disease process.

Necrotizing Enterocolitis in Older Infants. Although NEC is typically a disease that affects preterm infants, several independent groups have reported a tendency for early onset of NEC in term and near-term infants. In these patients, the pattern of disease was found to be different from that found in premature infants. Specifically, NEC in older infants typically is localized to the end of the small intestine and beginning of the colon, suggestive of an ischemic pathophysiologic. There are four pertinent associations that are observed in term infants that develop NEC: congenital heart disease, in utero growth restriction, polycythemia, and perinatal hypoxic-ischemic events. As with NEC in preterm infants, NEC in older patients is also associated with formula consumption and is very rare in exclusively breastfed infants. Patients with NEC at full term typically present with bloody stools and may be characterized by rapid onset of symptoms and a fulminant course. Thus, although it is true that NEC is typically a disease of premature babies, in the appropriate setting, NEC can develop at any age.

Spontaneous Intestinal Perforation Versus Necrotizing Enterocolitis. In addition to NEC, preterm infants with intestinal pathology may develop spontaneous intestinal perforation (SIP). SIP is a distinct clinical entity from NEC, and it is essentially a perforation in the terminal ileum. The histopathology of SIP is different from NEC. Specifically, the mucosa is intact and not necrotic, there is no sign of ischemia, and the submucosa is thinned at the site of perforation. In contrast to NEC, pneumatosis intestinalis is absent in SIP. Moreover, the demographics of NEC and SIP are slightly different, in that patients with SIP tend to be slightly more premature, smaller, and more likely to have been on inotropic support. SIP occurs in two separate time points, both within a few days after birth and approximately 10 days later, and in all cases, free air will be present, but pneumatosis will be absent. Because patients with SIP have isolated disease without necrosis or systemic inflammation, they tend to have a better outcome and are likely to respond better to peritoneal drainage. In short, the diagnosis of SIP versus NEC has important prognostic significance. Treatment for SIP should primarily be surgical, with intestinal resection and stoma creation, followed by stoma reversal once the child is stable.

In both SIP and NEC, the timing of stoma closure is a matter of ongoing debate. Whereas in the past, pediatric surgeons typically waited until the child reached 5 kg or so, experience indicates that there is no benefit in waiting this long, and children tolerate stoma closure very well when they are at much lower weights. One approach is to close the stoma when the calculated gestational age is approximately 38 to 40 weeks, which will, on average, be at approximately 6 weeks after the initial surgery. This time point is selected based on the observation that proinflammatory gene expression has normalized by then, and NEC recurrence is very unlikely.

Outcome. Survival in patients with NEC is dependent on the stage of disease, the extent of prematurity, and the presence of associated comorbidities. Survival by stage has recently been shown to be approximately 85%, 65%, and 35% for stages I, II, and III, respectively. Strictures develop in 20% of medically or surgically treated patients, and a contrast enema is mandatory before reestablishing intestinal continuity. If all other factors are favorable, the ileostomy is closed when the child is between 2 and 2.5 kg. At the time of stoma closure, the entire intestine should be examined to search for areas of NEC. Patients who develop massive intestinal necrosis are at risk of developing short bowel syndrome, particularly when the total length of the viable intestinal segment is less than 40 cm. These patients require TPN to provide adequate calories for growth and development, and may develop parenteral nutrition–associated cholestasis and hepatic fibrosis. In a significant number of these patients, transplantation of the liver and small bowel may be required.

Short Bowel Syndrome

Short bowel syndrome (SBS) is an extremely morbid condition with an increasing incidence. Various congenital and perinatal acquired conditions such as gastrochisis, malrotation, atresia, and NEC may lead to SBS. Medical and surgical treatment options carry high dollar and human costs and morbidities including multiple infections and hospitalizations for vascular access, liver failure in conjunction with parenteral nutrition–associated cholestasis, and death. Medical centers that have developed multidisciplinary clinics focused on treating children with short bowel syndrome have achieved significant success in
Intussusception

Intussusception is the leading cause of intestinal obstruction in the young child. It refers to the condition whereby a segment of intestine becomes drawn into the lumen of the more proximal bowel. The process usually begins in the region of the terminal ileum, and extends distally into the ascending, transverse, or descending colon. Rarely, an intussusception may prolapse through the rectum.

The cause of intussusception is not clear, although one hypothesis suggests that hypertrophy of the Peyer’s patches in the terminal ileum from an antecedent viral infection acts as a lead point. Peristaltic action of the intestine then causes the bowel distal to the lead point to invaginate into itself. Idiopathic intussusception occurs in children between the ages of approximately 6 and 24 months of age. Beyond this age group, one should consider the possibility that a pathologic lead point maybe present. These include polyps, malignant tumors such as lymphoma, enteric duplication cysts or Meckel’s diverticulum. Such intussusceptions are rarely reduced by air or contrast enema, and thus the lead point is identified when operative reduction of the intussusception is performed.

Clinical Manifestations. Since intussusception is frequently preceded by a gastrointestinal viral illness, the onset may not be easily determined. Typically, the infant develops paroxysms of crampy abdominal pain and intermittent vomiting. Between attacks, the infant may act normally, but as symptoms progress, increasing lethargy develops. Bloody mucus ("currant-jelly" stool) may be passed per rectum. Ultimately, if reduction is not accomplished, gangrene of the intussusceptum occurs, and perforation may ensue. On physical examination, an elongated mass is detected in the right upper quadrant or epigastrium with an absence of bowel in the right lower quadrant (Dance’s sign). The mass may be seen on plain abdominal X-ray but is more easily demonstrated on air or contrast enema.

Treatment. Patients with intussusception should be assessed for the presence of peritonitis and for the severity of systemic illness. Following resuscitation and administration of IV antibiotics, the child is assessed for suitability to proceed with radiographic versus surgical reduction. In the absence of peritonitis, the child should undergo radiographic reduction. If peritonitis is present, or if the child appears systemically ill, urgent laparotomy is indicated.

In the stable patient, the air enema is both diagnostic and may be curative, and it is the preferred method of diagnosis and treatment of intussusception. Air is introduced with a manometer, and the pressure that is administered is carefully monitored. Under most instances, this should not exceed 120 mmHg. Successful reduction is marked by free reflux of air into multiple loops of small bowel and symptomatic improvement as the infant suddenly becomes pain free. Unless both of these signs are observed, it cannot be assumed that the intussusception is reduced. If reduction is unsuccessful, and the infant remains stable, the infant should be brought back to the radiology suite for a repeat attempt at reduction after a few hours. This strategy has improved the success rate of nonoperative reduction in many centers. In addition, hydrostatic reduction with barium may be useful if pneumatic reduction is unsuccessful. The overall success rate of radiographic reduction varies based on the experience of the center, and it is typically between 60% and 90%.

If nonoperative reduction is successful, the infant may be given oral fluids after a period of observation. Failure to reduce the intussusception mandates surgery. which can be approached through an open or laparoscopic technique. In an open procedure, exploration is carried out through a right lower quadrant incision, delivering the intussuscepted mass into the wound. Reduction usually can be accomplished by gentle distal pressure, where the intussusceptum is gently milked out of the intussuscipiens (Fig. 39-20). Care should be taken not to pull the bowel out, as this can cause damage to the bowel wall. The blood supply to the appendix is often compromised, and appendectomy is therefore often performed. If the bowel is frankly gangrenous, resection and primary anastomosis is performed. In experienced hands, laparoscopic reduction may be performed, even in very young infants. This is performed using a 5-mm laparoscope placed in the umbilicus, and two additional 5 mm ports in the left and right lower quadrants. The bowel is inspected, and if it appears to be viable, reduction is performed by milking the bowel or using gentle traction, although this approach is normally discouraged during manual reduction. Atraumatic bowel graspers allow the bowel to be handled without injuring it.

IV fluids are continued until the postoperative ileus subsides. Patients are started on clear liquids, and their diet is advanced as tolerated. Of note, recurrent intussusception occurs in 5% to 10% of patients, independent of whether the bowel is reduced radiographically or surgically. Patients present with recurrent symptoms in the immediate postoperative period. Treatment involves repeat air enema, which is successful in most cases. In patients who experience three or more episodes of intussusception, the presence of a pathologic lead point should be suspected and carefully evaluated using contrast studies. After the third episode of intussusception, many pediatric surgeons will perform an exploratory laparotomy to reduce the bowel and to resect a pathologic lead point if identified.

Appendicitis

Presentation. Correct diagnosis of appendicitis in children can be one of the most humbling and challenging tasks facing the pediatric surgeon. The classical presentation is known to all students and practitioners of surgery: generalized abdominal pain that localizes to the right lower quadrant followed by nausea, vomiting, fever, and localized peritoneal irritation in the region of McBurney’s point. When children present in this
manner, there should be little diagnostic delay. The child should be made NPO, administered IV fluids and broad-spectrum antibiotics, and brought to the operating room for an appendectomy. However, children often do not present in this manner. The coexistence of nonspecific viral syndromes and the inability of young children to describe the location and quality of their pain often result in diagnostic delay. As a result, children with appendicitis often present with perforation, particularly those who are under 5 years of age. Perforation increases the length of hospital stay and makes the overall course of the illness significantly more complex.

**Diagnosis of Appendicitis in Children.** There have been significant improvements in the role of radiographic studies in the diagnosis of acute appendicitis. While CT is quite reliable in making the diagnosis, US is very useful when performed in experienced centers and good visualization of the appendix is achieved. MRI may be performed where available with high specificity and sensitivity—and avoidance of radiation. US is very useful for excluding ovarian causes of abdominal pain. Despite these radiographic measures, the diagnosis of appendicitis remains largely clinical, and each clinician should develop his or her own threshold to operate or to observe the patient. A reasonable practice guideline is as follows. When the diagnosis is clinically apparent, appendectomy should obviously be performed with minimal delay. Localized right lower quadrant tenderness associated with low-grade fever and leukocytosis in boys should prompt surgical exploration. In girls, ovarian or uterine pathology must also be considered. When there is diagnostic uncertainty, the child may be observed, rehydrated, and reassessed. In girls of menstruating age, an US may be obtained to exclude ovarian pathology (cysts, torsion, or tumor). If all studies are negative, yet the pain persists, and the abdominal findings remain equivocal, diagnostic laparoscopy may be employed to determine the etiology of the abdominal pain. The appendix should be removed even if it appears to be normal, unless another pathologic cause of the abdominal pain is definitively identified and the appendectomy would substantially increase morbidity.

**Surgical Treatment of Appendicitis.** The definitive treatment for acute appendicitis is appendectomy. Prior to surgery, it is important that patients receive adequate IV fluids in order to correct dehydration that commonly develops as a result of fever and vomiting in patients with appendicitis. Patients should also be started on antibiotics (such as a second-generation cephalosporin). Most surgeons will perform a laparoscopic appendectomy, which may have some advantage over removing the appendix through a single, larger incision. During the laparoscopic appendectomy, a small incision is made at the umbilicus, and two additional incisions are made in the lower abdomen. The appendix is typically delivered through the umbilicus, and all incisions are then closed, with dissolvable sutures. If the appendix is not ruptured, the patient may start drinking liquids shortly after waking up from the operation, and may be advanced to a solid diet the next day. In general, the same steps are taken when appendectomy is performed through an open approach. The most common complication after appendectomy is a surgical site infection. Other risks—including bleeding or damage to other structures inside the abdomen—are extremely rare. Recovery from surgery is dependent upon the individual patient. Most children are back to school approximately 1 week from surgery and usually are allowed to return to full physical activity after 2 to 3 weeks. During the recovery period, over-the-counter pain medication may be required. Older patients tend to require a longer time for full recovery.

**Management of the Child With Perforated Appendicitis.** The signs and symptoms of perforated appendicitis can closely mimic those of gastroenteritis and include abdominal pain, vomiting, and diarrhea. Alternatively, the child may present with symptoms of intestinal obstruction. An abdominal mass may be present in the lower abdomen. When the symptoms have been present for more than 4 or 5 days, and an abscess is suspected, it is reasonable to obtain a computerized tomogram of the abdomen and pelvis with IV, oral, and rectal contrast in order to visualize the appendix and the presence of an associated abscess, phlegmon, or fecalith (Fig. 39-21).

An individualized approach is necessary for the child who presents with perforated appendicitis. When there is evidence of generalized peritonitis, intestinal obstruction or evidence of systemic toxicity, the child should undergo appendectomy. This should be delayed only for as long as is required to ensure adequate fluid resuscitation and administration of broad-spectrum antibiotics. The operation can be performed through an open or through a laparoscopic approach. One distinct advantage of the laparoscopic approach is that it provides excellent visualization of the pelvis and all four quadrants of the abdomen. At the time of surgery, adhesions are gently lysed, abscess cavities are drained and the appendix is removed. Drains are seldom used, and the skin incisions can be closed primarily. If a fecalith is identified outside the appendix on computerized tomography, every effort should be made to retrieve it and to remove it along with the appendix, if at all possible. Often, the child in whom symptoms have been present for more than 4 or 5 days will present with an abscess without evidence of generalized peritonitis. Under these circumstances, it is appropriate to perform image-guided percutaneous drainage of the abscess followed by broad-spectrum antibiotic therapy. The inflammation will generally subside within several days, and the appendix can be safely removed as an outpatient 6 to 8 weeks later. If the child’s symptoms do not improve, or if the abscess is not amenable to percutaneous drainage, then laparoscopic or open appendectomy and abscess drainage is required. Patients who present with a phlegmon in the region of a perforated appendix may be managed in a similar manner. In general, children who are younger...
than 4 or 5 years of age do not respond as well to an initial nonoperative approach because their bodies do not localize or isolate the inflammatory process. Thus, these patients are more likely to require early surgical intervention. Patients who have had symptoms of appendicitis for no more than 4 days should probably undergo “early” appendectomy because the inflammatory response is not as excessive during that initial period and the procedure can be performed safely.

**Nonoperative Management of Acute Appendicitis.** Despite the fact that surgical removal of the acutely inflamed appendix is effective in all cases, there has been a growing recognition that certain children will respond to antibiotics alone and thus avoid surgery. Several trials have shown that acute appendicitis may be treated with antibiotics alone effectively in nearly 80% of patients. However, the failure rate is considered unacceptably high for many patients, who effectively will have suffered a delay from definitive care. Furthermore, the heterogeneity of disease presentation, and varying degree of illness severity, make it quite difficult to predict who will respond to antibiotics alone. This question is currently being answered in the United States in the form of a randomized controlled trial that is recruiting over 1500 patients in eight states, which will be divided into antibiotic therapy versus surgery (ClinicalTrials.gov, identifier NCT02800785).

**Other Causes of Abdominal Pain That Mimic Appendicitis in Children.** As mentioned earlier, appendicitis can be one of the most difficult diagnoses to establish in children with abdominal pain, in part because of the large number of diseases that present in a similar fashion. Patients with urinary tract infection can present very similarly to those with appendicitis. However, patients with urinary tract infection are less likely to present with vomiting and are likely to also experience difficulty with urination, characterized by pressure, burning, and frequency. Constipation may be commonly confused with appendicitis in its earliest stages. However, patients with constipation rarely have fever and will not have abnormalities in their blood work. Ovarian torsion can mimic appendicitis, given the severe abdominal pain that accompanies this condition. However, patients with ovarian torsion are generally asymptomatic until the acute onset of severe pain. By contrast, patients with appendicitis generally experience gradual onset of pain associated with nausea and vomiting. Finally, children and young adults are always at risk for the development of gastroenteritis. However, unlike appendicitis, patients with gastroenteritis generally present with persistent vomiting and occasionally diarrhea, which precedes the onset of the abdominal pain.

**Intestinal Duplications**

Duplications represent mucosa-lined structures that are in continuity with the gastrointestinal tract. Although they can occur at any level in the gastrointestinal tract, duplications are found most commonly in the ileum within the leaves of the mesentery. Duplications may be long and tubular but usually are cystic masses. In all cases, they share a common wall with the intestine. Symptoms associated with enteric duplication cysts include recurrent abdominal pain, emesis from intestinal obstruction, or hematochezia. Such bleeding typically results from ulceration in the duplication or in the adjacent intestine if the duplication contains ectopic gastric mucosa. On examination, a palpable mass is often identified. Children may also develop intestinal obstruction. Torsion may produce gangrene and perforation.

The ability to make a preoperative diagnosis of enteric duplication cyst usually depends on the presentation. CT, US, and technetium pertechnetate scanning can be very helpful. Occasionally, a duplication can be seen on small bowel follow-through or technetium imaging. In the case of short duplications, resection of the cyst and adjacent intestine with end-to-end anastomosis can be performed. If resection of long duplications would compromise intestinal length, multiple enterotomies and mucosal stripping in the duplicated segment will allow the walls to collapse and become adherent. An alternative method is to divide the common wall using the GIA stapler, forming a common lumen. Patients with duplications who undergo complete excision without compromise of the length of remaining intestine have an excellent prognosis.

**Meckel’s Diverticulum**

A Meckel’s diverticulum is a remnant of the embryonic omphalomesenteric (vitelline) duct. It is located on the antimesenteric border of the ileum, usually within 2 ft of the ileocecal valve (Fig. 39-22). It may be found incidentally at surgery or may present with inflammation masquerading as appendicitis. Perforation of a Meckel’s diverticulum may occur if the outpouching becomes impacted with food, leading to distention and necrosis. Occasionally, bands of tissue extend from the Meckel’s diverticulum to the anterior abdominal wall, and these may represent lead points around which internal hernias may develop. This is an important cause of intestinal obstruction in the older child who has a scarless abdomen. Similar to duplications, ectopic gastric mucosa may produce ileal ulcerations that bleed and lead to the passage of maroon-colored stools. Pancreatic mucosa may also be present. Diagnosis may be made by technetium pertechnetate scans when the patient presents with bleeding. Treatment is surgical. If the base is narrow and there is no mass present in the lumen of the diverticulum, a wedge resection of the diverticulum with transverse closure of the ileum can be performed. A linear stapler is especially useful in this circumstance. When a mass of ectopic tissue is palpable, if the base is wide, or when there is inflammation, it is preferable to perform a resection of the involved bowel and end-to-end ileoileostomy.

**Mesenteric Cysts**

Mesenteric cysts are similar to duplications in their location within the mesentry. However, they do not contain any mucosa or muscular wall. Chylous cysts may result from congenital

![Figure 39-22. Operative photograph showing the presence of a Meckel’s diverticulum (arrow).](image)
lymphatic obstruction. Mesenteric cysts can cause intestinal obstruction or may present as an abdominal mass. The diagnosis may be made by abdominal US or CT. Treatment involves surgical excision. This may require resection of the adjacent intestine, particularly for extensive, multicystic lesions. In cases where complete excision is not possible due to the close proximity to vital structures, partial excision or marsupialization should be performed.

**Hirschsprung’s Disease**

**Pathogenesis.** In his classic textbook entitled *Pediatric Surgery*, Dr. Orvar Swenson, who is eponymously associated with one of the classic surgical treatments for Hirschsprung’s disease, described this condition as follows: “Congenital megacolon is caused by a malformation in the pelvic parasympathetic system which results in the absence of ganglion cells in Auerbach’s plexus of a segment of distal colon. Not only is there an absence of ganglion cells, but the nerve fibers are large and excessive in number, indicating that the anomaly may be more extensive than the absence of ganglion cells.” This narrative of Hirschsprung’s disease is as accurate today as it was more than 50 years ago and summarizes the essential pathologic features of this disease: absence of ganglion cells in Auerbach’s plexus and hypertrophy of associated nerve trunks. The cause of Hirschsprung’s disease remains incompletely understood, although current thinking suggests that the disease results from a defect in the migration of neural crest cells, which are the embryonic precursors of the intestinal ganglion cell. Under normal conditions, the neural crest cells migrate into the intestine from cephalad to caudal. The process is completed by the 12th week of gestation, but the migration from midtransverse colon to anus takes 4 weeks. During this latter period, the fetus is most vulnerable to defects in migration of neural crest cells. This may explain why most cases of aganglionosis involve the rectum and rectosigmoid. The length of the aganglionic segment of bowel is therefore determined by the most distal region that the migrating neural crest cells reach. In rare instances, total colonic aganglionosis may occur.

Recent studies have shed light on the molecular basis for Hirschsprung’s disease. Patients with Hirschsprung’s disease have an increased frequency of mutations in several genes, including *GDNF*, its receptor *Ret*, or its coreceptor *Gfra-1*. Moreover, mutations in these genes also lead to aganglionic megacolon in mice, which provides the opportunity to study the function of the encoded proteins. Initial investigations indicate that *GDNF* promotes the survival, proliferation, and migration of mixed populations of neural crest cells in culture. Other studies have revealed that *GDNF* is expressed in the gut in advance of migrating neural crest cells and is chemotactic for neural crest cells in culture. These findings raise the possibility that mutations in the *GDNF* or *Ret* genes could lead to impaired neural crest migration in utero and the development of Hirschsprung’s disease.

**Clinical Presentation.** The incidence of sporadic Hirschsprung’s disease is 1 in 5000 live births. There are reports of increased frequency of Hirschsprung’s disease in multiple generations of the same family. Occasionally, such families have mutations in the genes described earlier, including the *Ret* gene. Because the aganglionic colon does not permit normal peristalsis to occur, the presentation of children with Hirschsprung’s disease is characterized by a functional distal intestinal obstruction. In the newborn period, the most common symptoms are abdominal distention, failure to pass meconium, and bilious emesis. Any infant who does not pass meconium beyond 48 hours of life must be investigated for the presence of Hirschsprung’s disease. Occasionally, infants present with a dramatic complication of Hirschsprung’s disease called *enterocolitis*. This pattern of presentation is characterized by abdominal distention and tenderness, and it is associated with manifestations of systemic toxicity that include fever, failure to thrive, and lethargy. Infants are often dehydrated and demonstrate a leukocytosis or increase in circulating band forms on hematologic evaluation. On rectal examination, forceful expulsion of foul-smelling liquid feces is typically observed and represents the accumulation of stool under pressure in an obstructed distal colon. Treatment includes rehydration, systemic antibiotics, nasogastric decompression, and rectal irrigations while the diagnosis of Hirschsprung’s disease is being confirmed. In children that do not respond to nonoperative management, a decompressive stoma is required. It is important to ensure that this stoma is placed in ganglion-containing bowel, which must be confirmed by frozen section at the time of stoma creation.

In approximately 20% of cases, the diagnosis of Hirschsprung’s disease is made beyond the newborn period. These children have severe constipation, which has usually been treated with laxatives and enemas. Abdominal distention and failure to thrive may also be present at diagnosis.

**Diagnosis.** The definitive diagnosis of Hirschsprung’s disease is made by rectal biopsy. Samples of mucosa and submucosa are obtained at 1 cm, 2 cm, and 3 cm from the dentate line. This can be performed at the bedside in the neonatal period without anesthesia, as samples are taken in bowel that does not have somatic innervation and is thus not painful to the child. In older children, the procedure should be performed using IV sedation. The histopathology of Hirschsprung’s disease is the absence of ganglion cells in the myenteric plexuses, increased acetylcholinesterase staining, and the presence of hypertrophied nerve bundles.

It is important to obtain a barium enema in children in whom the diagnosis of Hirschsprung’s disease is suspected. This test may demonstrate the location of the transition zone between the dilated ganglionic colon and the distal constricted aganglionic rectal segment. Our practice is to obtain this test before instituting rectal irrigations if possible so that the difference in size between the proximal and distal bowel is preserved. Although the barium enema can only suggest, but not reliably establish, the diagnosis of Hirschsprung’s disease, it is very useful in excluding other causes of distal intestinal obstruction. These include small left colon syndrome (as occurs in infants of diabetic mothers), colonic atresia, meconium plug syndrome, or the unused colon observed in infants after the administration of magnesium or tocolytic agents. The barium enema in total colonic aganglionosis may show a markedly shortened colon. Some surgeons have found the use of rectal manometry helpful, particularly in older children, although it is relatively inaccurate.

**Treatment.** The diagnosis of Hirschsprung’s disease requires surgery in all cases. The classic surgical approach consisted of a multiple stage procedure. This included a colostomy in the newborn period, followed by a definitive pull-through operation after the child was over 10 kg. There are three viable options for the definitive pull-through procedure that are currently used. Although individual surgeons may advocate one procedure over another, studies have demonstrated that the outcome after each type of operation is similar. For each of
the operations that is performed, the principles of treatment include confirming the location in the bowel where the transition zone between ganglionic and aganglionic bowel exists, resecting the aganglionic segment of bowel, and performing an anastomosis of ganglionicated bowel to either the anus or a cuff of rectal mucosa (Fig. 39-23).

It is now well established that a primary pull-through procedure can be performed safely, even in the newborn period. This approach follows the same treatment principles as a staged procedure and saves the patient from an additional surgical procedure. Many surgeons perform the intra-abdominal dissection using the laparoscope. This approach is especially useful in the newborn period as this provides excellent visualization of the pelvis. In children with significant colonic distention, it is important to allow for a period of decompression using a rectal tube if a single-staged pull-through is to be performed. In older children with very distended, hypertrophied colon, it may be prudent to perform a colostomy to allow the bowel to decompress prior to performing a pull-through procedure. However, it should be emphasized that there is no upper age limit for performing a primary pull-through.

Of the three pull-through procedures performed for Hirschsprung’s disease, the first is the original Swenson procedure. In this operation, the aganglionic rectum is dissected in the pelvis and removed down to the anus. The ganglionic colon is then anastomosed to the anus via a perineal approach. In the Duhamel procedure, dissection outside the rectum is confined to the retrorectal space, and the ganglionic colon is anastomosed posteriorly just above the anus. The anterior wall of the ganglionic colon and the posterior wall of the aganglionic rectum are anastomosed, using a stapler. Although both of these procedures are extremely effective, they are limited by the possibility of damage to the parasympathetic nerves that are adjacent to the rectum. To circumvent this potential problem, Soave’s procedure involves dissection entirely within the rectum. The rectal mucosa is stripped from the muscular sleeve, and the ganglionic colon is brought through this sleeve and anastomosed to the anus. This operation may be performed completely from below. In all cases, it is critical that the level at which ganglionicated bowel exists be determined. Most surgeons believe that the anastomosis should be performed at least 5 cm from the point at which ganglion cells are found. This avoids performing a pull-through in the transition zone, which is associated with a high incidence of complications due to inadequate emptying of the pull-through segment. Up to one-third of patients who undergo a transition zone pull-through will require a reoperation.

The main complications of all procedures include postoperative enterocolitis, constipation, and anastomotic stricture. There is also a reported incidence of recurrent Hirschsprung’s disease, which may reflect either residual aganglionic bowel left behind after the pull-through, or the presence of ischemia in the pulled-through segment leading to ganglion cell loss. Long-term results with the three procedures are comparable and generally excellent in experienced hands. These three procedures also can be adapted for total colonic aganglionosis in which the ileum is used for the pull-through segment.

Anorectal Malformations

Anatomic Description. Anorectal malformations describe a spectrum of congenital anomalies that include imperforate anus and persistent cloaca. Anorectal malformations occur in approximately 1 in 5000 live births and affect males and females almost equally. The embryologic basis includes failure of descent of the urorectal septum. The level to which this septum descends determines the type of anomaly that is present, which subsequently influences the surgical approach.

In patients with imperforate anus, the rectum fails to descend through the external sphincter complex. Instead, the rectal pouch ends “blindly” in the pelvis, above or below the levator ani muscle. In most cases, the blind rectal pouch communicates more distally with the genitourinary system or with the perineum through a fistulous tract. Traditionally, anatomic
SPECIFIC CONSIDERATIONS

PART II

Figure 39-24. Low imperforate anus in a male. Note the well-developed buttocks. The perineal fistula was found at the midline raphe.

Imperforate anus in a female. A catheter has been placed into the fistula, which is in the vestibule of the vagina.

Description of imperforate anus has been characterized as either “high” or “low” depending on whether the rectum ends above the levator ani muscle complex or partially descends through this muscle (Fig. 39-24). Based upon this classification system, in male patients with high imperforate anus the rectum usually ends as a fistula into the membranous urethra. In females, high imperforate anus often occurs in the context of a persistent cloaca. In both males and females, low lesions are associated with a fistula to the perineum. In males, the fistula connects with the median raphe of the scrotum or penis. In females, the fistula may end within the vestibule of the vagina, which is located immediately outside the hymen or at the perineum.

Because this classification system is somewhat arbitrary, Peña proposed a classification system that specifically and unambiguously describes the location of the fistulous opening. In men, the fistula may communicate with: (a) the perineum (cutaneous perineal fistula); (b) the lowest portion of the posterior urethra (rectourethral bulbar fistula); (c) the upper portion of the posterior urethra (rectourethral prostatic fistula); or (d) the bladder neck (rectovesical fistula). In females, the urethra may open to the perineum between the female genitalia and the center of the sphincter (cutaneous perineal fistula) or into the vestibule of the vagina (vestibular fistula) (Fig. 39-25).

In both sexes, the rectum may end in a completely blind fashion (imperforate anus without fistula). In rare cases, patients may have a normal anal canal, yet there may be total atresia or severe stenosis of the rectum.

The most frequent defect in males is imperforate anus with rectourethral fistula, followed by rectoperineal fistula, then rectovesical fistula or rectobladder neck. In females, the most frequent defect is the rectovestibular defect, followed by the cutaneous perineal fistula. The third most common defect in females is the persistent cloaca. This lesion represents a wide spectrum of malformations in which the rectum, vagina, and urinary tract meet and fuse into a single common channel. On physical examination, a single perineal orifice is observed, and it is located at the place where the urethra normally opens. Typically, the external genitalia are hypoplastic.

Associated Malformations. Approximately 60% of patients have an associated malformation. The most common is a urinary tract defect, which occurs in approximately 50% of patients. Skeletal defects are also seen, and the sacrum is most commonly involved. Spinal cord anomalies especially tethered cord are common, particularly in children with high lesions. Gastrointestinal anomalies occur, most commonly esophageal atresia. Cardiac anomalies may be noted, and occasionally patients present with a constellation of defects as part of the VACTERL syndrome (described earlier).

Management of Patients With Imperforate Anus. Patients with imperforate anus are usually stable, and the diagnosis is readily apparent. Despite the obstruction, the abdomen is initially not distended, and there is rarely any urgency to intervene. The principles of management center around diagnosing the type of defect that is present (high vs. low), and evaluating the presence of associated anomalies. It may take up to 24 hours before the presence of a fistula on the skin is noted, and thus it is important to observe the neonate for some period of time before definitive surgery is undertaken. All patients should therefore have an orogastric tube placed and be monitored for the appearance of meconium in or around the perineum or in the urine. Investigation for associated defects should include an US of the abdomen to assess for the presence of urinary tract anomaly. Other tests should include an echocardiogram and spinal radiographs. An US of the spine should be performed to look for the presence of a tethered cord. To further classify the location of the fistula as either “high” versus “low,” a lateral abdominal radiograph can be obtained with a radiopaque marker on the perineum. By placing the infant in the inverted position, the distance between the most distal extent of air in the rectum and the perineal surface can be measured. This study is imprecise, however, and may add little to the overall management of these patients.

The surgical management of infants with imperforate anus is determined by the anatomic defect. In general, when a low lesion is present, only a perineal operation is required without a colostomy. Infants with a high lesion require a colostomy in the newborn period, followed by a pull-through procedure at approximately 2 months of age. When a persistent cloaca is present, the urinary tract needs to be carefully evaluated at the time of colostomy formation to ensure that normal emptying can occur and to determine whether the bladder needs to be drained by means of a vesicostomy. If there is any doubt about the type of lesion, it is safer to perform a colostomy rather than jeopardize the infant’s long-term chances for continence by an injudicious perineal operation.
The type of pull-through procedure favored by most pediatric surgeons today is the posterior sagittal anorectoplasty (PSARP procedure), as described by Peña and DeVries. This involves placing the patient in the prone jack-knife position, dividing the levator ani and external sphincter complex in the midline posteriorly, dividing the communication between the gastrointestinal tract and the urinary tract, and bringing down the rectum after sufficient length is achieved. The muscles are then reconstructed and sutured to the rectum. The outcome of 1192 patients who had undergone this procedure has been reviewed by Peña and Hong. Seventy-five percent of patients were found to have voluntary bowel movements, and nearly 40% were considered totally continent. As a rule, patients with high lesions demonstrate an increase incidence of incontinence, whereas those with low lesions are more likely to be constipated. Management of patients with high imperforate anus can be greatly facilitated using a laparoscopic assisted approach, in which the patient is operated on in the supine position, and the rectum is mobilized down to the fistulous connection to the bladder neck. This fistulous connection is then divided, and the rectum is completely mobilized down to below the peritoneal reflection. The operation then proceeds at the perineum, and the location of the muscle complex is determined using the nerve stimulator. A Veress needle is then advanced through the skin at the indicated site, with the laparoscope providing guidance to the exact intrapelvic orientation. Dilators are then placed over the Veress needle, the rectum is then pulled through this peritoneal opening, and an anoplasty is performed.

JAUNDICE

The Approach to the Jaundiced Infant

Jaundice is present during the first week of life in 60% of term infants and 80% of preterm infants. There is usually accumulation of unconjugated bilirubin, but there may also be deposition of direct bilirubin. During fetal life, the placenta is the principal route of elimination of unconjugated bilirubin. In the newborn infant, bilirubin is conjugated through the activity of \textit{glucoronyl transferase}. In the conjugated form, bilirubin is water soluble, which results in its excretion into the biliary system and then into the gastrointestinal tract. Newborns have a relatively high level of circulating hemoglobin and relative immaturity of the conjugating machinery. This results in a transient accumulation of bilirubin in the tissues, which is manifested as jaundice. Physiologic jaundice is evident by the second or third day of life and usually resolves within approximately 5 to 7 days. By definition, jaundice that persists beyond 2 weeks is considered pathologic.

Pathologic jaundice may be due to biliary obstruction, increased hemoglobin load, or to liver dysfunction. The workup of the jaundiced infant therefore should include a search for the following possibilities: (a) obstructive disorders, including biliary atresia, choledochal cyst, and inissipated bile syndrome; (b) hematologic disorders, including ABO incompatibility, \textit{Rh} incompatibility, \textit{spherocytosis}; (c) metabolic disorders, including \textit{α}-1 antitrypsin deficiency, \textit{galactosemia}; pyruvate kinase deficiency; and (d) congenital infection, including syphilis and rubella.

Biliary Atresia

Pathogenesis. Biliary atresia is a rare disease associated with significant morbidity and mortality. This disease is characterized by a fibroproliferative obliteration of the biliary tree which progresses toward hepatic fibrosis, cirrhosis, and end-stage liver failure. The incidence of this disease is approximately 1 in 8000 to 1 in 18,000. The etiology of biliary atresia is likely multifactorial. In the classic textbook, \textit{Abdominal Surgery of Infancy and Childhood}, Ladd and Gross described the cause of biliary atresia as an “arrest of development during the solid stage of bile duct formation.” Previously proposed theories on the etiology of biliary atresia have focused on defects in hepatogenesis, prenatal vasculogenesis, immune dysregulation, infectious agents, and exposure to toxins. More recently, genetic mutations in the \textit{cfc1} gene, implicated in left-right axis determinations, were identified in patients with biliary atresia-splenic malformation syndrome. Additionally, the detection of higher incidence of maternal microchimerism in the livers of males with biliary atresia has led to the suggestion that consequent expression of maternal antigens may lead to an autoimmune process leading to inflammation and obliteration of the biliary tree. Recent animal studies strongly implicate perinatal exposure to reovirus or rotavirus. Such viral exposure may lead to perportal inflammation mediated by interferon-γ and other cytokines.

Clinical Presentation. Infants with biliary atresia present with jaundice at birth or shortly thereafter. The diagnosis of biliary atresia is frequently not entertained by pediatricians in part because physiologic jaundice of the newborn is so common and biliary atresia is so uncommon. As such, it is not unusual for there to be a delay in diagnosis. However, infants with biliary atresia characteristically have acholic, pale gray appearing stools, secondary to obstructed bile flow. With further passage of time, these infants manifest progressive failure to thrive, and if untreated, develop stigmata of liver failure and portal hypertension, particularly splenomegaly and esophageal varices.

The oblitative process of biliary atresia involves the common duct, cystic duct, one or both hepatic ducts, and the gallbladder, in a variety of combinations. The histopathology of patients with biliary atresia includes inflammatory changes within the parenchyma of the liver, as well as fibrous deposition at the portal plates that is observed on trichrome staining of frozen tissue sections. In certain cases, bile duct proliferation may be seen, a relatively nonspecific marker of liver injury. Approximately 25% of patients with biliary atresia have coincidental malformations, often associated with polysplenia, and may include intestinal malrotation, preduodenal portal vein, and intrahepatic vena cava.

Diagnosis. In general, the diagnosis of biliary atresia is made utilizing a combination of studies, as no single test is sufficiently sensitive or specific. Fractionation of the serum bilirubin is performed to determine if the associated hyperbilirubinemia is conjugated or unconjugated. Workup commonly includes the analysis of TORCH infection titers as well as viral hepatitis. Typically, a US is performed to assess the presence of other causes of biliary tract obstruction, including choledochal cyst. The absence of a gallbladder is highly suggestive of the diagnosis of biliary atresia. However, the presence of a gallbladder does not exclude the diagnosis of biliary atresia because in approximately 10% of biliary atresia patients, the distal biliary tract is patent and a gall bladder may be visualized, even though the proximal ducts are atretic. It is important to note that the intrahepatic bile ducts are never dilated in patients with biliary atresia. In many centers, a nuclear medicine scan using technetium 99m IDA (DISIDA), performed after pretreatment of the patient with phenobarbital, has proven to be an accurate and reliable study.
If radionuclide appears in the intestine, there is patency of the biliary tree, and the diagnosis of biliary atresia is excluded. If radionuclide is concentrated by the liver but not excreted despite treatment with phenobarbital, and the metabolic screen, particularly $\alpha_1$-antitrypsin determination, is normal, the presumptive diagnosis is biliary atresia. A percutaneous liver biopsy might potentially distinguish between biliary atresia and other sources of jaundice such as neonatal hepatitis. When these tests point to or cannot exclude the diagnosis of biliary atresia, surgical exploration is warranted. At surgery, a cholangiogram may be performed if possible, using the gallbladder as a point of access. This may be performed using a laparoscope. The cholangiogram demonstrates the anatomy of the biliary tree, determines whether extrahepatic bile duct atresia is present, and evaluates whether there is distal bile flow into the duodenum. The cholangiogram may demonstrate hypoplasia of the extrahepatic biliary system. This condition is associated with hepatic parenchymal disorders that cause severe intrahepatic cholestasis, including $\alpha_1$-antitrypsin deficiency and biliary hypoplasia (Alagille’s syndrome). Alternatively, a cursory assessment of the extrahepatic biliary tree may clearly delineate the atresia.

**Inspissated Bile Syndrome.** This term is applied to patients with normal biliary tracts who have persistent obstructive jaundice. Increased viscosity of bile and obstruction of the canaliculi are implicated as causes. The condition has been seen in infants receiving parenteral nutrition, but it is also encountered in conditions associated with hemolysis, or in cystic fibrosis. In some instances, no etiologic factors can be defined. Neonatal hepatitis may present in a similar fashion to biliary atresia. This disease is characterized by persistent jaundice due to acquired biliary inflammation without obliteration of the bile ducts. There may be a viral etiology, and the disease is usually self-limited. In this case, cholangiography is both diagnostic and therapeutic.

**Treatment.** If the diagnosis of biliary atresia is confirmed intraoperatively, then surgical treatment is undertaken at the same setting. Currently, first-line therapy consists of creation of a hepatopancreatic duct fistula, as described by Kasai. The purpose of this procedure is to promote bile flow into the intestine. The procedure is based on Kasai’s observation that the fibrous tissue at the porta hepatis invests microscopically patent biliary ductules that, in turn, communicate with the intrahepatic ductal system (Fig. 39-26). Transecting this fibrous tissue at the portal plate, invariably encountered cephalad to the bifurcating portal vein, opens these channels and establishes bile flow into a surgically constructed intestinal conduit, usually a Roux-en-Y limb of jejunum (Fig. 39-27). Some authors believe that an intussuscepted antireflux valve is useful in preventing retrograde bile reflux, although the data suggest that it does not impact outcome. A liver biopsy is performed at the time of surgery to determine the degree of hepatic fibrosis that is present. The diameter of bile ducts at the portal plate is predictive of likelihood of long-term success of biliary drainage through the portoenterostomy. Numerous studies also suggest that the likelihood of surgical success is inversely related to the age at the time of portoenterostomy. Infants treated prior to 60 days of life are more likely to achieve successful and long-term biliary drainage than older infants. Although the outlook is less favorable for patients after the 12th week, it is reasonable to proceed with surgery even beyond this time point, as the alternative is certain liver failure. It is noteworthy that a significant number of patients have had favorable outcomes after undergoing portoenterostomy despite advanced age at time of diagnosis.

Bile drainage is anticipated when the operation is carried out early; however, bile flow does not necessarily imply cure. Approximately one-third of patients remain symptom free after portoenterostomy, the remainder require liver transplantation due to progressive liver failure. Independent risk factors that predict failure of the procedure include bridging liver fibrosis at the time of surgery and postoperative cholangitic episodes. A review of the data of the Japanese Biliary Atresia Registry (JBAR), which

![Figure 39-26. Operative photograph showing Kasai portoenterostomy. Arrows denote the site of the anastomosis. Note the engorged liver.](image)

![Figure 39-27. Schematic illustration of the Kasai portoenterostomy for biliary atresia. An isolated limb of jejunum is brought to the porta hepatitis and anastomosed to the transected ducts at the liver plate.](image)
includes the results of 1381 patients, showed that the 10-year survival rate was 53% without transplantation, and 66.7% with transplantation. A common postoperative complication is cholangitis. There is no effective strategy to completely eliminate this complication, and the effectiveness of long-term prophylactic antibiotics has not been fully resolved. The Childhood Liver Research and Education Network (ChiLDREn, formerly the Biliary Atresia Research Consortium) is an active consortium of 15 children’s hospitals in the United States, funded by the National Institutes of Health (NIH) that studies rare cholestatic liver diseases of infants and children (http://childrennetwork.org). An NIH-funded, randomized, double-blinded, placebo-controlled trial designed to determine if adjuvant steroids improve outcome of infants undergoing Kasai portoenterostomy has been completed. This trial showed that among infants with biliary atresia who have undergone hepatopancreatorenostomy, high-dose steroid therapy following surgery did not result in statistically significant treatment differences in bile drainage at 6 months, although a small clinical benefit could not be excluded. Steroid treatment was associated with earlier onset of serious adverse events in children with biliary atresia.

Previous authors have published merits of revising the portoenterostomy in select patients if drainage of bile stops. Recently, Bondoc et al reported on their experience with revision of portoenterostomies. Specifically, the authors reported on 183 patients who underwent Kasai portoenterostomy for biliary atresia, of which 24 underwent revision for recurrence of nondrainage after successful bypass. Of the patients who underwent revision for nondrainage, 75% ultimately achieved drainage after the second procedure, of which nearly 50% survived long term with their native livers. The authors conclude that in selected patients in which bile flow was established following the Kasai procedure and then lost, revision of the portoenterostomy is a reasonable treatment option with good success.

**Choledochal Cyst**

**Classification.** The term *choledochal cyst* refers to a spectrum of congenital biliary tract disorders that were previously grouped under the name *idiopathic dilation of the common bile duct*. After the classification system proposed by Alonso-Lej, five types of choledochal cyst are described. Type I cyst is characterized by fusiform dilatation of the bile duct. This is the most common type and is found in 80% to 90% of cases. Type II choledochal cysts appear as an isolated diverticulum protruding from the wall of the common bile duct. The cyst may be joined to the common bile duct by a narrow stalk. Type III choledochal cysts arise from the intraduodenal portion of the common bile duct and are also known as choledochoceles. Type IVA cysts consist of multiple dilatations of the intrahepatic and extrahepatic bile ducts. Type IVB choledochal cysts are multiple dilatations involving only the extrahepatic bile ducts. Type V (Carolí’s disease) consists of multiple dilatations limited to the intrahepatic bile ducts.

Choledochal cyst is most appropriately considered the predominant feature in a constellation of pathologic abnormalities that can occur within the pancreato-biliary system. Frequently associated with choledochal cyst is an anomalous junction of the pancreatic and common bile ducts. The etiology of choledochal cyst is controversial. Babbit proposed an abnormal pancreatic and biliary duct junction, with the formation of a “common channel” into which pancreatic enzymes are secreted. This process results in weakening of the bile duct wall by gradual enzymatic destruction, leading to dilatation, inflammation, and finally cyst formation. Not all patients with choledochal cyst demonstrate an anatomic common channel, which raises questions regarding the accuracy of this model.

**Clinical Presentation.** Choledochal cyst is more common in females than in males (4:1). Typically, these present in children beyond the toddler age group. The classic symptom triad consists of abdominal pain, mass, and jaundice. However, this complex is actually encountered in fewer than half of the patients. The more usual presentation is that of episodic abdominal pain, often recurring over the course of months or years, and generally associated with only minimal jaundice that may escape detection. If left undiagnosed, patients may develop cholangitis or pancreatitis. Cholangitis may lead to the development of cirrhosis and portal hypertension. Choledochal cyst can present in the newborn period, where the symptoms are very similar to those of biliary atresia. Often neonates will have an abdominal mass at presentation.

**Diagnosis.** Choledochal cyst is frequently diagnosed in the fetus at a screening prenatal US. In the older child or adolescent, abdominal US may reveal a cystic structure arising from the biliary tree. CT will confirm the diagnosis. These studies will demonstrate the dimensions of the cyst and define its relationship to the vascular structures in the porta hepatis, as well as the intrahepatic ductal configuration. Endoscopic retrograde cholangiopancreatography (ERCP) is reserved for patients in whom confusion remains after evaluation by less invasive imaging modalities. Magnetic resonance cholangiopancreatography may provide a more detailed depiction of the anatomy of the cyst and its relationship to the bifurcation of the hepatic ducts and into the pancreas.

**Treatment.** The cyst wall is composed of fibrous tissue and is devoid of mucosal lining. As a result, the treatment of choledochal cyst is surgical excision followed by biliary-enteric reconstruction. There is no role for internal drainage by cystenterostomy, which leaves the cyst wall intact and leads to the inevitable development of cholangitis. Rarely, choledochal cyst can lead to the development of a biliary tract malignancy. This provides a further rationale for complete cyst excision.

Resection of the cyst may be performed via open or laparoscopic approach, and where possible, requires circumferential dissection. The posterior plane between the cyst and portal vein must be carefully dissected to accomplish removal. The pancreatic duct, which may enter the distal cyst, is vulnerable to injury during distal cyst excision but can be avoided by avoiding entry into the pancreatic parenchyma. In cases were the degree of pericystic inflammation is dense, it may be unsafe to attempt complete cyst removal. In this instance, it is reasonable to dissect within the posterior wall of the cyst, which allows the inner lining of the back wall to be dissected free from the outer layer that directly overlies the portal vascular structures. The lateral and anterior cyst, as well as the internal aspect of the back wall, is removed, yet the outer posterior wall remains behind. Cyst excision is accomplished, and the proximal bile duct is anastomosed to the intestinal tract typically via a Roux-en-Y limb of jejunum. More recently, laparoscopic-assisted resections of choledochal cysts have been described. In these cases, the end-to-side jejunoojejunostomy is performed extracorporeally, but the remainder of the procedure is completed utilizing minimally invasive techniques.

The prognosis for children who have undergone complete excision of choledochal cyst is excellent. Complications include anastomotic stricture, cholangitis, and intrahepatic stone
DEFORMITIES OF THE ABDOMINAL WALL

Embryology of the Abdominal Wall

The abdominal wall is formed by four separate embryologic folds: cephalic, caudal, right, and left lateral folds. Each of these is composed of somatic and splanchnic layers and develops toward the anterior center portion of the coelomic cavity, joining to form a large umbilical ring that surrounds the two umbilical arteries, the vein, and the yolk sac or omphalomesenteric duct. These structures are covered by an outer layer of amnion, and the entire unit composes the umbilical cord. Between the 5th and tenth weeks of fetal development, the intestinal tract undergoes rapid growth outside the abdominal cavity within the proximal portion of the umbilical cord. As development is completed, the intestine gradually returns to the abdominal cavity. Contraction of the umbilical ring completes the process of abdominal wall formation.

Failure of the cephalic fold to close results in sternal deformities such as congenital absence of the sternum. Failure of the caudal fold to close results in exstrophy of the bladder and, in more extreme cases, exstrophy of the cloaca. Interruption of central migration of the lateral folds results in omphalocele. Gastrochisis, originally thought to be a variant of omphalocele, possibly results from a fetal accident in the form of intrauterine rupture of a hernia of the umbilical cord, although other hypotheses have been advanced.

Umbilical Hernia

Failure of the umbilical ring to close results in a central defect in the linea alba. The resulting umbilical hernia is covered by normal umbilical skin and subcutaneous tissue, but the fascial defect allows protrusion of abdominal contents. Hernias less than a centimeter in size at the time of birth usually will close spontaneously by 4 to 5 years of life and in most cases should not undergo early repair. Sometimes the hernia is large enough that the protrusion allows protrusion of abdominal contents. Hernias less than a centimeter in size at the time of birth usually will close spontaneously by 4 to 5 years of life and in most cases should not undergo early repair. Sometimes the hernia is large enough that the protrusion is disfiguring and disturbing to both the child and the family. In such circumstances, early repair may be advisable (Fig. 39-28).

Umbilical hernias are generally asymptomatic protrusions of the abdominal wall. They are generally noted by parents or physicians shortly after birth. All families of patients with umbilical hernia should be counseled about signs of incarceration, which is rare in umbilical hernias and more common in smaller (1 cm or less) rather than larger defects. Incarceration presents with abdominal pain, bilious emesis, and a tender, hard mass protruding from the umbilicus. This constellation of symptoms mandates immediate exploration and repair of the hernia to avoid strangulation. More commonly, the child is asymptomatic and treatment is governed by the size of the defect, the age of the patient, and the concern that the child and family have regarding the cosmetic appearance of the abdomen. When the defect is small and spontaneous closure is likely, most surgeons will delay surgical correction until 5 years of age. If closure does not occur by this time or a younger child has a very large or symptomatic hernia, it is reasonable to proceed to repair.

Repair of uncomplicated umbilical hernia is performed under general anesthesia as an outpatient procedure. A small curving incision that fits into the skin crease of the umbilicus is made, and the sac is dissected free from the overlying skin. The fascial defect is repaired with permanent or long-lasting absorbable, interrupted sutures that are placed in a transverse plane. The skin is closed using subcuticular sutures. The postoperative recovery is typically uneventful and recurrence is rare, but it is more common in children with elevated intraabdominal pressures, such as those with a VP shunt.

Patent Urachus

During the development of the coelomic cavity, there is free communication between the urinary bladder and the abdominal wall through the urachus, which exits adjacent to the omphalomesenteric duct. Persistence of this tract results in a communication between the bladder and the umbilicus. The first sign of a patent urachus is moisture or urine flow from the umbilicus. Recurrent urinary tract infection can result. The urachus may be partially obliterated, with a remnant beneath the umbilicus in the extraperitoneal position as an isolated cyst that may be identified by US. A urachal cyst usually presents as an inflammatory mass inferior to the umbilicus. Initial treatment is drainage of the infected cyst followed by cyst excision as a separate procedure once the inflammation has resolved.

In the child with a persistently draining umbilicus, a diagnosis of patent urachus should be considered. The differential diagnosis includes an umbilical granuloma, which generally responds to local application of silver nitrate. The diagnosis of patent urachus is confirmed by umbilical exploration. The urachal tract is excised and the bladder is closed with an absorbable suture. A patent vitelline duct may also present with umbilical drainage. In this circumstance, there is a communication with the small intestine, often at the site of a Meckel’s diverticulum. Treatment includes umbilical exploration with resection of the duct remnant (Fig. 39-29).

Omphalocele

Presentation. Omphalocele refers to a congenital defect of the abdominal wall in which the bowel and solid viscera are covered by peritoneum and amniotic membrane (Fig. 39-30). The umbilical cord inserts into the sac. Omphalocele can vary from a small defect with intestinal contents to giant omphalocele in which the abdominal wall defect measures 4 cm or more in diameter and contains liver. The overall incidence is approximately 1 in 5000
live births, with 1 in 10,000 that are giant omphaloceles. Omphalocele occurs in association with special syndromes such as extrophy of the cloaca (vesicointestinal fissure), the Beckwith-Wiedemann constellation of anomalies (macroglossia, macrosomia, hypoglycemia, and visceromegaly and omphalocele) and Cantrell’s Pentalogy (lower thoracic wall malformations [cleft sternum], ectopia cordis, epigastric omphalocele, anterior midline diaphragmatic hernia and cardiac anomalies). There is a 60% to 70% incidence of associated anomalies, especially cardiac (20–40% of cases) and chromosomal abnormalities. Chromosomal anomalies are more common in children with smaller defects. Omphalocele is associated with prematurity (10–50% of cases) and intrauterine growth restriction (20% of cases).

**Treatment.** Immediate treatment of an infant with omphalocele consists of attending to the vital signs and maintaining the body temperature. A blood glucose should be evaluated because of the association with Beckwith-Wiedemann. The omphalocele should be covered to reduce fluid loss, but moist dressings may result in heat loss and are not indicated. No pressure should be placed on the omphalocele sac in an effort to reduce its contents because this maneuver may increase the risk of rupture of the sac or may interfere with abdominal venous return. Prophylactic broad-spectrum antibiotics should be administered in case of rupture. The subsequent treatment and outcome is determined by the size of the omphalocele. In general terms, small to medium-sized defects have a significantly better prognosis than extremely large defects in which the liver is present. In these cases, not only is the management of the abdominal wall defect a significant challenge, but these patients often have concomitant pulmonary insufficiency that can lead to significant morbidity and mortality. If possible, and if the pulmonary status will permit it, a primary repair of the omphalocele should be undertaken. This involves resection of the omphalocele membrane and closure of the fascia. A layer of prosthetic material may be required to achieve closure. In infants with a giant omphalocele, the defect cannot be closed primarily because there is not adequate intraperitoneal domain to reduce the viscera (see Fig. 39-30). Some infants may have associated congenital anomalies that complicate surgical repair, and because cardiac anomalies are common, an echocardiogram should be obtained prior to any procedure. If repair is contraindicated, such as with a very large defect, a nonoperative approach can be used. The omphalocele sac can be treated with topical treatments, which serve to harden the sac to allow more protective coverage where muscle and skin cannot be used given the large defect. Various authors describe success with iodine-containing solutions, silver sulfadiazine, or saline, and some surgeons rotate these solutions because of the impact of iodine on the thyroid and the difficulty of cleaning off all of the silver sulfadiazine and its association with leukopenia. It typically takes 2 to 3 months before reepithelialization occurs. In the past, mercury compounds were used, but they have been discontinued because of associated systemic toxicity. After epithelialization has occurred, attempts should be made to achieve closure of the anterior abdominal wall but may be delayed by associated pulmonary insufficiency. Such procedures typically require complex measures to achieve skin closure, including the use of biosynthetic materials or component separation. In cases of giant omphalocele, prolonged hospitalization is typical. If the base is very narrow—which can occur even for babies with very large omphaloceles—it may be wise to open the base in order to allow the abdominal contents and the liver to reenter the abdominal cavity, and thereby achieve abdominal domain. This approach will, by necessity, require sewing in some synthetic material in order to achieve fascial closure, and prolonged hospitalization will be required to allow for skin coverage to occur. These patients require high amounts of caloric support, given the major demands for healing.

**Gastroschisis**

**Presentation.** Gastroschisis represents a congenital anomaly characterized by a defect in the anterior abdominal wall through which the intestinal contents freely protrude. Unlike omphalocele, there is no overlying sac, and the size of the defect is usually <4 cm. The abdominal wall defect is located at the junction of the umbilicus and normal skin, and is almost always to the right of the umbilicus (Fig. 39-31). The umbilicus becomes partly detached, allowing free communication with the...
abdominal cavity. The appearance of the bowel provides some information with respect to the in-utero timing of the defect. The intestine may be normal in appearance, suggesting that the rupture occurred relatively late during the pregnancy. More commonly, however, the intestine is thick, edematous, discolored, and covered with exudate, implying a more longstanding process. Progression to full enteral feeding is usually delayed, with diminished motility that may be related to these changes.

Unlike infants born with omphalocele, associated anomalies are not usually seen with gastroschisis except for a 10% rate of intestinal atresia. This defect can readily be diagnosed on prenatal US (Fig. 39-32). There is no advantage to performing a cesarean section instead of a vaginal delivery. In a decade long retrospective review, early deliver did not affect the thickness of bowel peel, yet patients delivered before 36 weeks had significantly longer length of stay in the hospital and time to enteral feeds. Based upon these findings, it is thought that fetal well-being should be the primary determinant of delivery for gastroschisis.

Treatment. All infants born with gastroschisis require urgent surgical treatment. Of equal importance, these infants require vigorous fluid resuscitation in the range of 160 to 190 cc/kg per day to replace significant evaporative fluid losses. In many instances, the intestine can be returned to the abdominal cavity, and a primary surgical closure of the abdominal wall is performed. Some surgeons believe that they facilitate primary closure with mechanical stretching of the abdominal wall, thorough orogastric suctioning with foregut decompression, rectal irrigation, and evacuation of meconium. Care must be taken to prevent markedly increased abdominal pressure during the reduction, which will lead to compression of the inferior vena cava, respiratory embarrassment, and abdominal compartment syndrome. To avoid this complication, it is helpful to monitor the bladder or airway pressures during reduction. In infants whose intestine has become thickened and edematous, it may be impossible to reduce the bowel into the peritoneal cavity in the immediate postnatal period. Under such circumstances, a plastic spring-loaded silo can be placed onto the bowel and secured beneath the fascia or a sutured silastic silo constructed. The silo covers the bowel and allows for graduated reduction on a daily basis as the edema in the bowel wall decreases (Fig. 39-33). It is important to ensure that the silo-fascia junction does not become a constricting point or “funnel,” in which case the intestine will be injured upon return to the peritoneum. In this case, the fascial opening must be enlarged. Surgical closure can usually be accomplished within approximately 1 to 2 weeks. A prosthetic piece of material may be required to bring the edges of the fascia together. If an atresia is noted at the time of closure, it is prudent to reduce the bowel at the first operation and return after several weeks once the edema has resolved to correct the atresia. Intestinal function does not typically return for several weeks in patients with gastroschisis. This is especially true if the bowel is thickened and edematous. As a result, these patients will require central line placement and institution of total parenteral nutrition in order to grow. Feeding advancement should be slow and typically requires weeks to arrive at full enteral nutrition.
Prune-Belly Syndrome

Clinical Presentation. Prune-belly syndrome refers to a disorder that is characterized by extremely lax lower abdominal musculature, dilated urinary tract including the bladder, and bilateral undescended testes (Fig. 39-34). The term prune-belly syndrome appropriately describes the wrinkled appearance of the anterior abdominal wall that characterizes these patients. Prune-belly syndrome is also known as Eagle-Barrett syndrome as well as the triad syndrome because of the three major manifestations. The incidence is significantly higher in males. Patients manifest a variety of comorbidities. The most significant is pulmonary hypoplasia, which can be unsurvivable in the most severe cases. Skeletal abnormalities include dislocation or dysplasia of the hip and pectus excavatum.

The major genitourinary manifestation in prune-belly syndrome is ureteral dilation. The ureters are typically long and tortuous and become more dilated distally. Ureteric obstruction is rarely present, and the dilation may be caused by decreased smooth muscle and increased collagen in the ureters. Additionally, up to eighty percent of these patients will have some degree of vesicoureteral reflux, which can predispose to urinary tract infection. Despite the marked dilatation of the urinary tract, most children with prune-belly syndrome have adequate renal parenchyma for growth and development. Factors associated with the development of long-term renal failure include the presence of abnormal kidneys on US or renal scan and persistent pyelonephritis.

Treatment. Despite the ureteric dilation, there is currently no role for ureteric surgery unless an area of obstruction develops. The testes are invariably intraabdominal, and bilateral orchidopexy can be performed in conjunction with abdominal wall reconstruction at 6 to 12 months of age. Despite orchidopexy, fertility in a boy with prune-belly syndrome is unlikely as spermatogenesis over time is insufficient. Deficiencies in the production of prostatic fluid and a predisposition to retrograde ejaculation contribute to infertility. Abdominal wall repair is accomplished through an abdominoplasty, which typically requires a transverse incision in the lower abdomen extending into the flanks.

Inguinal Hernia

An understanding of the management of pediatric inguinal hernias is a central component of modern pediatric surgical practice. Inguinal hernia repair represents one of the most common operations performed in children. The presence of an inguinal hernia in a child is an indication for surgical repair. The operation is termed a hernioplasty because it involves closing off the patent processus vaginalis. This is to be contrasted with the hernioplasty that is performed in adults, which requires a reconstruction of the inguinal floor.

Embryology. In order to understand how to diagnose and treat inguinal hernias in children, it is critical to understand their embryologic origin. It is very useful to describe these events to the parents, who often are under the misconception that the hernia was somehow caused by their inability to console their crying child, or the child’s high activity level. Inguinal hernia results from a failure of closure of the processus vaginalis; a finger-like projection of the peritoneum that accompanies the testicle as it descends into the scrotum. Closure of the processus vaginalis normally occurs a few months prior to birth. This explains the high incidence of inguinal hernias in premature infants. When the processes vaginalis remains completely patent, a communication persists between the peritoneal cavity and the groin, resulting in a hernia. Partial closure can result in entrapped fluid, which results in the presence of a hydrocele. A communicating hydrocele refers to a hydrocele that is in communication with the peritoneal cavity and can therefore be thought of as a hernia. Using the classification system that is typically applied to adult hernias, all congenital hernias in children are by definition indirect inguinal hernias. Children also present with direct inguinal and femoral hernias, although these are much less common.

Clinical Manifestation. Inguinal hernias occur more commonly in males than females (10:1) and are more common on the right side than the left. Infants are at high risk for incarceration of an inguinal hernia because of the narrow inguinal ring. Patients most commonly present with a groin bulge that is noticed by the parents as they change the diaper (Fig. 39-35).
Older children may notice the bulge themselves. On examination, the cord on the affected side will be thicker, and pressure on the lower abdomen usually will display the hernia on the affected side. The presence of an incarcerated hernia is manifested by a firm bulge that does not spontaneously resolve and may be associated with fussiness and irritability in the child. The infant that has a strangulated inguinal hernia will manifest an edematous, tender bulge in the groin, occasionally with overlying skin changes. The child will eventually develop intestinal obstruction, peritonitis, and systemic toxicity.

Usually an incarcerated hernia can be reduced. Occasionally this may require light sedation. Gentle pressure is applied on the sac from below in the direction of the internal inguinal ring. Following reduction of the incarcerated hernia, the child may be admitted for observation, and herniorrhaphy is performed within the next 24 hours to prevent recurrent incarceration. Alternatively, the child may be scheduled for surgery at the next available time slot. If the hernia cannot be reduced, or if evidence of strangulation is present, emergency operation is necessary. This may require a laparotomy and bowel resection.

When the diagnosis of inguinal hernia is made in an otherwise normal child, operative repair should be planned. Spontaneous resolution does not occur, and therefore a nonoperative approach cannot ever be justified. An inguinal hernia in a female infant or child frequently contains an ovary rather than intestine. Although the gonad usually can be reduced into the abdomen by gentle pressure, it often prolapses in and out until surgical repair is carried out. In some patients, the ovary and fallopian tube constitute one wall of the hernial sac (sliding hernia), and in these patients, the ovary can be reduced effectively only at the time of operation. If the ovary is irreducible, prompt hernia repair is indicated to prevent ovarian torsion or strangulation.

When a hydrocele is diagnosed in infancy and there is no evidence of a hernia, observation is proper therapy until the child is older than 12 months. If the hydrocele has not disappeared by 12 months, invariably there is a patent processus vaginalis, and operative hydrocelectomy with excision of the processus vaginalis is indicated. When the first signs of a hydrocele are seen after 12 months of age, the patient should undergo elective hydrocelectomy, which in a child is always performed through a groin incision. Aspiration of hydroceles is discouraged because almost all without a patent processus vaginalis will resorb spontaneously and those with a communication to the peritoneum will recur and require operative repair eventually. Transillumination as a method to distinguish between hydrocele and hernia is nonspecific. A noncommunicating hydrocele is better identified by palpation of a nonreducible oval structure that appears to have a blunt end below the external ring, indicating an isolated fluid collection without a patent connection to the peritoneum. Surgical Repair. The repair of a pediatric inguinal hernia can be extremely challenging, particularly in the premature child with incarceration. A small incision is made in a skin crease in the groin directly over the internal inguinal ring. Scarpa’s fascia is seen and divided. The external oblique muscle is dissected free from overlying tissue, and the location of the external ring is confirmed. The external oblique aponeurosis is then opened along the direction of the external oblique fibers over the inguinal canal. The undersurface of the external oblique is then cleared from surrounding tissue. The cremasteric fibers are separated from the cord structures and hernia sac, and these are then elevated into the wound. Care is taken not to grasp the vas deferens. The hernia sac is then dissected up to the internal ring and doubly suture ligated. The distal part of the hernia sac is opened widely to drain any hydrocele fluid. When the hernia is very large and the patient very small, tightening of the internal inguinal ring or even formal repair of the inguinal floor may be necessary, although the vast majority of children do not require any treatment beyond high ligation of the hernia sac.

Controversy exists regarding the role for exploration of an asymptomatic opposite side in a child with an inguinal hernia. Several reports indicate that frequency of a patent processus vaginalis on the side opposite the obvious hernia is approximately 30%, although this figure decreases with increasing age of the child. Management options include never exploring the opposite side, to exploring only under certain conditions such as in premature infants or in patients in whom incarceration is present. The opposite side may readily be explored laparoscopically. To do so, a blunt 3-mm trochar is placed into the hernia sac of the affected side. The abdominal cavity is insufflated, and the 2.7-mm 70° camera is placed through the trochar such that the opposite side is visualized. The status of the processes vaginalis on the opposite side can be visualized. However, the presence of a patent processus vaginalis by laparoscopy does not always imply the presence of a hernia.

There has been quite widespread adoption of laparoscopic approach in the management of inguinal hernias in children, especially those under the age of 2 years. This technique requires insufflation through the umbilicus and the placement of an extraperitoneal suture to ligate the hernia sac. Proponents of this procedure emphasize the fact that no groin incision is used, so there is a decreased chance of injuring cord structures, and that visualization of the contralateral side is achieved immediately. The long-term results of this technique have been quite excellent.

Inguinal hernias in children recur in less than 1% of patients, and recurrences usually result from missed hernia sacs at the first procedure, a direct hernia, or a missed femoral hernia. All children should have local anesthetic administered either by caudal injection or by direct injection into the wound. Spinal anesthesia in preterm infant decreases the risk of postoperative apnea when compared with general anesthesia.

**GENITALIA**

**Undescended testis**

**Embryology.** The term undescended testicle (cryptorchidism) refers to the interruption of the normal descent of the testis into the scrotum. The testicle may reside in the retroperitoneum, in the internal inguinal ring, in the inguinal canal, or even at the external ring. The testicle begins as a thickening on the urogenital ridge in the fifth to sixth week of embryologic life. In the seventh and eighth months, the testicle descends along the inguinal canal into the upper scrotum, and with its progress the processus vaginalis is formed and pulled along with the migrating testicle. At birth, approximately 95% of infants have the testicle normally positioned in the scrotum.

A distinction should be made between an undescended testicle and an ectopic testicle. An ectopic testis, by definition, is one that has passed through the external ring in the normal pathway and then has come to rest in an abnormal location overlying either the rectus abdominis or external oblique muscle, or the soft tissue of the medial thigh, or behind the scrotum in the perineum. A congenitally absent testicle results from failure of normal development or an intrauterine accident leading to loss of blood supply to the developing testicle.
Clinical Presentation. The incidence of undescended testes is approximately 30% in preterm infants, and 1% to 3% at term. For diagnosis, the child should be examined in the supine position, where visual inspection may reveal a hypoplastic or poorly rugated scrotum. Usually a unilateral undescended testicle can be palpated in the inguinal canal or in the upper scrotum. Occasionally, the testicle will be difficult or impossible to palpate, indicating either an abdominal testicle or congenital absence of the gonad. If the testicle is not palpable in the supine position, the child should be examined with his legs crossed while seated. This maneuver diminishes the cremasteric reflex and facilitates identification of the location of the testicle. If there is uncertainty regarding location of a testis, repeated evaluations over time may be helpful.

It is now established that cryptorchid testes demonstrate an increased predisposition to malignant degeneration. In addition, fertility is decreased when the testicle is not in the scrotum. For these reasons, surgical placement of the testicle in the scrotum (orchidopexy) is indicated. It should be emphasized that this procedure does improve the fertility potential, although it is never normal. Similarly, the testicle is still at risk of malignant change, although its location in the scrotum facilitates potentially earlier detection of a testicular malignancy. Other reasons to consider orchidopexy include the risk of trauma to the testicle located at the pubic tubercle and incidence of torsion, as well as the psychological impact of an empty scrotum in a developing male. The reason for malignant degeneration is not established, but the evidence points to an inherent abnormality of the testicle that predisposes it to incomplete descent and malignancy rather than malignancy as a result of an abnormal environment.

Treatment. Males with bilateral undescended testicles are often infertile. When the testicle is not present within the scrotum, it is subjected to a higher temperature, resulting in decreased spermatogenesis. Mengel and coworkers studied 515 undescended testicles by histology and demonstrated reduced spermatogonia after 2 years of age. It is now recommended that the undescended testicle be surgically repositioned by 1 year of age. Despite orchidopexy, the incidence of infertility is approximately two times higher in men with unilateral orchidopexy compared to men with normal testicular descent.

The use of chorionic gonadotropin occasionally may be effective in patients with bilateral undescended testes, suggesting that these patients are more apt to have a hormone insufficiency than children with unilateral undescended testicle. The combination of micro-penis and bilateral undescended testes is an indication for hormonal evaluation and testosterone replacement if indicated. If there is no testicular descent after a month of endocrine therapy, operative correction should be undertaken. A child with unilateral cryptorchidism should have surgical correction of the problem. The operation is typically performed through a combined groin and scrotal incision. The cord vessels are fully mobilized, and the testicle is placed in a dartos pouch within the scrotum. An inguinal hernia often accompanies a cryptorchid testis. This should be repaired at the time of orchidopexy.

Patients with a nonpalpable testicle present a challenge in management. The current approach involves laparoscopy to identify the location of the testicle. If the spermatic cord is found to traverse the internal ring or the testis is found at the ring and can be delivered into the scrotum, a groin incision is made and an orchidopexy is performed. If an abdominal testis is identified that is too far to reach the scrotum, a two-staged Fowler-Stephens approach is used. In the first stage, the testicular vessels are clipped laparoscopically, which promotes the development of new blood vessels along the vas deferens. Several months later, the second stage is performed during which the testis is mobilized laparoscopically along with a swath of peritoneum with collateralized blood supply along the vas. Preservation of the gubernacular attachments with its collaterals to the testicle may confer improved testicular survival following orchidopexy in over 90%. It is, nonetheless, preferable to preserve the testicular vessels whenever possible and complete mobilization of the testicle with its vessels intact.

Vaginal Anomalies
Surgical diseases of the vagina in children are either congenital or acquired. Congenital anomalies include a spectrum of diseases that range from simple defects (imperforate hymen) to more complex forms of vaginal atresia, including distal, proximal, and, most severe, complete. These defects are produced by abnormal development of müllerian ducts and/or urogenital sinus. The diagnosis is made most often by physical examination. Secretions into the obstructed vagina produce hydrocolpos, which may present as a large, painful abdominal mass. The anatomy may be defined using US. Pelvic magnetic resonance imaging provides the most thorough and accurate assessment of the pelvic structures. Treatment is dependent on the extent of the defect. For an imperforate hymen, division of the hymen is curative. More complex forms of vaginal atresia require mobilization of the vaginal remnants and creation of an anastomosis at the perineum. Laparoscopy can be extremely useful, both in mobilizing the vagina, in draining hydrocolpos, and in evaluating the internal genitalia. Complete vaginal atresia requires the construction of skin flaps or the creation of a neovagina using a segment of colon.

The most common acquired disorder of the vagina is the straddle injury. This often occurs as young girls fall on blunt objects which cause a direct injury to the perineum. Typical manifestations include vaginal bleeding and inability to void. Unless the injury is extremely superficial, patients should be examined in the operating room where the lighting is optimal and sedation can be administered. Examination under anesthesia is particularly important in girls who are unable to void, suggesting a possible urethral injury. Vaginal lacerations are repaired using absorbable sutures, and the proximity to the urethra should be carefully assessed. Prior to hospital discharge, it is important that girls are able to void spontaneously. In all cases of vaginal trauma, it is essential that the patient be assessed for the presence of sexual abuse. In these cases, early contact with the sexual abuse service is necessary so that the appropriate microbiologic and photographic evidence can be obtained.

Ovarian Cysts and Tumors
Pathologic Classification. Ovarian cysts and tumors may be classified as nonneoplastic or neoplastic. Nonneoplastic lesions include cysts (simple, follicular, inclusion, paraovarian, or corpus luteum), endometriosis, and inflammatory lesions. Neoplastic lesions are classified based on the three primordia that contribute to the ovary: mesenchymal components of the urogenital ridge, germinal epithelium overlying the urogenital ridge, and germ cells migrating from the yolk sac. The most common variety is germ cell tumors. Germ cell tumors are classified based on the degree of differentiation and the cellular components
involved. The least differentiated tumors are the dysgerminomas, which share features similar to the seminoma in males. Although these are malignant tumors, they are extremely sensitive to radiation and chemotherapy. The most common germ cell tumors are the teratomas, which may be mature, immature, or malignant. The degree of differentiation of the neural elements of the tumor determines the degree of immaturity. The sex cord stromal tumors arise from the mesenchymal components of the urogenital ridge. These include the granulosa-theca cell tumors and the Sertoli-Leydig cell tumors. These tumors often produce hormones that result in precocious puberty or hirsutism, respectively. Although rare, epithelial tumors do occur in children. These include serous and mucinous cystadenomas.

**Clinical Presentation.** Children with ovarian lesions usually present with abdominal pain. Other signs and symptoms include a palpable abdominal mass, evidence of urinary obstruction, symptoms of bowel obstruction, and endocrine imbalance. The surgical approach depends on the appearance of the mass at operation (i.e., whether it is benign-appearing or is suspicious for malignancy). In the case of a simple ovarian cyst, surgery depends on the size of the cyst and the degree of symptoms it causes. In general, large cysts (over 4–5 cm) in size should be resected, as they are unlikely to resolve, may be at risk of torsion, and may mask an underlying malignancy. Resection may be performed laparoscopically, and ovarian tissue should be spared in all cases.

**Surgical Management.** For ovarian lesions that appear malignant, it is important to obtain tumor markers including α-fetoprotein (teratomas), β-hCG (dysgerminoma), β-human chorionic gonadotropin (choriocarcinoma), and CA-125 (epithelial tumors). Although the diagnostic sensitivity of these markers is not always reliable, they provide material for postoperative follow-up and indicate the response to therapy. When malignancy is suspected, the patient should undergo a formal cancer operation. This procedure is performed through either a midline incision or a Pfannenstie approach. Ascites and peritoneal washings should be collected for cytologic study. The liver and diaphragm are inspected carefully for metastatic disease. An omentectomy is performed if there is any evidence of tumor present. Pelvic and para-aortic lymph nodes are biopsied, and the primary tumor is resected completely. Finally, the contralateral ovary is carefully inspected, and if a lesion is seen, it should be biopsied. Dysgerminomas and epithelial tumors may be bilateral in up to 15% of cases. The surgical approach for a benign lesion of the ovary should include preservation of the ipsilateral fallopian tube and preservation of the noninvolved ovary.

**Ovarian Cysts in the Newborn.** Ovarian cysts may be detected by prenatal US. The approach to lesions less than 4 cm should include serial US evaluation every 2 months or so as many of these lesions will resolve spontaneously. Consideration should be given to laparoscopic excision of cysts larger than 4 cm to avoid the risks of ovarian torsion or development of abdominal symptoms. For smaller lesions, resolution occurs by approximately 6 months of age. A laparoscopic approach is preferable in these cases. By contrast, complex cysts of any size require surgical intervention at presentation to exclude the possibility of malignancy.

**Ambiguous Genitalia**

**Embryology.** Normal sexual differentiation occurs in the sixth fetal week. In every fetus, wolffian (male) and müllerian (female) ducts are present until the onset of sexual differentiation. Normal sexual differentiation is directed by the sex determining region of the Y chromosome (SRY). This is located on the distal end of the short arm of the Y chromosome. SRY provides a genetic switch that initiates gonadal differentiation in the mammalian urogenital ridge. Secretion of Müllerian-inhibiting substance (MIS) by the Sertoli cells of the seminiferous tubules results in regression of the müllerian duct, the anlage of the uterus, Fallopian tubes, and the upper vagina. The result of MIS secretion therefore is a phenotypic male. In the absence of SRY in the Y chromosome, MIS is not produced, and the müllerian duct derivatives are preserved. Thus, the female phenotype prevails.

In order for the male phenotype to develop, the embryo must have a Y chromosome, the SRY must be normal without point mutations or deletions, testosterone and MIS must be produced by the differentiated gonad, and the tissues must respond to these hormones. Any disruption of the orderly steps in sexual differentiation may be reflected clinically as variants of the intersex syndromes.

These may be classified as (a) true hermaphroditism (with ovarian and testicular gonadal tissue), (b) male pseudohermaphroditism (testicles only), (c) female pseudohermaphroditism (ovarian tissue only), and (d) mixed gonadal dysgenesis (usually underdeveloped or imperfectly formed gonads).

**True Hermaphroditism** This represents the rarest form of ambiguous genitalia. Patients have both normal male and female gonads, with an ovary on one side and a testis on the other. Occasionally, an ovotestis is present on one or both sides. The majority of these patients have a 46,XX karyotype. Both the testis and the testicular portion of the ovotestis should be removed.

**Male Pseudohermaphroditism** This condition occurs in infants with an XY karyotype but deficient masculinization of the external genitalia. Bilateral testes are present, but the duct structures differentiate partly as phenotypic females. The causes include inadequate testosterone production due to biosynthetic error, inability to convert testosterone to dihydrotestosterone due to 5α-reductase deficiency or deficiencies in androgen receptors. The latter disorder is termed testicular feminization syndrome. Occasionally, the diagnosis in these children is made during routine inguinal herniorrhaphy in a phenotypic female at which time testes are found. The testes should be resected due to the risk of malignant degeneration, although this should be performed only after a full discussion with the family has occurred.

**Female Pseudohermaphroditism** The most common cause of female pseudohermaphroditism is congenital adrenal hyperplasia. These children have a 46,XX karyotype but have been exposed to excessive androgens in utero. Common enzyme deficiencies include 21-hydroxylase, 11-hydroxylase, and 3β-hydroxysteroid dehydrogenase. These deficiencies result in overproduction of intermediary steroid hormones, which results in masculinization of the external genitalia of the XX fetus. These patients are unable to synthesize cortisol. In 90% of cases, deficiency of 21-hydroxylase causes adrenocorticotropic hormone (ACTH) to stimulate the secretion of excessive quantities of adrenal androgen, which masculinizes the developing female (Fig. 39-36). These infants are prone to salt loss, and require cortisol replacement. Those with mineralocorticoid deficiency also require fluorocortisone replacement.

**Mixed Gonadal Dysgenesis** This syndrome is associated with dysgenetic gonads and retained müllerian structures. The typical karyotype is mosaic, usually 45XO,46XY. A high incidence of
malignant tumors occur in the dysgenetic gonads, most commonly gonadoblastoma. Therefore, they should be removed.

Management. In the differential diagnosis of patients with intersex anomalies, the following diagnostic steps are necessary: (a) evaluation of the genetic background and family history; (b) assessment of the anatomic structures by physical examination, US, and/or chromosome studies; (c) determination of biochemical factors in serum and urine to evaluate the presence of an enzyme defect; and (d) laparoscopy for gonadal biopsy. Treatment should include correction of electrolyte and volume losses, in cases of congenital adrenal hyperplasia, and replacement of hormone deficiency. Surgical assignment of gender should never be determined at the first operation. Although historically female gender had been assigned, there is abundant and convincing evidence that raising a genotypic male as a female has devastating consequences, not only anatomically but also psychosocially. This is particularly relevant given the role of pre- and postnatal hormones on gender imprinting and identity. In general terms, surgical reconstruction should be performed after a full genetic workup and with the involvement of pediatric endocrinologists, plastic surgeons, and ethicists with expertise in gender issues. Discussion with the family also plays an important role. This approach will serve to reduce the anxiety associated with these disorders and will help to ensure the normal physical and emotional development of these patients.

PEDIATRIC MALIGNANCY

Cancer is the second leading cause of death in children after trauma and accounts for approximately 11% of all pediatric deaths in the United States. The following description will be restricted to the most commonly encountered tumors in children.

Wilms’ Tumor

Clinical Presentation. Wilms’ tumor is the most common primary malignant tumor of the kidney in children. There are approximately 500 new cases annually in the United States, and most are diagnosed between 1 and 5 years with the peak incidence at age 3. Advances in the care of patients with Wilms’ tumor has resulted in an overall cure rate of roughly 90%, even in the presence of metastatic spread. The tumor usually develops in otherwise healthy children as an asymptomatic mass in the flank or upper abdomen. Frequently, the mass is discovered by a parent while bathing or dressing the child. Other symptoms include hypertension, hematuria, obstipation, and weight loss. Occasionally the mass is discovered following blunt abdominal trauma.

Genetics of Wilms’ Tumor. Wilms’ tumor can arise from both germline and somatic mutations and can occur in the presence or absence of a family history. Nearly 97% of Wilms’ tumors are sporadic in that they occur in the absence of a heritable or congenital cause or risk factor. When a heritable risk factor is identified, the affected children often present at an earlier age, and the tumors are frequently bilateral. Most of these tumors are associated with germline mutations. It is well established that there is a genetic predisposition to Wilms’ tumor in WAGR syndrome, which consists of Wilms’ tumor, aniridia, genitourinary abnormalities, and mental retardation. In addition, there is an increased incidence of Wilms’ tumor in certain overgrowth conditions, particularly Beckwith–Wiedemann syndrome and hemihypertrophy. WAGR syndrome has been shown to result from the deletion of one copy each of the Wilms’ tumor gene, WT1, and the adjacent aniridia gene, PAX6, on chromosome 11p13. Beckwith–Wiedemann syndrome is an overgrowth syndrome that is characterized by visceromegaly, macroglossia, and hyperinsulinemic hypoglycemia. It arises from mutations at the 11p15.5 locus. There is evidence to suggest that analysis of the methylation status of several genes in the 11p15 locus could predict the individual risk to the development of Wilms’ tumor. Importantly, most patients with Wilms’ tumor do not have mutations at these genetic loci.

Surgical Treatment. Before operation, all patients suspected of having Wilms’ tumor should undergo abdominal and chest computerized tomography. These studies characterize the mass, identify the presence of metastases, and provide information on the opposite kidney (Fig. 39-37). CT scanning also indicates the presence of nephrogenic rests, which are precursor lesions to Wilms’ tumor. An abdominal US should be performed to evaluate the presence of renal vein or vena caval extension.

The management of patients with Wilms’ tumor has been carefully analyzed within the context of large studies involving thousands of patients. These studies have been coordinated by the National Wilms’ Tumor Study Group (NWTSG) in North America and the International Society of Paediatric Oncology.
(SIOP), mainly involving European countries. Significant differences in the approach to patients with Wilms’ tumor have been highlighted by these studies. NWTSG supports a strategy of surgery followed by chemotherapy in most instances, whereas the SIOP approach is to shrink the tumor using preoperative chemotherapy. There are instances where preoperative chemotherapy is supported by both groups, including the presence of bilateral involvement or inferior vena cava involvement that extends above the hepatic veins and involvement of a solitary kidney by Wilms’ tumor. The NWTSG proponents argue that preoperative therapy in other instances results in a loss of important staging information, and therefore places patients at higher risk for recurrence; alternatively, it may lead to overly aggressive treatment in some cases and greater morbidity. However, the overall survival rates are not different between the NWTSG and SIOP approaches.

The goal of surgery is complete removal of the tumor. It is crucial to avoid tumor rupture or injury to contiguous organs. A sampling of regional lymph nodes should be included, and all suspicious nodes should be sampled. Typically, a large transverse abdominal incision is made, and a transperitoneal approach is used. The opposite side is carefully inspected to ensure that there is no disease present. Although historically this involved the complete mobilization of the contralateral kidney, current evidence indicates that preoperative, high-resolution CT scanning is of sufficient accuracy for the detection of clinically significant lesions if they are present. Provided only unilateral disease is present, a radical nephroureterectomy is then performed with control of the renal pedicle as an initial step. If there is spread above the hepatic veins, an intrathoracic approach may be required. If bilateral disease is encountered, both lesions are biopsied, and chemotherapy is administered followed by a nephron-sparing procedure.

Chemotherapy. Following nephroureterectomy for Wilms’ tumor, the need for chemotherapy and/or radiation therapy are determined by the histology of the tumor and the clinical stage of the patient (Table 39-3). Essentially, patients who have disease confined to one kidney completely excised surgically receive a short course of chemotherapy and can expect a 97% 4-year survival, with tumor relapse rare after that time. Patients with more advanced disease or with unfavorable histology receive more intensive chemotherapy and radiation. Even in stage IV, high cure rates may be achieved. The survival rates are worse in the small percentage of patients considered to have unfavorable histology.

Neuroblastoma
Clinical Presentation. Neuroblastoma is the third most common pediatric malignancy and accounts for approximately 10% of all childhood cancers. The vast majority of patients have advanced disease at the time of presentation, and unlike Wilms’ tumor, in which cure is expected in the vast majority of patients, the overall survival of patients with neuroblastoma is significantly lower. Over 80% of cases present before the age of 4 years, and the peak incidence is two years of age. Neuroblastomas arise from the neural crest cells and show different levels of differentiation. The tumor originates most frequently in the adrenal glands, posterior mediastinum, neck, or pelvis but can arise in any sympathetic ganglion. The clinical presentation depends on the site of the primary and the presence of metastases.

<table>
<thead>
<tr>
<th>Table 39-3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Staging of Wilms’ tumor</strong></td>
</tr>
<tr>
<td><strong>Stage I:</strong> Tumor limited to the kidney and completely excised.</td>
</tr>
<tr>
<td><strong>Stage II:</strong> Tumor that extends beyond the kidney but is completely excised. This includes penetration of the renal capsule, invasion of the soft tissues of the renal sinus, or blood vessels within the nephrectomy specimen outside the renal parenchyma containing tumor. No residual tumor is apparent at or beyond the margins of excision.</td>
</tr>
<tr>
<td><strong>Stage III:</strong> Residual nonhematogenous tumor confined to the abdomen. Lymph nodes in the abdomen or pelvis contain tumor. Peritoneal contamination by the tumor, such as by spillage or biopsy of tumor before or during surgery. Tumor growth that has penetrated through the peritoneal surface. Implants are found on the peritoneal surfaces. Tumor extends beyond the surgical margins either microscopically or grossly. Tumor is not completely resectable because of local infiltration into vital structures. The tumor was treated with preoperative chemotherapy with or without biopsy. Tumor is removed in greater than one piece.</td>
</tr>
<tr>
<td><strong>Stage IV:</strong> Hematogenous metastases or lymph node involvement outside the abdomino-pelvic region.</td>
</tr>
<tr>
<td><strong>Stage V:</strong> Bilateral renal involvement.</td>
</tr>
<tr>
<td><strong>International Neuroblastoma Staging System</strong></td>
</tr>
<tr>
<td>Stage 1: Localized tumor with complete gross resection, with or without microscopic residual disease</td>
</tr>
<tr>
<td>Stage 2A: Localized tumor with incomplete gross excision; representative ipsilateral nonadherent lymph nodes negative for tumor</td>
</tr>
<tr>
<td>Stage 2B: Localized tumor with or without complete gross excision, with ipsilateral nonadherent lymph nodes positive for tumor. Enlarged contralateral lymph nodes must be negative microscopically</td>
</tr>
<tr>
<td>Stage 3: Unresectable unilateral tumor crossing midline, with or without regional lymph node involvement; or localized unilateral tumor with contralateral regional lymph node involvement; or midline tumor</td>
</tr>
<tr>
<td>Stage 4: Any primary tumor with dissemination to distant lymph nodes, bone, bone marrow, liver, skin, and/or other organs</td>
</tr>
<tr>
<td>Stage 4S: In infants &lt;1 year of age; localized primary tumor with dissemination limited to skin, liver, and/or bone marrow</td>
</tr>
</tbody>
</table>

*Rupture or spillage confined to the flank, including biopsy of the tumor, is no longer included in stage II and is now included in stage III.*
Two-thirds of these tumors are first noted as an asymptomatic abdominal mass. The tumor may cross the midline, and a majority of patients will already show signs of metastatic disease. Occasionally, children may experience pain from the tumor mass or from bony metastases. Proptosis and periorbital ecchymosis may occur due to the presence of retrobulbar metastasis. Because they originate in paraspinal ganglia, neuroblastosmas may invade through neural foramina and compress the spinal cord, causing muscle weakness or sensory changes. Rarely, children may have severe watery diarrhea due to the secretion of vasoactive intestinal peptide by the tumor, or with paraneoplastic neurologic findings including cerebellar ataxia or opsoclonus/myoclonus. The International Neuroblastoma Staging System and the International Neuroblastoma Risk Group Staging System are provided in Table 39-3.

**Diagnostic Evaluation.** Since these tumors derive from the sympathetic nervous system, catecholamines and their metabolites will be produced at increased levels. These include elevated levels of serum catecholamines (dopamine, norepinephrine) or urine catecholamine metabolites: vanillylmandelic acid (VMA) or homovanillic acid (HVA). Measurement of VMA and HVMA in serum and urine aids in the diagnosis and in monitoring adequacy of future treatment and recurrence. The minimum criterion for a diagnosis of neuroblastoma is based on one of the following: (a) an unequivocal pathologic diagnosis made from tumor tissue by light microscopy (with or without immunohistology, electron microscopy, or increased levels of serum catecholamines or urinary catecholamine metabolites); (b) the combination of bone marrow aspirate or biopsy containing unequivocal tumor cells and increased levels of serum catecholamines or urinary catecholamine metabolites as described earlier.

The patient should be evaluated by abdominal computerized tomography, which may show displacement and occasionally obstruction of the ureter of an intact kidney (Fig. 39-38). Prior to the institution of therapy, a complete staging workup should be performed. This includes radiograph of the chest, bone marrow biopsy, and radionuclide scans to search for metastases. Any abnormality on chest X-ray should be followed up with CT of the chest.

**Prognostic Indicators.** A number of biologic variables have been studied in children with neuroblastoma. An open biopsy is required in order to provide tissue for this analysis. Hyperdiploid tumor DNA is associated with a favorable prognosis, and \( N\text{-}myc \) amplification is associated with a poor prognosis regardless of patient age. The Shimada classification describes tumors as either favorable or unfavorable histology based on the degree of differentiation, the mitosis-karyorrhexis index, and the presence or absence of schwannian stroma. In general, children of any age with localized neuroblastoma and infants younger than 1 year of age with advanced disease and favorable disease characteristics have a high likelihood of disease-free survival. By contrast, older children with advanced-stage disease have a significantly decreased chance for cure despite intensive therapy. For example, aggressive multiagent chemotherapy has resulted in a 2-year survival rate of approximately 20% in older children with stage IV disease. Neuroblastoma in the adolescent has a worse long-term prognosis regardless of stage or site and, in many cases, a more prolonged course.

**Surgery.** The goal of surgery is complete resection. However, this is often not possible at initial presentation due to the extensive locoregional spread of the tumor at the time of presentation. Under these circumstances, a biopsy is performed, and preoperative chemotherapy is provided based upon the stage of the tumor. After neoadjuvant treatment has been administered, surgical resection is performed. The principal goal of surgery is to obtain at least 95% resection without compromising major structures. Abdominal tumors are approached through a transverse incision. Thoracic tumors may be approached through a posterolateral thoracotomy or through a thoracosopic approach. These may have an intraspinal component. In all cases of intrathoracic neuroblastoma, particularly those at the thoracic inlet, it is important to be aware of the possibility of a Horner’s syndrome (anhidrosis, ptosis, meiosis) developing. This typically resolves, although it may take many months to do so.

**Neuroblastoma in Infants.** Spontaneous regression of neuroblastoma has been well described in infants, especially in those with stage 4S disease. Regression generally occurs only in tumors with a near triploid number of chromosomes that also lack \( N\text{-}myc \) amplification and loss of chromosome 1p. Recent studies indicate that infants with asymptomatic, small, low-stage neuroblastoma detected by screening may have tumors that spontaneously regress. These patients may be observed safely without surgical intervention or tissue diagnosis.

**Rhabdomyosarcoma**

Rhabdomyosarcoma is a primitive soft tissue tumor that arises from mesenchymal tissues. The most common sites of origin include the head and neck (36%), extremities (19%), genitourinary tract (2%), and trunk (9%), although the tumor can arise virtually anywhere. The clinical presentation of the tumor depends on the site of origin. The diagnosis is confirmed with incisional or excisional biopsy after evaluation by MRI, CT scans of the affected area and the chest, and bone marrow biopsy. The tumor grows locally into surrounding structures and metastasizes widely to lung, regional lymph nodes, liver, brain, and bone marrow. The staging system for rhabdomyosarcoma is based upon the TNM system, as established by the Soft Tissue Sarcoma Committee of the Children’s Oncology Group. It is shown in Table 39-4. Surgery is an important component of the staging strategy and involves biopsy of the lesion and evaluation of lymphatics. Primary resection should be undertaken when complete excision can be performed without causing disability. If this is not possible, the lesion is biopsied, and intensive chemotherapy is administered. It is important to plan the biopsy so that it does not interfere with subsequent resection. After the
tumor has decreased in size, resection of gross residual disease should be performed. Radiation therapy is effective in achieving local control when microscopic or gross residual disease exists following initial treatment. Patients with completely resected tumors of embryonal histology do well without radiation therapy, but radiation therapy benefits patients with group I tumors with alveolar or undifferentiated histology.

**Prognosis.** The prognosis for rhabdomyosarcoma is related to the site of origin, resectability, presence of metastases, number of metastatic sites, and histopathology. Primary sites with more favorable prognoses include the orbit and nonparameningeal head and neck, paratestis and vagina (nonbladder, nonprostate genitourinary), and the biliary tract. Patients with tumors less than 5 cm in size have improved survival compared to children with larger tumors, while children with metastatic disease at diagnosis have the poorest prognosis. Tumor histology influences prognosis and the embryonal variant is favorable while the alveolar subtype is unfavorable.

**Teratoma**

Teratomas are tumors composed of tissue from all three embryonic germ layers. They may be benign or malignant, they may arise in any part of the body, and they are usually found in midline structures. Thoracic teratomas usually present as an anterior mediastinal mass. Ovarian teratomas present as an abdominal mass often with symptoms of torsion, bleeding, or rupture. Retroperitoneal teratomas may present as a flank or abdominal mass.

Most tumors are identified at birth and are benign. Malignant yolk sac tumor histology occurs in a minority of these tumors. Complete resection of the tumor as early as possible is essential. The rectum and genital structures are often distorted by the tumor but usually can be preserved in the course of resection. Perioperative complications of hypothermia and hemorrhage can occur with massive tumors and may prove lethal. This is of particular concern in small, preterm infants with large tumors. The cure rate is excellent if the tumor is excised completely.

**Sacroccocygeal Teratoma.** Sacroccocygeal teratoma usually presents as a large mass extending from the sacrum in the newborn period. Diagnosis may be established by prenatal US. In fetuses with evidence of hydrops and a large sacroccocygeal teratoma, prognosis is poor; thus, prenatal intervention has been advocated in such patients. The mass may be as small as a few centimeters in diameter or as massive as the size of the infant (Fig. 39-39). The tumor has been classified based upon the location and degree of intrapelvic extension. Lesions that grow predominantly into the presacral space often present later in childhood. The differential diagnosis consists of neural tumors, lipoma, and myelomeningoceles.

Most tumors are identified at birth and are benign. Malignant yolk sac tumor histology occurs in a minority of these tumors. Complete resection of the tumor as early as possible is essential. The rectum and genital structures are often distorted by the tumor but usually can be preserved in the course of resection. Perioperative complications of hypothermia and hemorrhage can occur with massive tumors and may prove lethal. This is of particular concern in small, preterm infants with large tumors. The cure rate is excellent if the tumor is excised completely.

**Table 39-4**

<table>
<thead>
<tr>
<th>STAGE</th>
<th>SITES</th>
<th>T SIZE</th>
<th>N</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Orbit, nonparameningeal head and neck, genitourinary (other than kidney, bladder, and prostate), and biliary</td>
<td>T1 or T2</td>
<td>a or b</td>
<td>Any</td>
</tr>
<tr>
<td>2</td>
<td>Bladder/prostate, extremity, cranial parameningeal, other</td>
<td>T1 or T2</td>
<td>a</td>
<td>N0 or NX</td>
</tr>
<tr>
<td>3</td>
<td>Bladder/prostate, extremity, cranial parameningeal, other</td>
<td>T1 or T2</td>
<td>a</td>
<td>N1</td>
</tr>
<tr>
<td>4</td>
<td>All</td>
<td>T1 or T2</td>
<td>a or b</td>
<td>Any</td>
</tr>
</tbody>
</table>

T1 = tumor confined to anatomic site of origin; T2 = tumor extension and/or fixed to surrounding tissues; a = ≤5 cm; b = >5 cm; N0 = regional nodes not clinically involved; N1 = regional nodes clinically involved; NX = regional node status unknown; M0 = no distant metastasis; M1 = metastasis present.

Clinical group:
- Group 1: Localized disease, completely resected, no regional lymph node involvement.
- Group 2: Localized disease, gross total resection but microscopic residual disease; or regional lymph nodes involved.
- Group 3: Localized disease with gross residual disease after incomplete resection or biopsy only.
- Group 4: Metastatic disease at diagnosis.

**Figure 39-39.** Sacroccocygeal teratoma in a 2-day-old boy.
The majority of patients who develop recurrent disease are salvagable with subsequent platinum-based chemotherapy.

**Liver Tumors**

More than two-thirds of all liver tumors in children are malignant. There are two major histologic subgroups: hepatoblastoma and hepatocellular carcinoma. The age of onset of liver cancer in children is related to the histology of the tumor. Hepatoblastoma is the most common malignancy of the liver in children, with most of these tumors diagnosed before 4 years of age. Hepatocellular carcinoma is the next most common, with a peak age incidence between 10 and 15 years. Malignant mesenchymomas and sarcomas are much less common but constitute the remainder of the malignancies. The finding of a liver mass does not necessarily imply that a malignancy is present. Nearly 50% of all masses are benign, and hemangiomas are the most common lesion.

Most children with a liver tumor present with an abdominal mass that is usually painless, which the parents note while changing the child’s clothes or while bathing the child. The patients are rarely jaundiced but may complain of anorexia and weight loss. Most liver function tests are normal. AFP levels are increased in 90% of children with hepatoblastomas but much less commonly in other liver malignancies. Radiographic evaluation of these children should include an abdominal CT scan to identify the lesion and to determine the degree of local invasiveness (Fig. 39-40). For malignant appearing lesions, a biopsy should be performed unless the lesion can be completely resected easily. Hepatoblastoma is most often unifocal, while hepatocellular carcinoma is often extensively invasive or multicentric. If a hepatoblastoma is completely removed, the majority of patients survive, but only a minority of patients have lesions amenable to complete resection at diagnosis.

A staging system based on postsurgical extent of tumor and surgical resectability is shown in Table 39-5. The overall survival rate for children with hepatoblastoma is 70%, but it is only 25% for hepatocellular carcinoma. Children diagnosed with stage I and II hepatoblastoma have a cure rate of greater than 90% compared to 60% for stage III and approximately 20% for stage IV. In children diagnosed with hepatocellular carcinoma, those with stage I have a good outcome, whereas stages III and IV are usually fatal. The fibrolamellar variant of hepatocellular carcinoma may have a better prognosis.

**Surgery.** The abdominal CT scan usually will determine the resectability of the lesion, although occasionally this can only be determined at the time of exploration. Complete surgical resection of the tumor is the primary goal and is essential for cure. For tumors that are unresectable, preoperative chemotherapy should be administered to reduce the size of the tumor and improve the possibility for complete removal. Chemotherapy is more successful for hepatoblastoma than for hepatocellular carcinoma. Areas of locally invasive disease, such as the diaphragm, should be resected at the time of surgery. For unresectable tumors, liver transplantation may be offered in select patients. The fibrolamellar variant of hepatocellular carcinoma may have a better outcome with liver transplantation than other hepatocellular carcinomas.

**TRAUMA IN CHILDREN**

Injury is the leading cause of death among children older than 1 year. In fact, trauma accounts for almost half of all pediatric deaths, more than cancer, congenital anomalies, pneumonia, heart disease, homicide, and meningitis combined. Death from unintentional injuries accounts for 65% of all injury-related deaths in children younger than 19 years. Motor vehicle collisions are the leading cause of death in people age 1 to 19 years, followed by homicide or suicide (predominantly with firearms) and drowning. Each year, approximately 20,000 children and teenagers die as a result of injury in the United States. For every child who dies from an injury, it is calculated that 40 others are hospitalized and 1120 are treated in emergency departments. An estimated 50,000 children acquire permanent disabilities each year, most of which are the result of head injuries. Thus, the problem of pediatric trauma continues to be one of the major threats to the health and well-being of children.

Specific considerations apply to trauma in children that influence management and outcome. These relate to the mechanisms of injury, the anatomic variations in children compared to adults, and the physiologic responses.

**Mechanisms of Injury**

Most pediatric trauma is blunt. Penetrating injuries are seen in the setting of gun violence, falls onto sharp objects, or penetration by glass after falling through windows. Age and gender significantly influence the patterns of injury. Male children between 14 and 18 years of age are exposed to contact sports, gun violence, and in some jurisdictions drive motor vehicles. As a result, they have a different pattern of injury than younger children, characterized by higher injury severity scores. In the infant and toddler age group, falls are a

![Figure 39-40. Computed tomography of the abdomen showing a hepatocellular carcinoma in a 12-year-old boy.](image)
common cause of severe injury. Injuries in the home are extremely common. These include falls, near-drownings, caustic ingestion, and nonaccidental injuries.

**Initial Management**

The goals of managing the pediatric trauma patient are similar to those of adults and follow Advanced Trauma Life Support guidelines as established by the American College of Surgeons Committee on Trauma. Airway control is the first priority. In a child, respiratory arrest can proceed quickly to cardiac arrest. It is important to be aware of the anatomic differences between the airway of the child and the adult. The child has a large head, shorter neck, smaller and anterior larynx, floppy epiglottis, short trachea, and large tongue. The size of the endotracheal tube can be estimated by the formula \((\text{age} + 16)/4\). It is important to use uncuffed endotracheal tubes in children younger than 8 years in order to minimize tracheal trauma. After evaluation of the airway, breathing is assessed. It is important to consider that gastric distention from aerophagia can severely compromise respirations. A nasogastric tube should therefore be placed early during the resuscitation if there is no head injury suspected, or an orogastric tube in cases of head injury. Pneumothorax or hemothorax should be treated promptly. When evaluating the circulation, it is important to recognize that tachycardia is usually the earliest measurable response to hypovolemia. Other signs of impending hypovolemic shock in children include changes in mentation, delayed capillary refill, skin pallor, and hypothermia. IV access should be rapidly obtained once the patient arrives in the trauma bay. The first approach should be to use the antecubital fossae. If this is not possible, a cut-down into the saphenous at the groin can be performed quickly and safely. Intraosseous cannulation can provide temporary access in children and young adults until IV access is established. US-guided central line placement in the groin or neck should be considered in patients in whom large bore peripheral IV access is not obtained. Blood is drawn for cross-match and evaluation of liver enzymes, lipase, amylase, and hematologic profile after the IV lines are placed.

In patients who show signs of volume depletion, a 20 mL/kg bolus of saline or lactated Ringer’s should be promptly given. If the patient does not respond to three boluses, blood should be transfused (10 mL/kg). The source of bleeding should be established. Common sites include the chest, abdomen, pelvis, extremity fractures, or large scalp wounds. These should be carefully sought. Care is taken to avoid hypothermia by infusing warmed fluids and by using external warming devices.

**Evaluation of Injury**

All patients should receive an X-ray of the cervical spine, chest, and abdomen with pelvis. All extremities that are suspicious for fracture should also be evaluated by X-ray. Plain cervical spine films are preferable to performing routine neck CT scans in the child, as X-rays provide sufficient anatomic detail. But if a head CT is obtained, it may be reasonable to obtain images down to C-2 since odontoid views in small children are difficult to obtain. In most children, it is possible to diagnose clinically significant cervical spine injuries using this approach while minimizing the degree of radiation exposure. Screening blood work that includes AST, ALT, and amylase/lipase is useful for the evaluation of liver and pancreatic injuries. Significant elevation in these tests requires further evaluation by CT scanning. The child with significant abdominal tenderness and a mechanism of injury that could cause intra-abdominal injury should undergo abdominal CT scanning using IV and oral contrast in all cases. There is a limited role for diagnostic peritoneal lavage (DPL) in children as a screening test. However, this can be occasionally useful in the child who is brought emergently to the operating room for management of significant intracranial hemorrhage. At the time of craniotomy, a DPL, or alternatively, a diagnostic laparoscopy, can be performed concurrently to identify abdominal bleeding. Although focused abdominal US (FAST exam) is extremely useful in the evaluation of adult abdominal trauma, it is not widely accepted in the management of pediatric blunt abdominal trauma. In part, this relates to the widespread use of nonoperative treatment for most solid-organ injuries. Thus, a positive abdominal US scan would not alter this approach in a hemodynamically stable patient.

**Injuries to the Central Nervous System**

The central nervous system (CNS) is the most commonly injured organ system and is the leading cause of death among injured children. In the toddler age group, nonaccidental trauma is the most common cause of serious head injury. Findings suggestive of abuse include the presence of retinal hemorrhage on fundoscopic evaluation and intracranial hemorrhage without evidence of external trauma (indicative of a shaking injury) and fractures at different stages of healing on skeletal survey. In older children, CNS injury occurs most commonly after falls and bicycle and motor vehicle collisions. The initial head CT can often underestimate the extent of injury in children. Criteria for head CT include any loss of consciousness or amnesia to the trauma, or inability to assess the CNS status as in the intubated patient. Patients with mild, isolated head injury (GCS 14-15) and negative CT scans can be discharged if their neurologic status is normal after 6 hours of observation. Young children and those in whom there is multisystem involvement should be admitted to the hospital for observation. Any change in the neurologic status warrants neurosurgical evaluation and repeat CT scanning. In patients with severe head injury (GCS 8 or less), urgent neurosurgical consultation is required. These patients are evaluated for intracranial pressure monitoring and for the need to undergo craniotomy.

**Thoracic Injuries**

The pediatric thorax is pliable due to incomplete calcification of the ribs and cartilages. As a result, blunt chest injury commonly results in pulmonary contusion, although rib fractures are infrequent. Diagnosis is made by chest radiograph and may be associated with severe hypoxia requiring mechanical ventilation. Pulmonary contusion usually resolves with careful ventilator management and judicious volume resuscitation. Children who have sustained massive blunt thoracic injury may develop traumatic asphyxia. This is characterized by cervical and facial petechial hemorrhages or cyanosis associated with vascular engorgement and subconjunctival hemorrhage. Management includes ventilation and treatment of coexisting CNS or abdominal injuries. Penetrating thoracic injuries may result in damage to the lung or to major disruption of the bronchi or great vessels.

**Abdominal Injuries**

In children, the small rib cage and minimal muscular coverage of the abdomen can result in significant injury after seemingly minor trauma. The liver and spleen in particular are relatively unprotected and are often injured after direct abdominal trauma. Duodenal injuries are usually the result of blunt trauma, which may arise from child abuse or injury from a bicycle handlebar. Duodenal hematomas usually resolve without surgery.
Small intestinal injury usually occurs in the jejunum in the area of fixation by the ligament of Treitz. These injuries are usually caused by rapid deceleration in the setting of a lap belt. There may be a hematoma on the anterior abdominal wall caused by a lap belt, the so-called seat belt sign (Fig. 39-41A). This should alert the caregiver to the possibility of an underlying small bowel injury (Fig. 39-41B), as well as to a potential lumbar spine injury (Chance fracture).

The spleen is injured relatively commonly after blunt abdominal trauma in children. The extent of injury to the spleen is graded (Table 39-6), and the management is governed by the injury grade. Current treatment involves a nonoperative approach in most cases, even for grade 4 injuries, assuming the patient is hemodynamically stable. This approach avoids surgery in most cases. All patients should be placed in a monitored unit, and type-specific blood should be available for transfusion. When nonoperative management is successful, as it is in most cases, an extended period of bed rest is prescribed. This optimizes the chance for healing and minimizes the likelihood of reinjury. A typical guideline is to keep the children on extremely restricted activity for 2 weeks longer than the grade of spleen injury (i.e., a child with a grade 4 spleen injury receives 6 weeks of restricted activity). In children who have an ongoing fluid requirement, or when a blood transfusion is required, exploration should not be delayed. At surgery, the spleen can often be salvaged. If a splenectomy is performed, prophylactic antibiotics and immunizations should be administered to protect against overwhelming post splenectomy sepsis. The liver is also commonly injured after blunt abdominal trauma. A grading system is used to characterize hepatic injuries (Table 39-7), and nonoperative management is usually successful (Fig. 39-42). Recent studies have shown that associated injuries are more significant predictors of outcome in children with liver injuries than the actual injury grade. Criteria for surgery are similar to those for splenic injury and primarily involve hemodynamic instability. The intraoperative considerations in the management of massive hepatic injury are similar in children and adults. Renal contusions may occur after significant blunt abdominal trauma. Nonoperative management is usually successful, unless patients are unstable due to active renal bleeding. It is important to confirm the presence of a normal contralateral kidney at the time of surgery.

**FETAL INTERVENTION**

One to the most exciting developments in the field of pediatric surgery has been the emergence of fetal surgery. In general terms, performance of a fetal intervention may be justified in the setting where a defect is present that would cause devastating consequences to the infant if left uncorrected. For the vast majority of congenital anomalies, postnatal surgery is the preferred modality. However, in specific circumstances, fetal surgery may offer the best possibility for a successful outcome.

### Table 39-6

**Grading of splenic injuries**

<table>
<thead>
<tr>
<th>Grade I</th>
<th>Subcapsular hematoma, &lt;10% surface area capsular tear, &lt;1 cm in depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade II</td>
<td>Subcapsular hematoma, nonexpanding, 10%–50% surface area; intraparenchymal hematoma, nonexpanding, &lt;2 cm in diameter; capsular tear, active bleeding, 1–3 cm, does not involve trabecular vessel</td>
</tr>
<tr>
<td>Grade III</td>
<td>Subcapsular hematoma, &gt;50% surface area or expanding; intraparenchymal hematoma, &gt;2 cm or expanding; laceration &gt;3 cm in depth or involving trabecular vessels</td>
</tr>
<tr>
<td>Grade IV</td>
<td>Ruptured intraparenchymal hematoma with active bleeding; laceration involving segmental or hilar vessels producing major devascularization (&gt;25% of spleen).</td>
</tr>
<tr>
<td>Grade V</td>
<td>Shattered spleen; hilar vascular injury that devascularizes spleen</td>
</tr>
</tbody>
</table>

### Table 39-7

**Liver injury grading system**

<table>
<thead>
<tr>
<th>Grade I</th>
<th>Capsular tear &lt;1 cm in depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade II</td>
<td>Capsular tear 1–3 cm in depth, &lt;10 cm length</td>
</tr>
<tr>
<td>Grade III</td>
<td>Capsular tear &gt;3 cm in depth</td>
</tr>
<tr>
<td>Grade IV</td>
<td>Parenchymal disruption 25%–75% of hepatic lobe or 1–3 Couinaud’s segments</td>
</tr>
<tr>
<td>Grade V</td>
<td>Parenchymal disruption &gt;75% of hepatic lobe or &gt;3 Couinaud’s segments within a single lobe, injury to retrohepatic vena cava</td>
</tr>
</tbody>
</table>

Specific Considerations

III

Fetal Surgery for Myelomeningocele

Myelomeningocele refers to a spectrum of anomalies in which portions of the spinal cord are uncovered by the spinal column. This leaves the neural tissue exposed to the injurious effects of the amniotic fluid, as well as to trauma from contact with the uterine wall. Nerve damage ensues, resulting in varying degrees of lower extremity paralysis as well as bowel and bladder dysfunction. Initial observations indicated that the extent of injury progressed throughout the pregnancy, which provided the rationale for fetal intervention. The current in utero approach for the fetus with myelomeningocele has focused on obtaining coverage of the exposed spinal cord. The efficacy of in utero treatment versus postnatal repair was recently compared in a large multicenter trial as described earlier and showed that prenatal surgery for myelomeningocele reduced the need for shunting and improved motor outcomes at 30 months but was associated with maternal and fetal risks. The results of this study have paved the way for the acceptance of in utero repair of myelomeningocele in certain centers with the experience and expertise to perform this procedure safely.

The EXIT Procedure

The EXIT procedure is an abbreviation for ex utero intrapartum treatment. It is utilized in circumstances where airway obstruction is predicted at the time of delivery due to the presence of a large neck mass, such as a cystic hygroma or teratoma (Fig. 39-43), or congenital tracheal stenosis. The success of the procedure is dependent upon the maintenance of utero-placental perfusion for a sufficient duration to secure the airway. To achieve this, deep uterine relaxation is obtained during a caesarian section under general anesthesia. Uterine perfusion with warmed saline also promotes relaxation and blood flow to the placenta. On average, between 20 and 30 minutes of placental perfusion can be achieved. The fetal airway is secured either by placement of an orotracheal tube or performance of a tracheotomy. Once the airway is secured, the cord is cut, and a definitive procedure may be performed to relieve the obstruction in the postnatal period. In general terms, cystic neck masses such as lymphangiomas have a more favorable response to an EXIT procedure as compared to solid tumors, such as teratomas, particularly in premature infants.

Figure 39-42. Abdominal computed tomography in a child demonstrating a grade 3 liver laceration (arrows).

Fetal Surgery for Lower Urinary Tract Obstruction

Lower urinary tract obstruction refers to a group of diseases characterized by obstruction of the distal urinary system. Common causes include the presence of posterior urethral valves and urethral atresia, as well as other anomalies of the urethra and bladder. The pathologic effects of lower urinary tract obstruction lie in the resultant massive bladder distention that occurs, which can lead to reflux hydronephrosis. This may result in oligohydramnios, and cause limb contractures, facial anomalies (Potter sequence), and pulmonary hypoplasia. Carefully selected patients with lower urinary tract obstruction may benefit from vesicoamniotic shunting. By relieving the obstruction and improving renal function, fetal growth and lung development may be preserved.

Figure 39-43. The EXIT procedure (ex utero intrapartum treatment) in a 34-week gestation age baby with a large cervical teratoma. Intubation is being performed while the fetus is on placental support.
BIBLIOGRAPHY

Entries highlighted in bright blue are key references.


PART II
SPECIFIC CONSIDERATIONS


Pritchard-Jones K. Controversies and advances in the management of Wilms’ tumour. *Arch Dis Child.* 2002;87:241-244.


The anatomic structures that generally require urologic management include the kidneys, adrenal glands, ureters, bladder, prostate, seminal vesicles, vas deferens, penis, urethra, scrotum, and testes. These organs are located in retroperitoneal or extraperitoneal spaces. However, a transperitoneal approach may be utilized to access the kidney, ureters, bladder, or retroperitoneal lymph nodes during certain urologic operations.

**Kidney and Adrenal Gland**

The kidneys are paired retroperitoneal organs that are invested in a fibro-fatty layer of tissue known as Gerota’s fascia. This natural barrier helps to tamponade bleeding and thus may provide renal and hemodynamic protection in cases of renal trauma or spontaneous renal hemorrhage. It also may assist in preventing tumor invasion into surrounding structures in the case of large renal masses. These organs are located in retroperitoneal or extraperitoneal spaces. However, a transperitoneal approach may be utilized to access the kidney, ureters, bladder, or retroperitoneal lymph nodes during certain urologic operations.

The kidneys are borderer posterolaterally by the quadratus lumborum muscle and posteromedially by the psoas muscle. Additionally, the diaphragm drapes across the posterior aspect of the superior pole of each kidney.

The left kidney is bordered anterolaterally by the spleen and descending colon. The pancreatic tail borders the anteromedial left kidney. The right kidney is bordered anterolaterally by the liver and the ascending colon. The second portion of the duodenum may be encountered near the right renal vessels and thus sometimes requires anteromedial reflection, known as the Kocher maneuver, to achieve intraoperative vascular control during right renal surgery.

The kidneys are end organs, which are responsible for their vulnerability to infarction. The renal arteries extend from the aorta and then branch into several segmental arteries and arterioles before becoming glomeruli. Each artery runs posterior to their respective renal vein. Occasionally, an accessory renal artery will arise, but in general, each kidney receives a single main renal artery. Each renal vein drains directly into the IVC and is located anteriorly to its respective renal artery when entering the kidney. The right renal vein is much shorter than the left and does not receive collateral venous drainage. The left renal vein passes anteriorly to the aorta and receives drainage from the left gonadal vein, the left inferior adrenal vein, and a lumbar vein.

The collecting system of the kidney begins as minor calyces near the renal papillae and then coalesces into major calyces. Major calyces join to form the renal pelvis, which then tapers down to the ureteropelvic junction (UPJ), from which the ureter emanates. The pelvis is located posterior to its respective renal artery.

The adrenal gland is superomedial to its respective kidney within Gerota’s fascia. Adrenal arterial supply arises from multiple sources: the inferior phrenic artery, aortic branches, and renal arterial branches. Venous drainage mirrors arterial supply. On the right side, the adrenal gland drains directly into the IVC. The right adrenal vein can be quite short (<1 cm) and can be
Key Points

1. Most small ureteral calculi will pass spontaneously or with the use of medical expulsive therapy, but larger stones (>6 mm) are better treated with ureteral stenting or lithotripsy.

2. Benign prostatic hyperplasia can be managed effectively with medical therapy or minimally invasive endoscopic and robotic surgical techniques depending on the urinary symptoms, patient bother, prostate size, and patient’s therapeutic choice.

3. Patients with recurrent urethral stricture after endoscopic treatment are unlikely to derive sustained benefit from future endoscopic therapies and should be referred for urethral reconstruction.

4. The vast majority of renal trauma can be treated conservatively, with early surgical intervention reserved for persistent bleeding, renal vascular, or ureteral injuries.

5. Extraperitoneal bladder ruptures can be treated conservatively, but intraperitoneal ruptures typically require surgical repair.

6. Testicular torsion is an emergency where successful testicular salvage is inversely related to the delay in repair, so cases with a high degree of clinical suspicion should not wait for a radiologic diagnosis.

7. Fournier’s gangrene is a rapidly progressive and potentially lethal condition that requires aggressive débridement and close follow-up due to the frequent need for repeat débridement.

8. The management of early stage prostate cancer has changed significantly, with a much greater emphasis on risk stratification. Low risk patients are largely treated with active surveillance.

9. Treatments for urinary incontinence and voiding dysfunction are varied depending on the etiology, severity, and bother of the symptom. Behavior modification, bladder retraining, and medical therapies can all be effective in improving symptoms without the need for surgery.

The bladder is located extraperitoneally in the pelvis and posterior to the aorta and renal artery, and the distal blood supply inserts laterally and arises from the surrounding iliac vessels and their branches. The arterial supply enters via a fatty layer of tissue around the ureter, and thus surgical preservation of the periureteral tissue is essential to maintain vascularization and achieve successful ureteral reconstruction.

The ureters initially course along the psoas muscle and then run distally along the pelvic sidewall. They generally pass posterior to the uterine arteries, making them susceptible to injury during hysterectomy. The ureters enter the bladder laterally and pass through the bladder wall at an oblique angle, which helps prevent reflux of urine during bladder filling. The ureters propel urine into the bladder via the ureteral orifices.

Bladder and Prostate

The bladder is located extraperitoneally in the pelvis and posterior to the pubis. A portion of the bladder dome is draped by peritoneum, and rupture or injury at this location can result in intraperitoneal urine leakage and subsequent chemical peritonitis. The average adult bladder holds approximately 500 mL of urine; however, in rare cases, capacity can reach up to or greater than 1000 mL, in which case the bladder extends towards the umbilicus. The sigmoid colon lies adjacent to the bladder and can fistulize to the lateral wall or dome of the bladder in cases of diverticulitis or colon cancer. The rectum lies posteriorly to the bladder in men, and the uterus and vagina lie posteriorly to the bladder in women.

The prostate is a walnut-shaped gland that encircles the urethra and is located in males immediately beneath the bladder neck. Smooth muscle fibers distribute throughout the gland, which can contract and facilitate bladder outlet obstruction. The average prostate measures approximately 30 mL in volume. Puboprostatic ligaments suspend the prostate to the pubis, and in the instance of pelvic trauma, shearing forces can cause disruption of the posterior urethra (known as pelvic fracture urethral injury). The external urethral sphincter houses the membranous urethra and sits just below the apex of the prostate. Vasculature to the bladder and prostate arises from the superior and inferior vesical arteries, which branch from the internal iliac arteries.

Penis

The penis is comprised of three bodies: two corpora cavernosa, which are responsible for erection, and the corpus spongiosum, which surrounds the urethra and gives rise to the glans penis. These three structures are all encased by skin and dartos fascia, as well as an inner investing layer of fascia called Buck’s fascia. The corpora cavernosa are spongy sinusoidal bodies that expand with parasympathetic neural stimulation to create an erection. Thick fascia, called tunica albuginea, assists in producing rigidity during erection. Each corpus cavernosum features a centrally located cavernosal artery, which arises from the penile artery. A porous septum separates the two corpora and allows for transcorporal blood exchange. The corpus spongiosum is located on the ventrum of the penis. The corpus spongiosum lacks a tough fascia similar to tunica albuginea and thus does not exhibit the same rigidity during erection.

Scrotum and Testes

The scrotum is a potential space that surrounds the testes, epididymis, and spermatic cords. The scrotum is comprised of many layers aside from skin and dartos fascia, and each derives from a particular layer of the anterior abdominal wall. The external spermatic fascia arises from the external oblique fascia, the cremasteric fascia arises from the internal oblique fascia, and the internal spermatic fascia arises from the transversus abdominis fascia. The testes are separated from the scrotal layers by the visceral and parietal layers of the tunica vaginalis, between which hydroceles form. The spermatic cord contains...
the vas deferens, the venous pampiniform plexus, and arterial blood supply to the superior pole of the testis via three separate sources. The testicular artery arises directly from the aorta; the deferential artery, which supplies the vas deferens, arises from the internal iliac artery; and the cremasteric artery, which supplies the cremaster muscle, arises from the external iliac artery. The presence of multiple arterial sources provides collateral flow and prevents ischemia in the event of injury to a particular vascular branch. The venous pampiniform plexus can dilate to form a palpable or visible varicocele, which can serve as an etiology of chronic testicular pain or infertility.

**INFECTION**

**Cystitis**
Uncomplicated cystitis usually presents as new onset urinary frequency, urgency, and dysuria. Patients may also report lower back pain, suprapubic pain, foul-smelling urine, or gross hematuria. Urinalysis with microscopy assists with diagnosis by confirming the presence of pyuria, hematuria, and bacteriuria. Office dipstick may be helpful, as the presence of nitrates reflects bacterial colonization and the presence of leukocyte esterase reflects pyuria. Risk factors for the development of uncomplicated cystitis include female gender, sexual activity, and use of spermicides.\(^1\) Three days of antibiotics are generally sufficient for treatment of uncomplicated cystitis. Fluoroquinolones and trimethoprim-sulfamethoxazole are well tolerated and are easily available. Nitrofurantoin, which is also commonly used for uncomplicated cystitis, requires 5 days of treatment. Men with uncomplicated cystitis should undergo 7 days of treatment.

Complicated cystitis may arise in the setting of structural or functional urinary tract abnormalities, recent urinary tract instrumentation, recent antimicrobial use, immunosuppressed states, pregnancy, or hospital-acquired infection. Symptoms may be similar to uncomplicated cystitis but can progress to pyelonephritis if left untreated. Elderly or very young patients tend to exhibit lethargy, change in mental status, or anorexia, which may confound the diagnosis of a urinary tract infection. Patients may require hospitalization if febrile or if symptoms are severe. Treatment consists of 10 to 14 days of antibiotics. Fluoroquinolones or trimethoprim-sulfamethoxazole are usually effective and should be administered based on culture results and/or regional bacteriograms. Asymptomatic bacteriuria does not require treatment unless detected during pregnancy or if urinary tract instrumentation is planned.\(^1\)

**Pyelonephritis**
Pyelonephritis arises when a bladder infection ascends proximally along the ureters to the renal parenchyma. It may also result from hematogenous spread, such as in the case of intravenous drug abuse or in patients with bacteremia from other sources. Patients with pyelonephritis may present with fevers, flank pain, nausea, vomiting, and lower urinary tract symptoms. Physical exam may reveal tenderness of the costovertebral angle. Patients may appear toxic, with poor oral intake. Laboratory evaluation may reveal leukocytosis with elevated neutrophils. Urinalysis usually demonstrates the presence of pyuria and bacteriuria, and urine culture should be sent prior to starting broad-spectrum antibiotics. Imaging should be considered to rule out obstruction, which could prolong the recovery period despite appropriate antimicrobial treatment.

Acute pyelonephritis requires 7 to 14 days of antibiotic therapy. Mild or moderate cases, even if febrile, can safely be treated as an outpatient with oral antibiotics. Fluoroquinolones and trimethoprim-sulfamethoxazole are ideal for treating pyelonephritis. Nitrofurantoin should not be used as it does not penetrate renal parenchyma. Patients with concern for sepsis or inability to tolerate oral intake may require hospitalization with IV antibiotics while awaiting culture results. Fevers may persist for up to 72 hours despite appropriate treatment. The presence of persistent fevers or symptoms after this time period warrants cross-sectional imaging to rule out renal or perinephric abscess. Treatment for renal or perinephric abscess usually consists of percutaneous drainage and broad-spectrum IV antibiotics.

**Prostatitis**
Acute prostatitis is marked by fever, suprapubic or perineal pain, and new onset lower urinary tract symptoms, namely dysuria, frequency, urgency, changes in stream caliber, or difficulty emptying the bladder. It is most often caused by urinary pathogens. Digital rectal exam may reveal a tender and soft prostate. Bladder drainage with a Foley or suprapubic tube may be required if urinary retention is present. Treatment consists of a long-term course (4–6 weeks) of antibiotics. If not treated in a timely fashion, acute prostatitis can develop into severe sepsis or a prostatic abscess. Prostatic abscesses may require drainage via a transurethral approach or transrectal needle aspiration.

Chronic prostatitis may be bacterial or abacterial. Symptoms in both cases include perineal, suprapubic, or penile pain, along with urinary frequency, urgency, or change in stream caliber. Men may also report pain in the groin, lower back, or testes. Fever is not observed in chronic prostatitis, and onset may occur over many months. Patients with chronic bacterial prostatitis may also report recurrent UTIs, with cultures consistently exhibiting the same bacteria. Differentiation between the two etiologies requires culture of expressed prostatic secretion to confirm the presence or absence of bacteria. Treatment of chronic bacterial prostatitis includes long-term antibiotics and \(\alpha\)-blockers.

Chronic abacterial prostatitis is also known as chronic pelvic pain syndrome (CPPS). Symptoms are similar to chronic bacterial prostatitis, but generally do not respond well to long-term antibiotics for treatment. It is generally somewhat more difficult to achieve symptomatic relief when treating CPPS, and options include \(\alpha\)-blockers, NSAIDs, neuromodulators, and/or pelvic floor physical therapy.\(^2\)

**Epididymo–Orchitis**
Epididymitis refers to inflammation of the epididymis. In most cases of bacterial infection, the testis is also affected, thus is encompassed by the term “epididymo–orchitis.” Common etiologies include sexually transmitted infection, especially in younger males, or urinary tract infection, which is more commonly seen in older males. Other possible etiologies include underlying congenital urologic abnormality or incomplete bladder emptying. Symptoms include pain and swelling of the epididymis and testis. Some men may report nausea or vomiting, which arises as a result of irritation of the spermatic cord. Urinary symptoms may be present, but absence of symptoms does not rule out bacterial epididymo–orchitis. Physical exam generally reveals a tender, swollen epididymis and testis. Scrotal skin erythema or reactive hydrocele may be present as well. A complete blood count should be performed to rule out leukocytosis, and urinalysis with urine culture should be collected prior to initiation of antibiotics. Urethral swab should be performed if sexually transmitted infection is a possible etiology. The clinical presentation of testicular torsion can be quite similar to that of...
epididymo-orchitis. It may be quite difficult to clinically differentiate the two entities, but one should keep in mind that the onset of torsion tends to be slightly more acute (within 4–8 hours) than that of epididymo-orchitis (which generally arises over the course of 24–48 hours). Scrotal ultrasound can assist in diagnosis; however, in cases of severe orchitis, testicular flow can be compromised, which may raise concern for torsion. Scrotal exploration should be considered in any equivocal case: a missed torsion can result in testicular loss secondary to necrosis.

Treatment of epididymo-orchitis consists of single dose of ceftriaxone and azithromycin if there is concern for sexually transmitted infection, as well as 14 days of oral antibiotic therapy, NSAIDs, and scrotal support. If the patient exhibits fevers or toxic presentation, hospitalization with IV antibiotics may be required.

**Balanitis and Balanoposthitis**

Balanitis refers to inflammation of the glans penis. Balanoposthitis arises when the foreskin is also involved. Common etiologies include fungal infection, bacterial infection, contact dermatitis, or local trauma. Exam reveals a diffusely erythematous and warm glans penis, with inner preputial erythema as well if balanoposthitis is present. Treatment includes appropriate hygiene, topical antibiotics or antifungals, and occasionally topical steroids. If there is an inappropriate response to treatment, the differential diagnosis should include malignancy, psoriasis, or infectious agents such as HPV.3

**URINARY TRACT OBSTRUCTION**

**Urolithiasis**

Renal stone disease is a common problem that is a major health care burden to society today. The prevalence of stone disease in the United States has increased over the past several decades as reported by the National Health and Nutrition Examination Survey (NHANES), and was estimated at 8.8% for the period between 2007 and 2010.4 This prevalence has increased with factors such as global warming, poor diet choices, and the obesity trend. Overall, the total estimated annual expenditure for individuals with claims for a diagnosis of urolithiasis was almost $2.1 billion in 2000, representing a 50% increase since 1994.5 Risk factors for stone formation include dietary habits, family history, white race, geographical location or occupational exposure to heat/dehydration, intestinal disease, and male gender, although the gender gap is decreasing.6 More recently, stone formation has also been associated with obesity, metabolic syndrome, and diabetes mellitus.7,8

Stones are most commonly composed of calcium oxalate. Other stone compositions include calcium phosphate, uric acid, cystine, medication-related, and infectious stones (struvite or carbonate apatite) or a mix thereof. Stone composition can vary based on a number of underlying pathophysiological processes. For example, hyperoxaluria may be seen in patients who have undergone small bowel resection, particularly the terminal ileum. This can result in an increase in unabsorbed fatty acids and bile salts which undergo saponification by binding with calcium in the bowel. The increase in unbound oxalate is absorbed by the large intestine and subsequently excreted in the urine, favoring the formation of calcium oxalate stones. Uric acid stones will form in a context of acidic urinary pH, low urinary volume, and high oral intake of purines. Countering these factors by alkalizing the urine and increasing urine output may lead to dissolution of uric acid stones and reduced further formation.9 Proteus species, Klebsiella species, and other urease-producing bacteria metabolize urea into ammonium and bicarbonate. The alkaline milieu (pH >7) predisposes to infectious (struvite) stones with the precipitation of magnesium, ammonium, and phosphate (Fig. 40-1).

Evaluation for first-time stone formers should include a complete medical history and physical exam, basic metabolic panel, calcium, uric acid, urinalysis and culture, and radiographic imaging. A noncontrast computed tomography (CT) scan is the most sensitive (98%) and specific (97%) exam to detect urolithiasis10 and can provide additional anatomical information useful for surgical planning, although its use in recurrent stone formers should be balanced by cost and radiation exposure. Low-dose CT is currently the preferred imaging study for patients with a body mass index (BMI) <30. This imaging study uses less than one-third of the estimated effective ionizing radiation dose (3 mSv) compared to standard dose noncontrast CT (10 mSv),10 while maintaining excellent sensitivity (95%) and specificity (97%).11 Plain abdominal X-ray can be used to follow radiopaque stones such as calcium-containing stones or struvite stones, although at times struvite can be difficult to see on plain X-ray, especially when the fragments are small. Uric acid and triamterene stones are radiolucent on plain abdominal X-ray but will be visible on noncontrast CT. A full metabolic evaluation with a 24-hour urine collection is indicated in recurrent stone formers, high-risk stone formers, or interested first-time stone formers.12

The natural history of stones is variable and depends primarily on their size and location. Smaller and more distal stones are much more likely to pass spontaneously without the need for surgical intervention.13,14 Patients with ureteral stones ≤10 mm can be offered a period of observation if their pain is well controlled without signs of infection or renal insufficiency. α-Blockers, which inhibit ureteral peristalsis, have been shown in meta-analyses to be particularly useful in patients with distal ureter stones ≤10 mm, improving the rate of stone passage from 54% to 77%,15,16 with shortened time to expulsion and fewer colic episodes.17
Patients who have not passed their stone after a 4- to 6-week observation period, those with larger stones, or those who desire immediate intervention, may be offered one of three definitive surgical interventions: shockwave lithotripsy (SWL), ureteroscopy (URS), or percutaneous nephrolithotomy (PCNL). Open surgical management of stones has been relegated to historic interest for the most part with less than 1% of stone surgery needing to be done open with access to modern endourologic equipment. The choice of the procedure will depend primarily on stone-related factors (e.g., stone size, location, and composition/density), and patient-related factors (e.g., comorbidities, coagulopathy, obesity, renal anatomy, and surrounding structures).

Shockwave lithotripsy is the procedure associated with the least morbidity and the lowest complication rate but is also associated with a lower success rate at treating stones as a single procedure and requires the patient to pass the stone fragments afterwards.15,16 The modality can be used for stones in the proximal ureter (particularly if <10 mm) or non–lower-pole renal stones <2 cm.15,16 The stone is located under fluoroscopic guidance, which is coupled to an extracorporeal lithotripter aimed at the stone. The stone is fragmented in a completely noninvasive manner. Complications associated with this procedure include subcapsular or perinephric renal hematoma and ureteral obstruction by stone fragments (“Steinstrasse”; Fig. 40-2). Ureteroscopy is the procedure of choice for patients with middle or distal ureteral stones. It also has a higher success rate than SWL in treating >10-mm proximal ureteral stones and renal stones.15,16 This procedure involves advancing a semi-rigid or flexible ureteroscope to the level of the stone and fragmenting it under direct visualization, often using a holmium:YAG laser. The surgeon is able to visualize the stone during fragmentation and thereby has some control over how small the fragments are. In addition, stone fragments may also be actively removed with a small nitinol stone basket. This is where the procedure may have an advantage over SWL. However, many patients have a ureteral stent placed after ureteroscopy, and, although temporary, this remains a major source of morbidity for the patient. Specific complications of URS include ureteral injury or stricture. PCNL is reserved for patients with larger or more complex stone burden, and requires a percutaneous tract into the kidney. Most stones larger than 2 cm are treated with PCNL although there is a role for PCNL for smaller stones located in the lower pole of the collecting system.18,19 More powerful lithotripters (pneumatic, ultrasound) and larger instruments (stone graspers) can be used to fragment and remove these larger stones through the percutaneous tract. Complications include injury to adjacent organs, acute and delayed renal bleeding due to pseudoaneurysm or arteriovenous fistula formation, sepsis, or renal pelvis perforation.

General preventative measures include correcting dietary habits, particularly increasing fluid intake to produce ≥2.5 liters of urine per day, limiting sodium, reducing animal protein intake, and monitoring foods high in oxalate. Medical therapy such as thiazide diuretics (helpful for hypercalciuria), urinary alkalization with potassium citrate, or allopurinol may also be indicated depending on the clinical situation.12

**Benign Prostatic Hyperplasia**

Benign prostatic hyperplasia (BPH) refers to the histological findings of smooth muscle and fibroblast/epithelial cell proliferation in the transition zone of the prostate. Lower urinary tract symptoms (LUTS) may be secondary to benign prostatic enlargement (BPE) causing progressive bladder outlet obstruction but may also be due to numerous other conditions (e.g., urethral stricture, infection, overactive or neurogenic bladder, malignancy). Although some male patients with LUTS may have BPE, not all patients with an enlarged prostate have LUTS. The prevalence of LUTS attributed to BPH in men over the age of 50 is estimated at 50% to 75% and increases with age with a prevalence of 80% in men over the age of 70.20 The treatment modalities have dramatically evolved over the past decades, with medical management typically used for first-line therapy. Endoscopic and minimally invasive techniques are used for those failing or intolerant of medical therapy.

Men with BPH/LUTS are evaluated with a complete history and physical exam including digital rectal exam. LUTS should be clearly defined, in addition to their severity and degree of bother. Validated questionnaires to quantify the patient’s symptoms and degree of bother include the American Urological Association Symptom Index (AUA-SI) and the International Prostate Symptom Score (IPSS).21,22 Complications of BPH such as urinary retention, incontinence, renal failure, hematuria, or recurrent infections should also be considered. Basic workup includes a urinalysis and culture to rule out infection. After an informative discussion about the risks and benefits of prostate cancer screening, a serum PSA is measured when life expectancy is >10 years and if the diagnosis of prostate cancer will alter management.23 Other diagnostic testing such as cystoscopy, cytology, postvoid residual (PVR), urodynamics, and radiologic imaging of the prostate, although not done routinely, may be required in patients with a definite indication (e.g., hematuria), uncertain diagnosis, poor response to therapy, or for surgical planning.24

The first line of treatment is most commonly pharmacotherapy for those men with bothersome symptoms. 2α-Blockers work by relaxing the smooth muscle of the prostate and bladder neck. All α-blocker agents are equally effective,25 and their side effects may include orthostatic hypotension, dizziness, asthenia, headache, nasal congestion, and retrograde ejaculation. Their effect is usually seen within days. Five-α reductase inhibitors (5-ARIs) block the conversion of...
testosterone to dihydrotestosterone (DHT), the hormone primarily responsible for BPH progression. These reduce prostatic size by 20% to 25%,\(^\text{26}\) but their effects are seen only after 4 to 6 months. Side effects include erectile dysfunction, decreased libido, and, rarely, gynecomastia. 5-ARIs, but not \(\alpha\)-blockers, can alter disease progression as demonstrated by two landmark trials, the MTOPS\(^\text{27}\) and CombAT\(^\text{28}\) trials. These trials evaluated combination therapy using \(\alpha\)-blockers and 5-ARIs. Patients on 5-ARIs, particularly those with larger prostates, had a reduced risk of both developing acute urinary retention and requiring surgical intervention. More recently, daily phosphodiesterase-5 inhibitors, which are most often used for erectile dysfunction (ED), have now been approved for treating patients with BPH. These can be particularly valuable in patients with concomitant ED.\(^\text{29,30}\)

Surgical modalities for BPH continue to evolve towards less invasive endoscopic procedures. Transurethral resection of the prostate (TURP) remains the mainstay of endoscopic procedures, with low treatment failure and complication rates.\(^\text{31}\) TURP syndrome is associated with prolonged use of hypotonic irrigation fluid, resulting in fluid overload and dilutional hyponatremia. Symptoms include nausea/vomiting, bradycardia and hypertension, pulmonary edema, mental status changes, and rarely death. Other endoscopic modalities used today include bipolar TURP and various laser procedures (e.g., Ho:YAG laser enucleation of the prostate, Ho:YAG laser ablation of the prostate, and photoselective vaporization of the prostate) with the goal of enucleating or vaporizing prostatic tissue. Normal saline is used for irrigation with these modalities, which greatly reduces the risk of TURP syndrome. Generally, laser procedures have been associated with shorter catheterization time and length of stay with comparable improvements in LUTS to open prostatectomy or TURP.\(^\text{32-34}\) Open, and more recently laparoscopic and robotic simple prostatectomy can also be performed for patients with moderate-severe, bothersome LUTS due to BPH. These are usually reserved for patients with larger prostatic volumes (>100 cc), or patients requiring concomitant bladder surgery (e.g., bladder diverticulectomy or stones).\(^\text{23}\)

**Urethral Stricture**

A urethral stricture is an area of scarring or fibrosis that causes concentric narrowing of the urethra, impeding the flow of urine as it drains from the bladder. Strictures occur at a prevalence of 0.9% of the population in the United States.\(^\text{35}\) Causes of urethral stricture disease include trauma (19%), iatrogenic causes (33%), inflammatory causes (15%), and idiopathic causes (33%).\(^\text{36}\) Symptoms of urethral stricture disease include incomplete emptying, weak urinary stream, urinary urgency/frequency, and pain.\(^\text{37,38}\)

The anatomy of the urethra in men can be divided into the following segments proceeding from cephalad to caudal: prostatic, membranous, bulbous (the area between the pelvic floor and the penoscoral junction), and penile. A stricture can occur in any segment of the urethra, but it is most common in the bulbar urethra.

Options to treat urethral stricture disease can be divided into two general categories: endoscopic and surgical reconstruction. Endoscopic treatments include a urethral dilation or stricture incision with a cystoscope. The latter is referred to as a direct vision internal urethrotomy. The success rate of one endoscopic attempt to treat a urethral stricture is around 30%.\(^\text{39}\)

The success of repeat endoscopic treatments of a urethral stricture drops to 13%, and recurrent dilations have been associated with the need for more complex reconstructive surgeries for definitive management.\(^\text{39,40}\) For that reason, common practice is to attempt one endoscopic intervention prior to referral for reconstructive surgery.

Surgical reconstruction of the urethra, referred to as a urethroplasty, can be divided into two general categories: excisional and tissue substitution. An excisional repair involves resection of the stricturetured segment of the urethra, and direct anastomosis of the two healthy urethral ends. This repair technique is generally reserved for membranous strictures and short bulbar strictures. Tissue substitution involves augmenting a narrowed urethral lumen with free tissue grafts. The most common tissue substitute is buccal (oral) mucosal graft.

**Other Causes of Obstruction**

Retroperitoneal fibrosis (RPF) is a rare cause of ureteric obstruction secondary to an inflammatory and fibrotic process of the retroperitoneal structures. Most cases (>70%) are idiopathic. Identifiable causes in the remaining cases include periarterial inflammation due to aneurysms, medications (e.g., methysergide, ergot derivatives, \(\beta\)-blockers, phenacetin), infections (e.g., tuberculosis, schistosomiasis), and malignancy (e.g., lymphoma, multiple myeloma, sarcoma). Symptoms are nonspecific and may include general abdominal discomfort or back pain, flank pain due to ureteral obstruction, or lower extremity edema due to vena caval compression. Laboratory abnormalities such as normocytic anemia, an elevated C-reactive protein, or ESR are identified in about two-thirds of cases.\(^\text{41}\) The classic radiological findings consist of a well-defined retroperitoneal soft tissue mass encasing the great vessels with medialization of the ureters. Contrast enhancement on CT scan, magnetic resonance imaging (MRI), and positron emission tomography (PET) scan can also be used to monitor disease activity and assess response to treatment.\(^\text{42}\)

Patients with symptomatic renal obstruction, renal insufficiency, or signs of infection should be decompressed with either ureteral stents or nephrostomy and monitored for postobstructive diuresis. Biopsy of the retroperitoneal mass to exclude malignancy should be considered prior to commencing treatment. Steroid therapy remains the mainstay of medical treatment, although other immunosuppressive agents have been described.\(^\text{43}\) If medical treatment fails, open or minimally invasive bilateral ureterolysis with intraperitonealization or omental wrapping of the ureters is indicated.

Ureteral obstruction secondary to tumor (benign or malignant) is commonly encountered. Ureteral stenting can be tried initially, but it fails in approximately one-half of cases.\(^\text{44}\) Other strategies such as percutaneous nephrostomy, ureteral stenting in tandem, metallic, and metal-mesh stents have been described. Metallic stents may be more cost-effective due to less frequent stent exchanges,\(^\text{45,47}\) although cost savings may be offset by the limited life expectancy in this patient population.\(^\text{44}\)

**GENITOURINARY TRAUMA**

Genitourinary (GU) trauma is rare. Approximately 10% of victims of abdominal trauma will have a urologic injury.\(^\text{48}\) Any portion of the GU tract can be injured including the following: kidneys, ureters, bladder, urethra, and the external genitalia including the testicles. Mechanisms of trauma parallel other injury mechanisms, the majority of which include blunt and penetrating injuries. This section will be divided into the management of each organ involved in the GU system.
The American Association for the Surgery of Trauma (AAST) renal trauma grading system

<table>
<thead>
<tr>
<th>GRADE</th>
<th>DESCRIPTION</th>
<th>MANAGEMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Contusion or nonenlarging subcapsular perirenal hematoma</td>
<td>Generally managed conservatively.</td>
</tr>
<tr>
<td>2</td>
<td>Perinephric hematoma without obvious parenchymal laceration on CT, or a &lt;1 cm laceration into the cortex of the kidney</td>
<td>Generally managed conservatively in a stable patient.</td>
</tr>
<tr>
<td>3</td>
<td>&gt;1 cm laceration into the cortex without involvement of the collecting system</td>
<td>Generally managed conservatively in a stable patient.</td>
</tr>
<tr>
<td>4</td>
<td>A deep laceration into the collecting system with evidence of urinary extravasation on CT, or a segmental renal artery or vein injury with contained hematoma, or partial vessel laceration, or vessel thrombosis</td>
<td>Can be observed expectantly in the stable patient, but may require subsequent urgent or delayed repair. Renal artery embolization may be an option for those who fail conservative therapy.</td>
</tr>
<tr>
<td>5</td>
<td>Renal pedicle injury or multiple deep renal lacerations (“shattered kidney”)</td>
<td>Patients often require surgical exploration, but stable patients with only parenchymal injury may be safely treated conservatively.</td>
</tr>
</tbody>
</table>

CT = computed tomography.

Kidneys

The prime goal of renal trauma management is preservation of renal function. Renal trauma has become largely nonoperative in modern times, especially in the setting of low- to intermediate-grade renal injuries from a blunt mechanism of action. The role of angioembolization through vascular and interventional radiology has further increased this nonoperative management.59,60

The first goal of renal trauma is to accurately grade the renal injury. The gold standard test to diagnose and stage a renal injury includes a CT scan with IV contrast, with delayed images. In most centers, this is referred to as a “CT urogram,” in which delayed contrast imaging delineates the upper urinary tract collecting system. Criteria that would mandate renal imaging include the presence of gross hematuria, microscopic hematuria with hypotension, and mechanisms increasing the prevalence of renal injury (sudden deceleration injuries, flank contusion, etc). The American Association for the Surgery of Trauma (AAST) renal trauma grading system is described in Table 40-1.51,52

The management of renal injuries depends not only on the grade but also on the injury mechanism and clinical symptoms. Absolute indications for surgical or radiological intervention on renal trauma include life-threatening hemorrhage, renal pedicle avulsion, or pulsatile/expanding retroperitoneal hematoma. Moreover, those suffering penetrating renal trauma with a retropertioneal hematoma should undergo exploration when hemodynamic instability exists.

In a hemodynamically stable patient with a renal injury, renal trauma should be initially observed. Data suggests that this approach may even be feasible in the setting of isolated, penetrating renal injuries.53 Conservative management entails bed rest and hemodynamic monitoring. Patients with a grade 4 renal injury (Fig. 40-3A to D) should be treated in the same manner, and a repeat CT scan should be done to make certain that the urinary extravasation has resolved.54 Otherwise, urinoma and subsequent abscess formation may occur. If urinary extravasation is persistent, placement of a ureteral stent or nephrostomy tube should be considered.

Across the board, the most common surgery for renal surgery in modern times is unfortunately a nephrectomy.55 Early renal vascular control may minimize nephrectomy rates.56 This is accomplished by isolating the renal vascular medially prior to opening the perinephric hematoma. If uncontrolled bleeding is encountered once the hematoma is opened, occlusion of the renal vasculature can be performed. At that time, a renorrhaphy can be safely done as can a nephrectomy in the setting of a grade 5 renal injury.

Ureters

There is no association between the magnitude of ureteral injury and the degree of hematuria that is present.57 A high index of suspicion is required. Diagnosis requires either a CT urogram, IVP, or a cystoscopy with a retrograde pyelogram. Unlike renal injury, the ureters more commonly are injured through iatrogenic mechanisms. Common surgical procedures in which the ureters are injured include gynecological, colorectal, and urological surgeries. The repair of ureteric injuries depends on the time of identification from initial injury, location, and length of the injured ureteral segment involved.

Iatrogenic ureteral injuries should be initially managed with ureteral stent placement when possible. When stenting is not feasible, open repair may be attempted when the patient presents shortly after injury. When stent placement is not feasible or when presentation is delayed, nephrostomy tube placement should be considered until formal repair can be safely done.

Ureteral injuries of traumatic origin (penetrating injuries, multiple intra-abdominal traumas) should be repaired during the index admission when possible. Hemodynamically stable patients undergoing laparotomy for other reasons in which a high index of suspicion of a ureteral injury is present should have ureteral exploration. Stable patients in this same situation that are identified to have a ureteral injury should have primary repair at the time of exploration. If a patient is hemodynamically unstable, the ureter can be ligated with subsequent nephrostomy tube placement. Ureteral repair can then be delayed until the patient is stable for surgery.
The definitive operative management of a ureteral injury depends on the location and the extent of devitalization. It is important to debride devitalized ends of the ureter, whether it is from a contusion via a gunshot wound or an iatrogenic thermal injury. Upper ureteral injuries that are short can generally be resected and anastomosed primarily. Ureteral mobilization with preservation of ureteral adventitia to maintain vascular supply can aid in bridging short defects. In modern times, more aggressive maneuvers to directly anastomose more proximal ureteral injuries to the bladder are possible. Maneuvers used to bridge the defect of ureteral length for direct anastomosis to the bladder include the following: bladder mobilization with ligation of the contralateral bladder pedicles, psoas hitch (tacking the bladder down to the ipsilateral psoas tendon), and the Boari flap with downward nephropexy. Creation of a Boari flap utilizes a tubularized flap of anterior bladder wall to bridge long defects. Bridging defects as high as the proximal ureter have been reported in association with this technique. When bladder-to-ureter anastomosis is not possible with these maneuvers, the remaining options include trans-ureteroureterostomy (anastomosing the injured ureter to the contralateral ureter), creation of an ileal ureter, or renal auto transplantation to the pelvis.

**Bladder**

The bladder can be injured through iatrogenic and classic traumatic mechanisms. Indications for bladder imaging include gross hematuria in the setting of injuries with a correlation for bladder injury. The most common clinical scenario is gross hematuria associated with a pelvic fracture, which is associated with a 29% chance of bladder laceration. Diagnosis of bladder injuries requires either a CT cystogram or a fluoroscopic cystogram. The sensitivities and specificities of these two modalities are similar. The bladder should be filled with approximately 300 cc of contrast for either of these imaging modalities.

Contrast may be visible at the site of injury, within the peritoneal space (Fig. 40-4A), or in the perivesical space (Fig. 40-4B). Simply capping the Foley catheter alone on a delayed excretory phase of abdominal CT imaging is insufficient to diagnose a bladder injury.

Two general categories of bladder injuries are extraperitoneal and intraperitoneal injuries. An intraperitoneal injury requires repair during the index admission after the patient has been resuscitated. Delayed repairs are associated with abdominal sepsis. Conversely, extra peritoneal injuries can generally be managed with Foley catheter drainage alone. Situations in which extraperitoneal bladder injuries should be treated with operative repair include complex injuries involving bone spicules from a pelvic fracture within the laceration and concurrent rectal or bladder lacerations, which increase the possibility of fistula formation. Bladder neck injuries should also be treated operatively during the index admission as these injuries occasionally do no heal with Foley catheter drainage alone. Repeat cystography should be done 7 to 14 days later prior to Foley removal to ensure that the laceration, or operative repair, has healed.

**Urethral Injuries**

Common mechanisms of trauma of the urethra include pelvic fracture associated injuries and straddle injuries. Pelvic fracture associated injuries occur at the level of the membranous urethra, whereas straddle injuries occur at the level of the bulb urethra. The clinical hallmark of a urethral injury is blood at the meatus. A retrograde urethrogram should be done when this clinical sign is present to diagnose an injury, prior to attempted Foley catheter placement (Fig. 40-5A).

The initial step in management of a urethral injury is bladder drainage to prevent urinoma formation and subsequent abscess formation. In general, this is accomplished through
placement of an SP tube. After stabilization, some centers perform “primary urethral alignment.” This is a dual antegrade and retrograde endoscopic procedure utilizing fluoroscopy to bridge the urethral defect and to place a Foley catheter across the injury. Subsequent restructure rates are high, but the severity of stricture formation may be less when primary alignment is performed. If patients are managed with an SP tube alone, the site of disruption leaves the patient with a urethral stricture and subsequent restructure. This requires a treatment with a urethroplasty after the patient’s period of convalescence has resolved (Fig. 40-5B).

Penetrating injuries to the anterior urethra are rare. In a hemodynamically stable patient with an uncomplicated injury, it is expert opinion to perform exploration with primary repair during index admission. Complicated injuries with extensive tissue devitalization should be managed with SP tube urinary diversion and delayed reconstruction.

External Genital Injuries
Penile fractures classically occur with excessive torqueing of the erect penis. This excessive torqueing results in rupture of the tunica albuginea, the fascial coating of the erectile bodies. Common symptoms include immediate detumescence with subsequent development of a hematoma. Clinical history and examination alone are sufficient to warrant surgical exploration with primary suture repair of the corporal body laceration. For equivocal cases, ultrasonography or an MRI may be done. Up to 10% of penile fractures are associated with urethral injuries. Blood at the meatus signifies the possibility of a coexisting urethral injury. This should be evaluated with either a retrograde urethrogram or cystoscopy at the time of repair.

Scrotal trauma generally occurs from a blunt mechanism. Injuries to the testis, epididymis, and spermatic cord may occur. Hematomas with subsequent ecchymosis are common with such injuries. Testicular rupture occurs with fracture of the fascial coating of the testicle, called the tunica albuginea. This may occur with blunt or penetrating mechanisms. The most specific findings on ultrasonography are loss of testicular contour and heterogeneous echotexture of parenchyma. The highest reported sensitivity for testicular rupture on ultrasound is 93%. With diagnosis of a testicular rupture or when a high index of

Figure 40-4. Intraperitoneal and extraperitoneal bladder injuries. A. During a computed tomography (CT) cystogram, intraperitoneal contrast is seen within the peritoneal space at the red arrow. B. During a CT cystogram, extravesical contrast is seen contained within the extraperitoneal space at the red arrow.

Figure 40-5. A. Retrograde urethrogram showing an area of narrowing at the double white arrow. This indicates a bulbar urethral stricture. B. After urethroplasty, a retrograde urethrogram demonstrates a normal-appearing and patent bulbar urethra at the arrow.
Acute Urinary Retention

Acute urinary retention (AUR) can happen in men or women and results from a variety of causes, although it most commonly occurs in men with benign prostatic hyperplasia (BPH). Other chronic causes of poor bladder emptying, such as diabetic neuropathy, urethral stricture, multiple sclerosis, or Parkinson’s disease, can result in episodes of complete urinary retention, often when the bladder becomes overdistended. This frequently occurs in the hospital setting when patients have limited mobility and are receiving medications that decrease bladder contractility, including opiates or anticholinergics. Constipation, a common side effect of those medications, can itself worsen urinary retention. Significant hematuria can result in the formation of blood clots, which may block the urethra and cause retention.

Although some patients receiving large doses of narcotics or those with chronically decompensated bladders may not experience discomfort, most patients with AUR have significant pain. Untreated severe urinary retention (often accompanied by overflow incontinence) may result in acute renal failure. Treatment should include placement of a urethral catheter as quickly as possible. However, BPH or urethral strictures often make the placement of a catheter difficult. For men with BPH, a coudé (French for curved) catheter is helpful in negotiating past the angulation in the prostatic urethra (Fig. 40-6A). The curved portion (which is angled in line with the balloon port) is maintained at the 12 o’clock position as it is passed through the urethra (Fig. 40-6B). A common mistake is to use a smaller catheter to bypass the enlarged prostate. However, a larger (18F to 20F) catheter is less flexible and is more likely to push into the bladder rather than curl in the prostatic urethra.

Smaller catheters, however, are quite useful for bypassing a urethral stricture. A urethral stricture should be suspected when the catheter meets resistance closer to the meatus, as many strictures occur in the distal urethra, which is narrower than the proximal portion. Using a 12F or 14F catheter often will allow the passage of the catheter into the bladder. If catheter placement is not successful, a urologic consultation should be requested. The urologist can either choose to (a) use a cystoscope, guidewire, and urethral dilators to dilate the stricture and place a Council-tip catheter via Seldinger technique; or (b) place a suprapubic tube approximately two fingerbreadths above the pubic symphysis. With regard to the suprapubic tube, ultrasound-guidance or aspiration with a fender needle should be used first to localize the bladder and avoid intra-abdominal contents, although bowel injury is unlikely with a distended bladder filling the pelvis. If hematuria is the cause of retention, continuous bladder irrigation often is necessary to prevent clot formation. This is done through a large three-way catheter that has an additional port for fluid inflow. Fluid is infused by gravity only because the use of higher pressure may result in bladder rupture if outflow is occluded.

Once the bladder is adequately drained, the cause of AUR should be addressed. For men with suspected BPH, an α-blocker such as tamsulosin should be started, and these have been shown to increase the likelihood of a successful trial without a catheter. Although finasteride and dutasteride (5α-reductase inhibitors) have been shown to reduce the incidence of urinary retention by 50%, they require several months to take effect and are most beneficial in large prostates; therefore, they will not provide significant benefit in the short term. Narcotics should be tapered as tolerated, and constipation should be treated.

Acute spinal cord compression, which is accompanied by saddle paresthesias, is a neurologic emergency that requires neurosurgical or orthopedic consultation. In most cases, except severe neurologic injuries, patients will be able to resume voiding, and the catheter can be removed after 1 to 2 days. Postvoid residuals should be checked with a portable ultrasound device (bladder scanner) or by “straight” catheterization to determine the residual amount of urine left after the patient tries to empty his or her bladder. In patients with severe liver dysfunction, the bladder scanner may inadvertently misinterpret ascites for urine. The inability to void or the presence of a postvoid residual over 200 mL is concerning for development of another episode of AUR. Patients may be given the option of an indwelling catheter for another few days with a subsequent voiding trial or to perform clean intermittent catheterization (CIC), whereby, after
predetermined intervals (4–6 hours) or after voiding attempts, the patient passes a catheter into the bladder and empties it. This is the preferred method because it reduces the likelihood of infections from indwelling catheters and may improve bladder functionality. However, most patients are resistant to this approach.

**Testicular Torsion**
The differential diagnosis of acute scrotal pain includes testicular torsion.74 This usually occurs in neonates or adolescent boys but may be observed in other age groups. The blood supply to the testicle is compromised due to twisting of the spermatic cord within the tunica vaginalis, resulting in ischemia to the epididymis and the testis. In newborns, an extravaginal torsion also can occur with twisting of the tunica vaginalis and spermatic cord together. Risk factors for torsion include undescended testis, testicular tumor, and a “bell-clapper” deformity—poor gubernacular fixation of the testicles to the scrotal wall.

Clinical history is vital for diagnosis.75 Patients describe a sudden onset of pain at a distinct point in time, with subsequent swelling. Physical examination may demonstrate a swollen, asymmetric scrotum with a tender, high-riding testicle. Children normally have a brisk cremasteric reflex that usually is lost in the setting of torsion. The diagnosis is made by clinical history and examination but can be supported by a Doppler ultrasound, which typically shows decreased intratesticular blood flow relative to the contralateral testis. If an ultrasound is not promptly available, timely surgical exploration should be performed.

Immediate surgical exploration can salvage an ischemic testis.76 At the time of surgery, the contralateral testes also must be explored and fixed to the dartos fascia due to the possibility that the same anatomic defect allowing torsion exists on the contralateral side. Midline (along the median raphe) or bilateral transverse scrotal incisions are made. Once the testis is detorsed, it should be assessed for viability after being given time for normal blood flow to resume. One can assess the blood flow using intraoperative Doppler or by incising the tunica vaginalis and observing tissue viability. The testes are fixed to the dartos fascia with a small, nonabsorbable suture on their medial, lateral, and dependent aspects, taking care to ensure that the spermatic cord is not twisted before doing so. An orchietomy should be performed to avoid later risk of abscess formation only if the testis is clearly necrotic.

**Fournier’s Gangrene**
Fournier’s gangrene is a necrotizing fasciitis of the male genitalia and perineum that can be rapidly progressive and fatal if not treated promptly (Fig. 40-7). The mortality rate has been reported to be as high as 67%.77 Risk factors for Fournier’s gangrene include perirectal abscesses, diabetes, obesity, and chronic alcoholism.78 The often polymicrobial infection spreads along dartos, Scarpa’s, and Colles’ fascia. Clinical signs include perineal and scrotal pain, inflammation, necrosis, and crepitus.78 The diagnosis is largely made on clinical suspicion; however, radiographic findings on CT imaging often assist with the diagnosis, including soft tissue air associated with fluid collections within the deep fascia.79

Prompt and aggressive surgical debridement of nonviable tissue and broad spectrum antibiotics are necessary to prevent further spread (Fig. 40-7A). Fecal diversion with endorectal tubes serve as an option for conservative fecal diversion.80 If there is damage to the external anal sphincter, patients may require a colostomy. Patients frequently require return trips to the operating room for further debridement.

Negative pressure wound therapy systems have been shown to reduce hospitalization time by aiding in wound healing.81 Reconstructive strategies involving skin grafting are needed when large tissue defects result from extensive tissue damage.

**Priapism**
Priapism is a persistent erection for greater than 4 hours unrelated to sexual stimulation.82 Priapism is divided into two types, based on the underlying pathophysiology. The most common type—low-flow/ischemic priapism—is a medical emergency. On examination, the penis is very tender, and both cavernosal bodies will be rigid while the glans will be flaccid. Decreased venous outflow with persistent inflow results in increased intracorporal pressure and tumescence, which is the normal process of erection. Diminished arterial inflow due to elevated
intrapenile pressure usually is brief under normal circumstances. Priapism is essentially a compartment syndrome. With prolonged erection (priapism), the sustained decrease in arterial inflow ultimately causes tissue hypoxia, acidosis, and edema and results in long-term fibrosis and impotence, and sometimes frank necrosis. Risk factors include sickle cell disease or trait, malignancy, medications, cocaine abuse, certain antidepressants, and total parenteral nutrition.82-84 If a cause is not identified, a hematologic workup is necessary to rule out malignancy or blood dyscrasias.

The management of priapism is rapid detumescence with the goal of preservation of future erectile function. The ability to achieve normal erections is directly related to the length of the episode of priapism. Ischemic priapism can be confirmed with a penile blood gas from the cavernosal bodies demonstrating hypoxic, acidic blood. Initial management can include systemic treatment of the underlying disorder (fluid and oxygen for sickle cell patients) but this should be done concurrently with an active treatment to reduce the priapism.82 The initial intervention may be therapeutic aspiration or injection of sympathomimetics (phenylephrine). Insertion of a large-gauge needle (16–21 gauge) into the lateral aspect of one corporal body allows thorough aspiration and irrigation of both corporal bodies because of widely communicating intercavernosal channels. Injection of phenylephrine (diluted 100–500 mcg/mL and given in 1 mL increments every 3–5 minutes for up to 1 hour before determining failure) into the corporal bodies works to cause vasoconstriction, but the patient should be monitored for acute hypertension and reflex bradycardia especially in patients with high cardiovascular risk.

A surgical shunt is sometimes necessary to resolve the episode if phenylephrine fails. Distal (corporoglanular) shunts should be performed first because they are the easiest to perform and the lowest amount of complications. A Winter shunt uses a large biopsy needle to create holes between the glans and corpora; however, if this fails, an operative procedure can be performed to remove the distal tips from each corpora (Al-Ghorab). Proximal shunts such as Grayhack (corporal-saphenous vein) or Quackel (proximal cavernosumspongiosum) shunts may be required in refractory cases.

The other form of priapism (high-flow/traumatic priapism) is rare and is related to penile or perineal trauma resulting in a cavernous artery–corporal body fistula. This form is not painful because it is not related to ischemia and can be managed conservatively with observation. Many cases will resolve with time; those that do not can undergo selective arterial embolization.82

Paraphimosis

Paraphimosis is a common problem that represents a true medical emergency for uncircumcised men. When the foreskin is retracted for prolonged periods, constriction of the glans penis may ensue. This is particularly likely in hospitalized patients who are confined to bed or who have altered mental status and are unable to respond to pain. Delay can be catastrophic as penile necrosis may occur due to ischemia. Penile blocks, pain medication, and sedation are sometimes necessary before manual reduction. It is useful to apply firm pressure to the edematous distal penis for several minutes.85 Although painful, this reduction in penile edema can be the key to success. With the fingers pulling the constricting band distally, the thumbs can push the glans penis back into normal location. Compression wraps have shown some benefit without the need for the physician to use hand compression.86 If the foreskin cannot be manually reduced, surgical intervention is required.

Emphysematous Pyelonephritis

Emphysematous pyelonephritis is a life-threatening infection that results from complicated pyelonephritis by gas-producing organisms. It is an acute necrotizing infection of the kidney that occurs predominantly in diabetic patients.87 Patients frequently present with sepsis and ketoacidosis. *Escherichia coli* appears to be the most frequent organism responsible for this infection. Patients require supportive care, IV antibiotics, and relief of any urinary tract obstruction. Third-generation cephalosporins have been suggested as the initial antibiotic of choice and fluoroquinolones avoided due to high rates of resistance.88 Emphysematous pyelonephritis can be subdivided based on the extent of infection. Cases where gas is isolated to the kidney frequently can be managed conservatively with the placement of a nephrostomy tube to allow drainage of purulent material. When there is extensive involvement of the perirenal tissue, conservative management may not be successful and strong consideration should be given to nephrectomy, particularly if the patient is displaying signs of sepsis.89,90

**Bladder Cancer**

**Epidemiology and Presentation.** In 2018, 81,190 men and women will be diagnosed with bladder cancer, and 17,240 will die from their disease.91 The disease is highly prevalent, with over 700,000 patients living with the disease in the United States as of 2016. Men have nearly three times the incidence of women. Tobacco use is the most frequent risk factor, followed by occupational exposure to various carcinogenic materials such as industrial solvents (e.g., aromatic amines). Other risk factors include arsenic, radiation, cyclophosphamide, and chronic exposure to foreign bodies (stones and catheters) and specific urinary parasites. The most common bladder cancer histology in the United States is urothelial carcinoma (UC), accounting for 90% of tumors, which tends towards a better prognosis as compared to the rarer forms, including squamous cell carcinoma (<10%), adenocarcinoma (1–2%), and small cell cancer (<1%). Unfortunately, there is no reliable screening test for bladder cancer, although patients felt to be at high risk may undergo urine sampling for microhematuria or abnormal cytology. Smoking cessation should be advised in all tobacco users as a preventive measure. The most common symptoms at presentation are hematuria (gross or microscopic) and/or irritable voiding (urgency, frequency, and dysuria). Office cystoscopy is an effective means to diagnose bladder cancer.

**Staging.** Clinical staging is completed with CT or MRI to assess intraabdominal nodal and visceral sites of metastasis. The upper tracts should be evaluated with CT urography or retrograde pyelography. Chest radiograph provides initial evaluation of the thorax and mediastinum. A bone scan should be obtained if the patient complains of bone pain, has known locally advanced or metastatic disease, or an unexplained elevation in the serum alkaline phosphatase level. Pathologic staging has been outlined by the American Joint Committee on Cancer.92

Transurethral resection of bladder tumor (TURBT) should include an examination under anesthesia (EUA) and sampling of the bladder muscular wall to fully assess depth of invasion. The presence of induration or a mass on EUA denotes extravesical
tumor extension and may alter the patient’s treatment plan. It may also be appropriate to biopsy multiple areas of mucosa to identify multifocal carcinoma in situ (CIS). Restaging TURBT within 2 to 6 weeks is recommended in the patient with incomplete, under-sampled, or uncertain resection. This is especially important in the patient with Tis, Ta, or T1 disease, as well as the patient with suspected T2 disease who is being considered for a bladder preservation treatment strategy. Invasion into the lamina propria and certainly the muscular wall demonstrates increased potential for distant metastases; muscle invasion is rarely treated completely with TURBT and requires additional therapy for adequate local control.

Recurrence rates of non–muscle-invasive bladder cancers are high, ranging from 50% to 70%.93 Adjuvant treatment strategies have thus been adopted after TURBT to reduce these rates. Intravesical chemotherapy used in conjunction with TURBT can reduce the risk of recurrence by 44% to 73% in patients with primary Ta and T1 tumors and by 38% to 65% in patients with recurrent Ta, T1, and Tis tumors when compared to TURBT alone.94 Intravesical immunotherapy using bacillus Calmette-Guérin (BCG) also provides a significant reduction in recurrence that is greater than 50% in this population. Despite improved rates of disease-free survival, standard induction courses of intravesical chemotherapy and immunotherapy do not improve disease-specific survival.94 However, when an induction course of BCG is followed by a series of maintenance doses consisting of weekly BCG given for 3 weeks at 3, 6, 12, 18, 24, 30, and 36 months after induction, disease-free and overall survival can be prolonged.95 In patients who fail an initial or maintenance course of intravesical therapy, it may be reasonable to try another agent; however, one must consider the risk of progression and not delay definitive treatment. Roughly 15% to 30% of patients presenting with non–muscle-invasive tumors will eventually progress to muscle invasion. Radical cystectomy remains the most effective single-modality treatment for patients with muscle-invasive bladder cancer, refractory high-risk non–muscle-invasive disease, and especially lymph node–negative disease with a reported 10-year recurrence-free survival of organ-confined lymph node–negative (<pT2N0) disease between 69% and 87%.94,95,96,97

**Surgical Considerations.** Cystectomy is indicated in the treatment of refractory NMIBC or to assert local control for muscle invasive bladder cancer (MIBC).98 Effective local control in the pelvis is achieved in 93% of cases with cystectomy. Indications for partial cystectomy are limited and generally apply to isolated tumors or those within diverticulum. Classic teaching suggests that patients with CIS should not be candidates, though the use of intravesical BCG to treat CIS may have broadened this application. For patients with MIBC, neoadjuvant systemic chemotherapy with M-VAC or gemcitabine and cisplatin (prior to cystectomy) offers a survival advantage when compared to radical cystectomy alone.98

Robotic approaches for cystectomy are increasingly used, but the urinary diversion is still usually performed through an open incision. The benefits of the robotic portion are decreased blood loss during the pelvic dissection (due to the pneumoperitoneum). However, recent evidence (randomized controlled trials of open vs. robot-assisted radical cystectomy) did not demonstrate any difference in oncologic efficacy or complication rates.

Complications of bladder cancer surgery involve bladder perforation during transurethral resection of the bladder tumor, which require catheter drainage for several days if small (common) or open repair if large and intraperitoneal (rare). Cystectomy and urinary diversion may result in prolonged ileus, bowel obstruction, intestinal anastomotic leak, urine leak, or rectal injury. A urine leak from the ureteroileal anastomoses is a common cause of ileus, intra-abdominal urinoma, abscess formation, and wound dehiscence. Deep venous thrombosis is common after cystectomy due to the advanced age of most patients, proximity of the iliac veins to the resection and lymph node dissection, and the presence of malignancy. The utility of subcutaneous heparin in the perioperative period can minimize the risk of venous thromboembolism. Contemporary series from high volume centers report readmission rates of 25%, complication rates of 50% to 60%, and perioperative mortality in the first 90 days at 5% to 10%.99,100

Urinary diversion can be accomplished using an incontinent or continent abdominal stoma or orthotopic continent reconstruction. The evolution of patient selection and surgical technique has led to improved outcomes for orthotopic diversion, although there are still patients who are better served with an ileal conduit. Motivated patients are considered for orthotopic neobladder diversion if they have a preoperative serum creatinine less than 2.0 mg/mL, normal preoperative bowel function, a negative urethral margin based on intraoperative frozen section at the time of cystectomy, and an intact sphincter after complete tumor resection.

Alternatives to cystectomy include observation, systemic chemotherapy, radiation therapy, or a combination of chemotherapy and radiation. These modalities may be required in patients who a poor surgical risk, who refuse surgery, or who are elderly.

Bladder preservation using radiation as the definitive therapy may be feasible in selected patients. In this context, modality therapy is preceded by aggressive TURBT and offers an improved rate of survival when performed in conjunction with chemotherapy. Up to 42% 5-year disease-specific survival can be achieved in patients with preserved bladders, with the best overall survival outcome in younger patients with lower stage tumors without lymphovascular or nodal involvement.

More recently, immunotherapeutic treatments have shown significant promise in the treatment of locally advanced and metastatic bladder cancer. Five agents have recently been approved for patients who have progressed on or after platinum-based chemotherapy or have progressed within 12 months of neoadjuvant or adjuvant treatment. These agents include PD-L1 inhibitors (atezolizumab, avelumab, durvalumab) and PD-1 inhibitors (nivolumab and pembrolizumab). Response rates for these agents are ~15% to 20% but may have extended median overall survival as much as 10.3 months when compared to chemotherapy.101,102

**Testicular Cancer**

Testicular cancer is the most common cancer in men age 20 to 40 years and the second most common cancer in young men age 15 to 19 years. Metastases to the testis (usually lymphoma in older men) are rare. In 2018 there were 9310 new cases and 400 deaths from the disease.93 The incidence of testis cancer varies around the world.103 It contains a heterogeneous group of tumors, of which 95% are germ cell tumors; the rest originate from stromal cells (Leydig or Sertoli cells). Germ cell tumors can be classified as either seminomatous or nonseminomatous. Seminoma constitutes more than 50% of all testis cancer. The
Stages I to IIA nonseminomatous testis cancer is potentially cured with RPLND or chemotherapy.\textsuperscript{111} Persistently high tumor markers after radical orchiectomy or high-stage metastatic germ cell tumors warrant systemic chemotherapy. Due to the high rates of teratoma or viable germ cell tumor, postchemotherapy bulky masses are resected by RPLND or other surgical procedures. The overall survival rate of localized disease is outstanding (99\% at 5 years). Patients with more advanced distant metastatic disease (stage III) have 75\% survival rates. The overall prognosis is generally better for seminomatous than nonseminomatous germ cell tumors.\textsuperscript{112}

**Surgical Considerations.** Radical orchiectomy is done through an inguinal incision extending from the external inguinal ring to the internal inguinal ring. The spermatic cord is ligated at the internal ring with long silk sutures for easier identification during a future RPLND. Integrity of the scrotal skin during orchiectomy is important. Complications of radical orchiectomy include scrotal hematoma, chronic pain, and hernia.

For RPLND, a midline incision is usually made from the xiphoid process to the pubic symphysis. All the lymphatic tissue is removed from the targeted areas using the classical split and roll technique, and all lumbar vessels are tied. Postganglionic sympathetic nerve sparing is possible in most cases for preservation of ejaculatory function.\textsuperscript{113} Robotic-assisted RPLND is growing, with faster recovery time and similar short term oncologic results.\textsuperscript{114} Complications after RPLND include bowel obstruction, excessive bleeding, chylous ascites, and ejaculatory dysfunction.

**Kidney Cancer**

Renal cell carcinoma (RCC) results in approximately 3.8\% of all new cancers, with an estimated 65,340 new cases and 14,970 deaths related to kidney cancer in 2018.\textsuperscript{91} Despite several advancements with immune-based and targeted molecular therapies demonstrating durable clinic responses, RCC still remains primarily a surgical disease and classically does not respond to conventional chemotherapy regimens or radiation therapy.

Most patients diagnosed with RCC in the modern era typically present with an incidentally discovered renal mass on abdominal radiographic imaging. Differential diagnosis of a renal mass includes malignant tumors (e.g., RCC, urothelial carcinoma, sarcomas, lymphoma, metastasis), benign tumors (e.g., cysts, angiomylipoma, oncocytoma), and inflammatory lesions (e.g., abscesses, xanthogranulomatous pyelonephritis, tuberculosis). Renal CT imaging with intravenous contrast remains the single most important radiographic test to delineate the nature of the mass. In general, any solid renal mass that enhances by more than 15 Hounsfield units is an RCC until proven otherwise. However, even if there is contrast enhancement on axial imaging, approximately 15\% to 30\% of solid renal masses are benign on final surgical pathology.\textsuperscript{115} Renal tumor biopsy can help distinguish between malignant or benign tumors, but this has not been widely adopted by the urological community, despite series showing their high diagnostic yield, concordance with surgical pathology, and safety.\textsuperscript{116-118} Biopsy remains particularly useful in patients considering surveillance or thermoablative therapy, or in patients with suspicion of metastasis or lymphoma.

Major recognized risk factors for RCC include smoking, obesity, and hypertension. Although most RCCs are discovered incidentally, some patients present with signs or symptoms
which may be the result of local tumor growth (e.g., flank pain, hematuria, perirenal hematoma), paraneoplastic syndromes (e.g., hypertension, weight loss, hypercalcemia, polycythemia/anemia, abnormal liver function tests), or metastatic disease. RCC metastasizes primarily to the lungs, lymph nodes, bone, liver, adrenal glands, and brain. Familial RCC subtypes with classical clinical manifestations are also well described. The von Hippel-Lindau disease, occurring as a result of a mutation in the tumor suppressor gene VHL (3p25-26), commonly manifests itself with clear cell RCC, pheochromocytomas, retinal angiomas, central nervous system hemangioblastomas, pancreatic cysts, and other tumors. Other familial syndromes include hereditary papillary RCC (papillary type 1 RCC), familial leiomyomatosis (papillary type 2 RCC), and Birt-Hogg-Dube syndrome (chromophobe RCC, hybrid oncocytic tumors, and oncocytoma). Familial RCC syndromes should be suspected in younger patients and patients with multicentric and/or bilateral tumors.

Clear cell RCC is the most common subtype, accounting for 70% to 80% of all RCCs. Papillary RCC occurs in 10% to 15%, type 1 being associated with a better prognosis, and type 2 a worse prognosis. Other subtypes include chromophobe RCC, collecting duct carcinoma, and unclassified type.

RCC may locally progress and cause invasion of the renal capsule and perirenal fat or the collecting system. RCC may also directly progress into the venous system in the form of a tumor thrombus that can extend into the IVC and into the right atrium. Staging is the single most important prognostic factor for RCC. Studies demonstrate a 70% to 90% 5-year survival rate for organ confined disease (stages I–II), compared to 0% to 10% for patients with systemic metastases (stage IV). Other important prognostic factors include histological subtype, tumor size, lymph node involvement, and site of metastases.

Management options for small renal masses (<4 cm) includes active surveillance, thermoablative techniques, or surgical excision (Fig. 40-9). Percutaneous or laparoscopic thermoablative techniques (cryoablation, radiofrequency ablation, high-intensity focused ultrasound) have been used to treat small renal masses, but they are associated with an increased risk of local recurrence.

Since the first laparoscopic radical nephrectomy described by Clayman et al in 1991, minimally invasive surgical approaches, including laparoscopy with robotic assistance, have virtually supplanted open procedures for localized RCC (Fig. 40-10). Partial nephrectomy is most appropriate for patients with small tumors, solitary kidney, bilateral tumors, or familial RCC. Some tumors may not be amenable to ablative therapies or partial nephrectomy, in which case radical nephrectomy would be employed.

Radical nephrectomy involves removal of the entire kidney with dissection external to Gerota’s fascia. The colon is retracted medially after incising the white line of Toldt, followed by meticulous hilar dissection with ligation of the renal artery and vein. The adrenal gland is usually spared unless the tumor involves the gland or is immediately adjacent to it. Lymphadenectomy remains controversial, and it is usually performed in patients with adenopathy on preoperative imaging or in patients with palpable lymph nodes intraoperatively. In partial nephrectomy, renal artery clamping is often performed to minimize blood loss while the tumor is excised. The goal is to remove the tumor with negative surgical margins while minimizing warm ischemia time to preserve as many functional nephrons as possible. With increasing experience, partial nephrectomy is now also performed on much more complex renal masses, including completely endophytic, central, and hilar tumors. Very large tumors or tumors with vena-caval thrombi can be removed robotically in experienced hands, but most are still removed using an open approach.

In minimally invasive surgery, both partial and radical nephrectomy can be done via either a transperitoneal or retroperitoneal approach. In open cases, a subcostal flank approach provides direct access to the retroperitoneum and is preferred for lower pole exposure, but it can limit access to the hilum, particularly with large renal masses. The anterior subcostal approach is preferred for larger renal masses. Bilateral anterior subcostal incisions (chevron incision) provides excellent vascular exposure (e.g., IVC thrombectomy, bilateral tumors). Midline incisions are usually reserved for renal trauma and for reconstructive procedures. Less commonly performed, the thoracoabdominal approach involves access usually above the 10th rib and is used for large upper pole or adrenal masses, IVC thrombectomy, or tumors involving adjacent structures. Complications include injury to adjacent organs, and for partial nephrectomy, pseudoaneurysms/arteriovenous fistula formation and delayed urinary leak.
Prostate Cancer

Prostate cancer is the most common noncutaneous cancer in men; 164,690 new cases of prostate cancer were diagnosed in 2018 and 29,430 men died from their disease. Screening for prostate cancer with detailed history, digital rectal examination, and serum prostate specific antigen (PSA) tests have changed the natural history of the disease. Since the introduction of prostate cancer screening in the mid-1980s, the incidence of metastatic prostate cancer has decreased by half. Currently 99% of newly diagnosed patients will survive more than 10 years.91

While early screening for African American patients or patients with a family history of prostate cancer is widely accepted, screening for all men is more controversial. Despite data from large randomized clinical trials showing a decrease in mortality after prostate cancer screening, the U.S. Preventive Services Task Force recommended against the routine use of prostate cancer screening.129 Its recommendation was based on the harm and toxicity of overtreatment of nonlethal disease.130 The American Urologic Association subsequently recommended informed and shared decision-making and screening for high-risk disease for men between the ages of 55 and 69 with a life expectancy more than 10 years.131

If the digital rectal examination is abnormal or if the PSA level is above expected for patients’ age and size of the prostate, a prostate biopsy is usually performed. Newer tests such as the 4K score, prostate health index, and PCA3 are sometimes used to inform the decision to proceed with biopsy. Recently, MRI fusion transrectal ultrasound-guided biopsy improved the accuracy of prostate biopsy.

Since most patients survive the disease, risk stratification systems are routinely utilized to guide staging and treatment. Clinical TNM stage, serum PSA levels, and the Gleason grading system are utilized in clinical practice. More recently, genetic testing on biopsy specimen was included in national guidelines. Historically, the Gleason scoring (GS) system included a primary and secondary score based on the most common and second most common histologic patterns. Grades range from 1 for the most differentiated to 5 for the least. The grades are added to create a resultant Gleason score.132 However, since no patients are assigned a score of less than 5 anymore, the grading system has been modified to a scale from 1 to 5. Grade one includes a GS of 3 + 3 = 6 or less, grade 2 for GS 3 + 4, grade 3 for GS 4 + 3, grade 4 for GS 4 + 4 and grade 5 for Gleason score of 9 or 10.133 Imaging studies like CT and bone scans are used to rule out metastatic disease in high-risk patients. The two most common sites of metastatic disease are pelvic/retroperitoneal lymph nodes and boney structures. Modern CT PET scans have a limited role at this point.

Treatment for localized prostate cancer is guided by cancer aggressiveness and patient’s preferences. Active surveillance is recommended for patients with low-risk disease grade 1–2, early-stage disease (cT1c), and small volume disease as determined by biopsy. Large prospective cohorts and randomized clinical trials have established the safety of this approach.134,135 The risk of progression to metastatic disease with close follow-up and repeat prostate biopsies is less than 2% in over 12 years. Radical prostatectomy and pelvic lymph node dissection (robotic, laparoscopic, or open), image modulated radiation therapy (IMRT), and brachytherapy are the standard of care for curative treatments. All provide equal cancer specific survival for low and intermediate risk cancers. For higher risk prostate cancer patients, both surgery and IMRT with androgen deprivation therapy provide excellent cancer control. Cryotherapy, or high intensity focused ultrasound (HIFU) and focal therapy are emerging options that may be acceptable for some patients with low-risk disease.

Level I evidence has established the role of adjuvant radiation therapy after radical prostatectomy for patients with positive surgical margins, extracapsular extension, and high-grade disease.136,137 After definitive treatment of localized prostate cancer, rising PSA is an extremely reliable indicator of recurrence or progression. However, it may take over 10 years for metastasis to appear on imaging studies.138 Once prostate cancer metastasizes, it is no longer curable. Medications that lower serum testosterone or androgen receptor blockers are able to control the disease, often for years. In addition, chemotherapy, immunotherapy, and radioisotope therapy at different stages of the disease increase the life expectancy of the patients or improve the quality of life. The cancer inevitably becomes resistant to these treatments. Nevertheless, patients with incurable prostate cancer can live many years, and a large number die of causes other than prostate cancer.

Over the past few years, we have witnessed major developments in the management of metastatic castrate resistant prostate cancer (mCRPC). New agents that interrupt androgen synthesis (e.g., abiraterone acetate)139,140 and new modulators of androgen receptors (e.g., enzalutamide)141,142 have significantly improved the life expectancy of patients with both androgen sensitive and resistant metastatic prostate cancer. Similarly, innovations in immunotherapy and chemotherapy delivery have advanced the management of advanced prostate cancer.

Surgical Considerations. Open radical retropubic prostatectomy is done through a lower midline incision from below the umbilicus to the pubic symphysis. After entering the space of Retzius, the external iliac, obturator, and internal iliac lymph nodes are removed. The cavernosal nerves located on the posterolateral surface of the prostate capsule are usually spared on the side(s) with low risk of extracapsular extension of the disease. Then the prostate is removed in a retrograde fashion, and the urethrovessical anastomosis is completed in an interrupted fashion.

Robotic radical prostatectomy using the da Vinci robotic surgical system (Fig. 40-11) is now the most common technique (over 90% of all patients in the United States) for the
surgical treatment of localized prostate cancer. Robotic surgery has lower blood loss and faster convalescence, less bladder neck contracture, and lower early postoperative complications. Some data show a faster return of continence and lower rates of erectile dysfunction. The most common postoperative complications include infection, urine leaks, ileus, lymphocele, and, very rarely, rectal or ureteral injury.

To minimize the impact of these side effects, researchers have used different ablative techniques to obliterc the areas of significant cancer. By avoiding the need for whole gland radiation or removal, these focal ablative therapies aim to balance the long-term impact on quality of life with survival. Laser, high-focused ultrasound, cryotherapy, and photodynamic ablations have showed similar results in early studies.

Urethral Cancer

Urethral carcinoma (UC) is a rare disease, the true incidence of which is unknown. It accounts for less than 1% of genitourinary cancers.143,144 It is a disease of the older adult. Risk factors include chronic inflammation from sexually transmitted diseases (human papillomavirus 16 and 18 in squamous cell carcinoma),145 chronic urethral stricture, and indwelling catheterization. Furthermore, urethral diverticulum and recurrent urinary tract infections increase the risk for women.

The majority of patients present with irritative and obstructive voiding symptoms, bleeding, or a palpable mass. Urothelial carcinoma is the most common histology; 29% of women have adenocarcinoma, and both genders can have squamous cell carcinoma. Untreated or refractory UC typically metastasizes through lymphatic channels to the inguinal and pelvic lymph nodes and hematologically to distant organs. Cytoscopic biopsy establishes the diagnoses. An MRI of the pelvis is extremely helpful for defining local extension of the disease while CT scans of the chest, abdomen, and pelvis identify metastatic disease. Finally, it is also important to evaluate the entire urinary tract.

The 5-year overall survival rates for distal urethral tumors is significantly better than for proximal cancers, 68% versus 40%, respectively.143,146 The median 5-year cancer-specific survival is approximately 46%.144 Prognosis is dictated by patients’ age, race, clinical stage, and location of the tumor.

If feasible, local endoscopic resection for low-volume, low-stage disease is preferable. Adjuvant intravesical instillation of Bacillus Calmette-Guérin (BCG) should be considered for patients with proximal noninvasive disease.147 Due to the paucity of robust data, management of locally advanced disease is more challenging. Either radical cystectomy or radiotherapy are acceptable options. Unfortunately, local recurrence rates are high after aggressive monotherapy (63%).148 More recent data support the use of multimodal therapy.149,150 Small series of combinations of perioperative chemotherapy, surgery, and radiation indicate the best cancer control.151

COMMON UROLOGIC CONDITIONS

Urinary Incontinence and Voiding Dysfunction

Urinary incontinence is defined as the involuntary loss of urine. This is more common in women than men for a variety of reasons, including anatomic differences such as a shorter urethra and risk factors such as childbirth. Many patients may also suffer from bothersome symptoms without leakage of urine such as overactive bladder (frequency and urgency of urination and often nocturia), or obstructive symptoms such as hesitancy, weak stream, and incomplete bladder emptying. These conditions can have a negative impact on quality of life,152–154 but they are also associated with serious health issues, including depression, anxiety, social isolation,155 and even falls and fractures in the elderly.156

Urinary incontinence can be divided into several categories, although patients (particularly women) may suffer from more than one type.157 Urge incontinence is the involuntary loss of urine associated with an urge to void. Stress leakage occurs with increases in intra-abdominal pressure, such as coughing or sneezing, and may relate to loss of sphincteric function, urethral hypermobility from pelvic floor laxity (often related to parity), or following prostate surgery in men. Overflow incontinence occurs in the setting of obstruction, with urine leakage occurring with movement causing overflow of urine from a distended bladder. Genitourinary fistulas typically result in the most severe form of incontinence with constant leakage of urine regardless of presence or absence of activity or movement. Examples include vesicovaginal or ureterovaginal fistulae most often due to gynecologic surgery, or rectourethral fistulae in men from cancer, radiation, or surgical intervention.

For urinary incontinence and voiding dysfunction, treatments vary depending on the etiology, severity, and bother of the symptom. Urge leakage and overactive bladder can be treated by (a) behavioral modification (timed voiding, adjustment to fluid intake, timing of diuretic medication, and improved constipation); (b) bladder retraining (pelvic floor physical therapy158); (c) medications (anticholinergics159 and β-3 agonists160); or (d) minimally invasive procedures (sacral neuromodulation,162 percutaneous tibial nerve stimulation,162 or bladder chemodenervation with detrusor botulinum toxin injection163).

Stress incontinence in women can be addressed by pelvic floor strengthening exercises, vaginally placed removable support with a pessary, injection of urethral bulking agent, or sling procedures using polypropylene mesh or autologous tissue. In men, stress leakage is due to either iatrogenic causes or neurologic disease. Treatments include strengthening exercises as in women, slings, or implantation of an artificial urinary sphincter. Overflow incontinence treatment is directed at the cause of obstruction, often benign prostatic enlargement in men, with bladder drainage, medications such as α-blockers or 5-α-reductase inhibitors, or surgical removal of the obstructing gland. When fistulas are present, adherence to surgical principles such as tension-free multilayer closure, nonoverlapping suture lines, and tissue interposition when possible offers the highest likelihood for success.

Erectile Dysfunction

Erectile dysfunction (ED) is defined as the inability to achieve and maintain an erection adequate for sexual intercourse. Formerly, this was known as a type of sexual dysfunction, but it is now understood that ED may be an early symptom of cardiovascular disease due to endothelial dysfunction. ED is a common disease for men later in life with a prevalence rate believed to range anywhere from 30% to 50% depending on age. Two large population-based studies, the Massachusetts Male Aging Study (MMAS) and the European Male Aging Study (EMAS), examined men age 40 to 79 years and found that ED rates increased with age.164,165
Erections are triggered via sexual stimulation setting off a cascade of events. Nitric oxide is released from nerve fibers and activating guanylyl cyclase leading to an increase in cyclic guanosine monophosphate (cGMP). The cGMP pathway leads to smooth muscle relaxation within the corpora cavernosa allowing blood to fill the lacunar spaces. Once the lacunar spaces are full, the expanded tissue compresses the subtunical venules thereby trapping blood within the penis and blocking venous outflow. Phosphodiesterase type-5 hydrolyzes cGMP to reverse the process.166

There are multiple mechanisms leading to ED including vasculogenic, neurogenic, iatrogenic, and psychologic, but often it is multifactorial. Vasculogenic ED can be a result of cardiovascular disease and endothelial dysfunction leading to cavernosal artery insufficiency. Diseases such as hypertension (odds ratio [OR] 1.35–3.04), diabetes (OR 2.57), dyslipidemia (OR 1.83), and tobacco abuse (OR 1.4) all may increase the risk for ED.167 Nerve injuries due to diseases (diabetes, Parkinson’s, multiple sclerosis, spinal cord injury) or surgery (radical prostatectomy, abdominoperineal resection, and other radical pelvic procedures) can lead to interruptions in the nerve signaling that causes nitric oxide release and therefore lead to ED. Iatrogenic causes may be a result of surgery (described earlier) or medication use, as in some antihypertensives, opiates, androgenics, and psychotherapeutics.168 Psychogenic ED, a common reaction to stress and anxiety, is a result of noradrenaline release causing smooth muscle contraction and thereby inhibiting erections.169

Treatment for ED begins with lifestyle modification by identifying any reversible risk factors such as stress/anxiety, medications, unhealthy diets, lack of exercise, and tobacco abuse.170 Medical therapy then begins with the use of phosphodiesterase type-5 inhibitors (PDE5i). These work by prolonging the activity of cGMP, leading to continued smooth muscle relaxation allowing more blood inflow into the penis. Common drugs include sildenafil, tadalafil, vardenafil, and avanafil. They differ in time to peak concentration (lowest in avanafil, sildenafil, and vardenafil), half-life (highest in tadalafil), and the impact of lipids in foods (sildenafil and vardenafil must be taken on an empty stomach). Common side effects include a headache, heartbeat, facial flushing, nasal congestion, and myalgias.171 Patients on nitrate-containing medications should not be given PDE5i due to the risk of severe hypotension. Vision related conditions like macular degeneration, retinitis pigmentosa, and nonarteritic anterior ischemic optic neuropathy are cause for increased awareness and possible ophthalmologic consult.172

Second-line options for ED include vacuum erection devices (VED), intracavernosal injections (ICI), and intraurethral suppositories. The VED is a mechanical device composed of a cylinder placed around the penis which then uses a vacuum to create negative pressure and pull blood into the penis. In order for blood to stay in the penis after the vacuum is released, a tight constriction band must be placed at the base of the penis. There is poor compliance due to difficulty with use and the common reactions of petechia, temporary paresthesia, color changes, and the penis being cold to touch.173,174 Alternatively, ICI uses vasoactive substances (prostaglandin E1 [alprostadil], papaverine, and phentolamine) either alone or in combination to trigger the erection cascade.168 Patients are trained to give themselves a self-injection when they want an erection, and it takes approximately 5 to 15 minutes until they are fully rigid if they respond. With ICI, there is greater concern for prolonged erection or priapism, so dose titration must be closely monitored. Intraurethral suppositories are composed of alprostadil in the form of a pellet which is then placed in the urethra and massaged for absorption. With suppository use, there are concerns about efficacy (only 46–65%) and compliance due to a burning sensation that limits the interest of some users.175,176

Third-line treatment of ED is with surgery placement of a penile prosthesis. There are three main types (malleable, two-piece, and three-piece). The malleable device does not inflate/dilate and merely bends in and out of position for intercourse. The two-piece and three-piece devices are inflatable and differ on the presence of a separate fluid reservoir. The two-piece device has the fluid maintained in the lower half of the penile cylinders, whereas, the three-piece device has a fluid reservoir placed in the pelvis or abdominal wall (Fig. 40-12). Overall, the inflatable prosthesis has high patient and partner satisfaction rates, >92% and >91%, respectively.177

**PEDIATRIC UROLOGY**

**Hypospadias**

Hypospadias, a condition which may be considered a form of incomplete maturation of the genitalia, is a common abnormality that occurs in 1 out of 250 to 300 newborn boys. The most obvious aspect of hypospadias is a urethral opening that is not at the tip of the glans, and 70% to 80% of affected babies will have a meatus on the mid to distal shaft or proximal glans. A lesser number will have more proximal openings, whether penoscrotal, scrotal, or perineal. In addition to an abnormally located meatus, boys usually have deficient ventral foreskin. Associated penile
curvature, more common in the severe varieties, is referred to as chordee.

No diagnostic studies are needed for the majority of boys with hypospadias as there is typically no increased risk of renal or bladder abnormalities. Children with associated cryptorchidism, especially with proximal hypospadias and a nonpalpable testis, have an increased risk of a having a coexisting disorder of sexual differentiation (DSD) and need to undergo a thorough evaluation including hormonal studies, karyotype, and pelvic ultrasonography.178

Distal hypospadias can usually be repaired in one stage with success rates of greater than 95%. Most would advocate a staged approach to proximal hypospadias with correction of penile curvature at the first stage and formal urethral reconstruction at the second.179 Adults with corrected hypospadias usually have normal sexual function and fertility.

**Urinary Tract Infections in Children**

Urinary tract infections (UTI) are common in children, and there is a greater chance of underlying anatomic abnormalities. Children may have conditions such as vesicoureteral reflux, ureteropelvic junction obstruction, ureterocele, or ectopic ureters as causes of these infections. Because of this association, in the past all children with febrile infections would undergo complete evaluations including renal ultrasonography (US) as well as invasive studies such as voiding cystourethrography (VCUG). However, defining pyelonephritis as having a positive renal cortical scan, only 30% to 40% of children with febrile UTI will have reflux. Thus the majority of children with febrile infections, and a greater percentage of those with afebrile infections (cystitis), will be anatomically normal.180 These data have led to a change in imaging guidelines for children with UTI.

Guidelines put out by the American Academy of Pediatrics have markedly changed the way children with infections are evaluated.181 These guidelines suggest that infants less than 2 months of age with febrile infections should undergo both a renal US and VCUG. Children between 2 months and 2 years who have their first documented infection only need to have a renal ultrasound performed. A VCUG is only needed if there are abnormalities detected on the ultrasound such as hydronephrosis, scarring, or other evidence of anatomic abnormality. A VCUG may also be performed if a child has recurrent infections despite empirical treatment. These guidelines do not address children older than 2 years of age but one can assume that similar algorithms of treatment would be appropriate.

There is now greater understanding that most children with UTIs, whether pyelonephritis or cystitis, have some element of bladder and/or bowel dysfunction as the major factor in the development of the infection. Thus, all children with UTIs need to have a thorough assessment of daily bladder and bowel habits. The latter may be difficult to ascertain in younger children, but bowel dysfunction, even subclinical, may be the most important factor in the development of UTIs. Behavioral therapies such as regular and complete voiding in conjunction with a bowel program should be considered the mainstay of the prevention of infections as opposed to prophylactic antibiotics.

**Prenatal Hydronephrosis**

Antenatal imaging will show hydronephrosis in nearly 1% of all babies. Though the majority of children will have benign hydronephrosis of no clinical significance, it may also be related to vesicoureteral reflux, ureteropelvic junction obstruction, ectopic ureter/ureterocele, and other upper tract abnormalities. Typically, nothing needs to be done for these children until after birth, at which point a baseline renal ultrasound can be performed. Other studies such as a VCUG or Lasix renal scans can then be done depending on the degree of dilation. Diagnosis of upper tract obstruction is usually based on progressive worsening of dilation or renal function on serial examinations.

Special consideration must be given for children with bilateral hydronephrosis or hydronephrosis associated with a solitary kidney, especially if linked to oligohydramnios. Since fetal urine production accounts for much of the amniotic fluid, low levels can be a sign of a severe abnormality of the urinary tract. Reduced amniotic fluid is of great consequence since normal lung development is dependent on normal amniotic fluid volumes and children with oligohydramnios can be born with significant pulmonary insufficiency. Boys with bilateral hydronephrosis and low amniotic fluid are at high risk for having posterior urethral valves (PUV). Boys with PUV have as much as a 25% risk of developing end stage renal disease at some point in their lives.182 Prenatal intervention such as placement of vesicoamniotic shunts have not been shown to reduce the risk of renal failure.

**Cryptorchidism**

Cryptorchidism or undescended testes (UDT) is a common condition occurring in 3% of full term and 30% of premature babies. Many of these testes will descend spontaneously due to the normal gonadotropin release that occurs in the first few months of life, so the true incidence is roughly 1% of boys. Untreated cryptorchidism will lead to testis damage, and there is evidence that permanent changes may occur by 3 years of age. Ideally, surgical treatment should occur prior to this age. UDT is usually an isolated finding, but it may occur as a part of a systemic condition such as Prader-Willi, Eagle-Barrett, or other such complex multisystem syndrome. Surgery is the treatment of choice; hormonal treatment has no role.

The consequences of untreated cryptorchidism include infertility and malignant degeneration. One study on fertility suggested that men with a history of unilateral cryptorchidism will have no difference in paternity rates compared to normal controls. In contrast, men with bilateral cryptorchidism have up to a 50% rate of infertility.183 There is data to suggest that orchidopexy in the first year of life is associated with better total sperm counts in adulthood.184 With regard to malignancy, untreated UDT has a fivefold increase risk of tumor development compared to the normal population. However, there is data to suggest that prepubertal orchidopexy is protective and that these boys only have a twofold greater risk.185

**REFERENCES**

Entries highlighted in bright blue are key references.


PATHOPHYSIOLOGY AND MECHANISMS OF DISEASE

The female reproductive system includes the external (vulva including the labia, clitoris, and vaginal opening) sex organs as well as the internal organs (uterus and cervix, fallopian tubes, and ovaries) that function in human reproduction. The female reproductive tract has a multitude of tightly regulated functions. The ovaries produce the ova (egg cells) and hormones necessary for maintenance of reproductive function. The fallopian tubes accommodate transit of an ovum to the uterus and provide a location for fertilization. The uterus accommodates an embryo that develops into the fetus. The cervix provides a barrier between the external and internal genital tract. Ongoing activities, such as angiogenesis and physiologic invasion, are necessary in order for the reproductive organs to fulfill their purpose and are usurped in disease. Immune surveillance is regulated in a fashion that allows implantation, placentation, and development of the fetus.

Because the pelvis contains a multitude of spatially and temporally varied functions, pathologies range from mechanical events, such as ovarian torsion or ruptured ectopic pregnancy, to infection, such as pelvic inflammatory disease, to mass effects, including leiomyomata and malignancy, that can present with similar and even overlapping symptoms and signs. An acute abdomen presentation in a woman of child bearing potential can range from pregnancy-related catastrophes, to appendicitis, to a hemorrhagic ovarian cyst.

The ongoing rupture, healing, and regrowth of the ovarian capsule and endometrium during the menstrual cycle use the same series of biologic and biochemic events that are also active in pathologic events such as endometriosis and endometriomas, mature teratomas, dysgerminomas, and progression to malignancy. Genetic abnormalities, both germ line and somatic, that may cause competence and/or promote disease are increasingly well understood. Incorporation of genetic and genomic information in disease diagnosis and assessment has altered how we diagnose and follow disease, in whom we increase our diligence in searching for disease, and ultimately how we use the drug and other therapeutic armamentarium available to the treating physician.

These points will be incorporated with surgical approaches into discussions of anatomy, diagnostic workup, infection, surgical and medical aspects of the obstetric patient, pelvic floor dysfunction, and neoplasms.

ANATOMY

Clinical gynecologic anatomy centers on the pelvis (L. basin). Apty named, the bowl-shaped pelvis houses the confluence and intersection of multiple organ systems. Understanding
those structural and functional relationships is essential for the surgeon and allows an appreciation for the interplay of sexual function and reproduction as well as a context for understanding gynecologic pathology.

**Structure and Support of the Pelvis and Genitalia**

The bony pelvis is comprised by the **sacrum** posteriorly and the **ischium**, **ilium**, and **pubic** bones anteromedially. It supports the upper body and transmits the stresses of weight bearing to the lower limbs in addition to providing anchors for the supporting tissues of the pelvic floor.1 The opening of the pelvis is spanned by the muscles of the pelvic diaphragm (Fig. 41-1). The muscles of the pelvic sidewall include the **iliacus**, the **psoas**, and the **obturator internus** muscle (Fig. 41-2). These muscles contract tonically and include, from anterior to posterior, bilaterally, the **pubococcygeus**, **puborectalis**, **iliococcygeus**, and **coccygeus** muscles. The first two of these muscles contribute fibers to the fibromuscular perineal body. The **urogenital hiatus** is bordered laterally by the **pubococcygeus** muscles and anteriorly by the **symphysis pubis**. It is through this muscular defect that the urethra and vagina pass, and it is the focal point for the study of disorders of pelvic support such as cystocele, rectocele, and uterine prolapse.

1. Gynecologic causes of acute abdomen include PID and tubo-ovarian abscess, ovarian torsion, ruptured ectopic pregnancy, septic abortion. Pregnancy must be ruled out early in assessment of reproductive age patients presenting with abdominal or pelvic pain.
2. The general gynecology exam must incorporate the whole physical examination in order to adequately diagnosis and treat gynecologic disorders.
3. Benign gynecologic pathologies that are encountered at the time of surgery include endometriosis, endometriomas, fibroids, and ovarian cysts.
4. It is critical that abnormal lesions of vulva, vagina, and cervix are biopsied for diagnosis before any treatment is planned; postmenopausal bleeding should always be investigated to rule out malignancy.
5. Pelvic floor dysfunction (pelvic organ prolapse, urinary and fecal incontinence) is common; 11% of women will undergo a reconstructive surgical procedure at some point in their lives.
6. Pregnancy confers important changes to both the cardiovascular system and the coagulation cascade. Trauma in pregnancy must be managed with these changes in mind.
7. Early-stage cervical cancer is managed surgically, whereas chemoradiation is preferred for stages Ib2 and above.
8. Risk-reducing salpingo-oopherectomy is recommended in women with BRCA1 or BRCA2 mutations.
9. Optimal debulking for epithelial ovarian cancer is a critical element in patient response and survival. The preferred postoperative therapy for optimally debulked advanced-stage ovarian epithelial ovarian cancer is intraperitoneal chemotherapy.
10. Long-term sequelae of intestinal and urologic injury can be avoided by intraoperative identification.

![Figure 41-1](image-url) Deeper muscles of the pelvic floor.
Vulva

The labia majora form the cutaneous boundaries of the lateral vulva and represent the female homologue of the male scrotum (Fig. 41-4). The labia majora are fatty folds covered by hair-bearing skin in the adult. They fuse anteriorly over the anterior prominence of the symphysis pubis, the mons pubis. The deeper portions of the adipose layers are called Colles fascia and insert onto the inferior margin of the perineal membrane, limiting spread of superficial hematomas inferiorly. Adjacent and medial to the labia majora are the labia minora, smaller folds of connective tissue covered laterally by non–hair-bearing skin and medially by vaginal mucosa. The anterior fusion of the labia minora forms the prepuce and frenulum of the clitoris; posteriorly, the labia minora fuse to create the fossa navicularis and posterior fourchette. The term vestibule refers to the area medial to the labia minora bounded by the fossa navicularis and the clitoris. Both the urethra and the vagina open into the vestibule. Skene’s glands lie lateral and inferior to the urethral meatus. Cysts, abscesses, and neoplasms may arise in these glands.

Erectile tissues and associated muscles are in the space between the perineal membrane and the vulvar subcutaneous tissues (see Fig. 41-1). The clitoris is formed by two crura and is suspended from the pubis. Overlying the crura are ischiocavernosus muscles, which run along the inferior surfaces of the ischiopubic rami. Extending medially from the inferior end of the ischiocavernosus muscles are the superficial transverse perinei muscles. These terminate in the midline in the perineal body, caudal and deep to the posterior fourchette. Vestibular bulbs lie just deep to the vestibule and are covered laterally by bulbocavernosus muscles. These originate from the perineal body and insert into the body of the clitoris. At the inferior end of the vestibular bulbs are Bartholin’s glands, which connect to the vestibular skin by ducts.

Vagina

The vagina is an elastic fibromuscular tube opening from the vestibule running superiorly and posteriorly, passing through the perineal membrane. The lower third is invested by the superficial and deep perineal muscles; it incorporates the urethra in its anterior wall and has a rich blood supply from the vaginal branches of the external and internal pudendal arteries. The upper two-thirds of the vagina are not invested by muscles. This portion lies in opposition to the bladder base anteriorly and the rectum and posterior pelvic cul-de-sac superiorly. The cervix opens into the posterior vaginal wall bulging into the vaginal lumen.

Uterus

The typically pear-shaped uterus consists of a fundus, cornua, body, and cervix. It lies between the bladder anteriorly and the rectosigmoid posteriorly. The endometrium lines the inside cavity and has a superficial functional layer that is shed with menstruation and a basal layer from which the new functional layer is formed. Sustained estrogenic stimulation without associated progestin maturation can lead to hyperplastic changes or carcinoma. Adenomyosis is a condition in which benign endometrial glands infiltrate into the muscle or myometrium of the uterus. The myometrium is composed of smooth muscle and the contraction of myometrium is a factor in menstrual pain and is essential in childbirth. The myometrium can develop benign smooth muscle neoplasms known as leiomyoma or fibroids.

Cervix

The cervix connects the uterus and vagina and projects into the upper vagina. The vagina forms an arched ring around the cervix described as the vaginal fornices—lateral, anterior, and posterior. The cervix is about 2.5–cm long with a fusiform endocervical canal lined by columnar epithelium lying between an internal and external os, or opening. The vaginal surface of the cervix is covered with stratified squamous epithelium, similar to that lining the vagina. The squamo-columnar junction, also referred to as the transformation zone, migrates at different stages of life and is influenced by estrogenic stimulation. The transformation zone develops as the columnar epithelium is replaced by squamous metaplasia. This transformation zone is
vulnerable to human papilloma virus (HPV) infection and resultant premalignant changes. These changes can be detected by microscopic assessment of cervical cytological (or Pap) smear. If the duct of a cervical gland becomes occluded, the gland tends to form a retention cyst or Nabothian follicle.

**Fallopian Tubes**

The bilateral fallopian tubes arise from the upper lateral cornua of the uterus and course posterolaterally within the upper border of the broad ligament. The tubes can be divided into four parts. The interstitial part forms a passage through the myometrium. The isthmus is the narrow portion extending about 3 cm from the myometrium. The ampulla is thin-walled and tortuous with its lateral end free of the broad ligament. The infundibulum is the distal end fringed by a ring of delicate fronds or fimbriae. The fallopian tubes receive the ovum after ovulation. Peristalsis carries the ovum to the ampulla where fertilization occurs. The zygote transits the tube over the course of 3 to 4 days to the uterus. Abnormal implantation in the fallopian tube is the most common site of ectopic pregnancies. The tubes may also be infected by ascending organisms, resulting in tubo-ovarian abscesses. Scarring of the fallopian tubes can lead to hydrosalpinx. Recent evidence suggests most high-grade serous ovarian cancer originates in the fallopian tubes.

**Ovaries**

The ovaries are attached to the uterine cornu by the proper ovarian ligaments, or the utero-ovarian ligaments. The ovaries are suspended from the lateral pelvis by their vascular pedicles, the infundibulopelvic ligaments (IP) or ovarian arteries. These are also called the suspensory ligaments of the ovaries, and correspond to the genital vessels in the male. The IP’s are paired branches from the abdominal aorta arising just below the renal arteries. They merge with the peritoneum over the psoas major muscle and pass over the pelvic brim and the external iliac vessels. The ovarian veins ascend at first with the ovarian arteries, then track more laterally. The right ovarian vein ascends directly into the inferior vena cava while the left vein drains into the left renal vein. Lymphatic drainage follows the arteries to the para-aortic lymph nodes. The ovaries are covered by a single layer of cells that is continuous with the mesothelium of the peritoneum. Beneath this is a fibrous stroma within which are embedded germ cells. At ovulation, an ovarian follicle ruptures through the ovarian epithelium.

**Fibrovascular Ligaments and Avascular Tissue Planes**

Figure 41-5 is a view of the internal genitalia and deep pelvis as one would approach the pelvis from a midline abdominal incision. The central uterus and uterine cervix are supported by the pelvic floor muscles (Fig. 41-5). They are suspended by
the lateral fibrous cardinal, or Mackenrodt’s ligament, and the uterosacral ligaments, which insert into the paracervical fascia medially and into the muscular sidewalls of the pelvis laterally. Posteriorly, the uterosacral ligaments provide support for the vagina and cervix as they course from the sacrum lateral to the rectum and insert into the paracervical fascia. Emanating from the uterine cornu and traveling through the inguinal canal are the round ligaments, eventually attaching to the subcutaneous tissue of the mons pubis. The peritoneum enfolding the adnexa (tube, round ligament, and ovary) is referred to as the broad ligament, which separates the pelvic cavity into an anterior and posterior component.

The peritoneal reflections in the pelvis anterior and posterior to the uterus are referred to as the anterior and posterior cul-de-sacs. The latter is also called the pouch or cul-de-sac of Douglas. On transverse section, seven avascular, and therefore important, surgical planes can be identified (Fig. 41-6). These include the right and left lateral paravesical and right and left pararectal spaces, and from anterior to posterior, the retropubic or prevesical space of Retzius and the vesicovaginal, rectovaginal, and retrorectal or presacral spaces.

These avascular tissue planes are often preserved and provide safe surgical access when the intraperitoneal pelvic anatomy is distorted by tumor, endometriosis, adhesions, or infection. Utilizing the avascular retroperitoneal planes, the ureter can be traced into the pelvis as it crosses the distal common iliac arteries laterally into the pararectal space and then courses inferior to the ovarian arteries and veins until crossing under the uterine arteries into the paravesical space just lateral to the cervix. After traveling to the cervix, the ureters course downward and medially over the anterior surface of the vagina before entering the base of the bladder in the vesicovaginal space.

Vasculature and Nerves of the Pelvis

The rich blood supply to the pelvis arises largely from the internal iliac arteries except for the middle sacral artery originating at the aortic bifurcation and the ovarian arteries originating from the abdominal aorta. There is also collateral flow and anastomoses to the pelvic vessels from the inferior mesenteric artery. The internal iliac, or hypogastric, arteries divide into anterior and posterior branches. The latter supply lumbar and gluteal branches. From the anterior division of the hypogastric arteries arise the obturator, uterine, pudendal, middle rectal, inferior gluteal, along with superior and middle vesical arteries (see Fig. 41-2).

The major motor nerves found in the pelvis are the sciatic, obturator, and femoral nerves (Fig. 41-3). Also important to the pelvic surgeon are the ilioinguinal, iliohypogastric and genitofemoral nerves, which arise as upper abdominal nerves, but are encountered on the most caudal portion of the anterior abdominal wall and the ventral portion of the external genitalia. Sympathetic fibers course along the major arteries and parasympathetics form the superior and inferior pelvic plexus. The pudendal nerve arises from S2–S4 and travels laterally, exiting the greater sciatic foramen, hooking around the ischial spine and sacrospinous ligament, and returning via the greater sciatic foramen. It travels through Alcock’s canal and becomes the sensory and motor nerve of the perineum (see Figs. 41-1 and 41-3). The motor neurons serve the tonically contracting urethral and anal sphincter, and direct branches from the S2–S4 nerves serve the levator ani muscles. During childbirth and other excessive straining, this tethered nerve (along with the levator ani muscles) is subject to stretch injury and is at least partially responsible for many female pelvic floor disorders.

EVALUATION AND DIAGNOSIS

Elements of a Gynecologic History

A complete history is a seminal part of any assessment (Table 41-1). Many gynecologic diseases can present with broad constitutional symptoms, occur secondary to other conditions, or be related to medications. A full history should include particular attention to family history, organ system history, including breast, gastrointestinal, and urinary tract symptoms, and a careful medication, anesthesia, and surgical history. The key elements of a focused gynecologic history include the following:

- Date of last menstrual period
- History of contraceptive and postmenopausal hormone use
- Obstetrical history
- Age at menarche and menopause (method of menopause, [e.g., drug, surgical])
- Menstrual bleeding pattern
- History of pelvic assessments, including cervical smear and HPV DNA results
- History of pelvic infections, including HPV and HIV status
- Sexual history
- Prior gynecologic surgery(s)

The Gynecologic Examination

For many young women, their gynecologist is their primary care physician. When that is the case, it is necessary that a full medical and surgical history be taken and that, in addition to the pelvic examination, the minimum additional examination should include assessment of the thyroid, breasts, and cardiopulmonary system. Screening, reproductive counseling, and age-appropriate health services should be available to women of all ages with or without a routine pelvic examination, but the decision to proceed with regular, annual pelvic examinations in otherwise healthy women is controversial. The U.S. Preventive Services Task Force recently evaluated the current evidence regarding the balance of benefits and harms of performing screening pelvic examinations in asymptomatic, nonpregnant adult women and concluded that the evidence is insufficient.

Figure 41-6. The avascular spaces of the female pelvis.
The pelvic examination starts with a full abdominal examination. Inguinal node evaluation is performed before placing the patient’s legs in the dorsal lithotomy position (in stirrups). A flexible, focused light source is essential, and vaginal instruments including speculums of variable sizes and shapes (Graves and Pederson), including pediatric sizes, are required to assure that the patient’s anatomy can be fully and comfortably viewed. A warmed lubricated speculum is inserted into the vagina and gently opened to identify the cervix if present or the vaginal apex if not. To avoid confusing the location of pelvic pain with immediate speculum exam, or if there is a concern that a malignancy is present, careful digital assessment of a vaginal mass and location may be addressed prior to speculum placement in order to avoid abrading a vascular lesion and inducing hemorrhage. The speculum would then be inserted just short of the length to the mass in order to view that area directly before advancing. An uncomplicated speculum exam includes examination of the vaginal sidewalls, assessment of secretions, including culture if necessary, and collection of the cervical cytologic specimen and HPV test if indicated (see “Common Screening”).

A bimanual examination is performed by placing two fingers in the vaginal canal; one finger may be used if patient has significant vaginal atrophy or has had prior radiation with stenosis (Fig. 41-7). Carefully and sequentially assess the size and shape of the uterus by moving it against the abdominal hand, and the adnexa by carefully sweeping the abdominal hand down the side of the uterus. The rectovaginal examination, consisting of one finger in the vagina and one in the rectal vault, is used to further examine and characterize the location, shape, fixation, size, and complexity of the uterus, adnexa, cervix, and anterior and posterior cul-de-sacs. The rectovaginal exam also allows examination of the uterosacral ligaments from the back of the uterus sweeping laterally to the rectal finger and the sacrum, as well as assessment of the rectum and anal canal for masses.

It is critical that presurgical assessments include a full general examination. This is particularly important with potential oncologic diagnoses or infectious issues in order to assure that the proposed surgery is both safe and appropriate. Issues such as sites of metastatic cancer or infection, associated bleeding and/
or clotting issues and history, and drug exposure, allergies, and current medications must be addressed.

**Commonly Used Testing**

**β-Human Chorionic Gonadotropin Testing.** Qualitative urinary pregnancy tests for human chorionic gonadotropin (β-hCG) are standard prior to any surgery in a woman of reproductive age and potential, regardless of contraception history. In addition, serum quantitative β-hCG testing is appropriate for evaluation of suspected ectopic pregnancy, gestational trophoblastic disease, or ovarian mass in a young woman. In the case of ectopic pregnancy, serial levels are required when a pregnancy cannot be identified in the uterine cavity by imaging. As a general rule, 85% of viable, very early intrauterine pregnancies will have at least a 66% rise in the β-hCG level over 48 hours.

**Microscopy of Vaginal Discharge.** During a speculum exam, a cotton-tipped applicator is used to collect the vaginal discharge; it is smeared on a slide with several drops of 0.9% normal saline to create a saline wet mount. A cover slide is placed and the slide is evaluated microscopically for the presence of mobile trichomonads (*Trichomonas vaginalis*) or clue cells (epithelial cells studded with bacteria, seen in bacterial vaginosis; Table 41-2). A potassium hydroxide (KOH) wet mount is the slide application of the collected vaginal discharge with 10% KOH; this destroys cellular elements. The test is positive for vaginal candidiasis when pseudohyphae are seen (see Table 41-2).

**Chlamydia/Gonorrhea Testing.** Nucleic acid amplification testing (NAAT) has emerged as the diagnostic test of choice for *N gonorrhoea* and *C trachomatis*. A vaginal swab, endocervical swab, and/or urine sample, can be used for this test.

**Cervical Cancer Screening and Prevention.** HPV infection is required for the development of epithelial cervical carcinomas (squamous and adenocarcinomas), and HPV DNA can be identified in virtually all primary cervical malignancies. HPV is a ubiquitous double-stranded DNA virus commonly acquired in the female lower genital tract through sexual contact. After entry into the cell, the HPV protein E6 degrades the tumor suppressor p53, resulting in deregulation of cell cycle arrest. E7 inactivates the tumor suppressor RB and releases E2F transcription factors, causing cellular hyperproliferation. More than 100 HPV types have been identified, and up to 40 of these subtypes infect the anogenital region. At least 12 are considered high-risk or oncogenic, and HPV genotypes 16 and 18 cause approximately 70% of cervical cancers worldwide.4

Recent cervical cytology guidelines have increased the intervals between screenings for most women given the known natural history of HPV-related cervical dysplasia progression to cancer and the high negative predictive value of a negative HPV test.6 The current recommendations call for cervical smear screening every 3 to 5 years in women ages 21 to 65 years. If an

---

**Table 41-2**

<table>
<thead>
<tr>
<th>Pathogen</th>
<th>BACTERIAL VAGINOSIS</th>
<th>VULVOVAGINAL CANDIDIASIS</th>
<th>TRICHOMONIASIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of vaginitis</td>
<td>40</td>
<td>30</td>
<td>20</td>
</tr>
<tr>
<td>pH</td>
<td>&gt;4.5</td>
<td>&lt;4.5</td>
<td>&gt;4.5</td>
</tr>
<tr>
<td>Signs and symptoms</td>
<td>Malodorous, adherent discharge</td>
<td>White discharge, vulvar erythema, pruritus, dyspareunia</td>
<td>Malodorous purulent discharge, vulvovaginal erythema, dyspareunia</td>
</tr>
<tr>
<td>Wet mount</td>
<td>Clue cells</td>
<td>Pseudohyphae or budding yeasts in 40% of cases</td>
<td>Motile trichomonads</td>
</tr>
<tr>
<td>KOH mount</td>
<td></td>
<td>Pseudohyphae or budding yeasts in 70% of cases</td>
<td></td>
</tr>
<tr>
<td>Amine test</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Treatment</td>
<td>Metronidazole 500 mg twice a day for 7 d or 2 g single dose, metronidazole or clindamycin vaginal cream</td>
<td>Oral fluconazole 150 mg single dose, vaginal antifungal preparations</td>
<td>Metronidazole 2 g single dose and treatment of partner</td>
</tr>
</tbody>
</table>

+ = positive; – = negative; KOH = potassium hydroxide.
HPV test performed at the same time also is negative, testing should be repeated every 5 years for women ages 30 to 65 years. Screening is not recommended for women age older than 65 or without a cervix (prior hysterectomy) unless they have a history of high-grade precancerous lesions. Women with a history of cervical dysplasia, HPV infection, or cervical cancer need more frequent screening based on their diagnosis. Primary high-risk HPV (hrHPV) screening is also an acceptable alternative to cytologic screening for women ages 30-65 because of an increased detection of high-grade squamous intraepithelial lesion (HSIL) and increased negative predictive value.6

**HPV Vaccine.** Three HPV vaccines have been approved by the U.S. Food and Drug Administration (FDA).7 In 2006, a quadrivalent (4vHPV) vaccine was approved that targets HPV 16 and 18, which cause 70% of cervical cancers, and HPV genotypes 6 and 11, which cause 90% of genital warts. In December 2014, a nine-valent vaccine (9vHPV) was introduced to replace the 4vHPV vaccine, which includes protection against the HPV strains covered by the first generation of 4vHPV as well as five other HPV strains responsible for 20% of cervical cancers (HPV-31, HPV-33, HPV-45, HPV-52, and HPV-58).7 The 9vHPV may be used to continue or complete a series started with a different HPV vaccine product. Vaccination with 9vHPV after completion of 4vHPV at least 12 months earlier is safe and may provide protection against additional HPV strains. A bivalent vaccine that targets HPV genotypes 16 and 18 with a different adjuvant that may have led to higher immunogenicity was approved in 2009 but is no longer marketed in the United States.

Vaccination generates high concentrations of neutralizing antibodies to HPV L1 protein, the antigen in all HPV vaccines. The vaccines are highly immunogenic, activating both humoral and cellular immune responses. Multiple randomized clinical trials have demonstrated nearly 100% efficacy in the prevention of the HPV subtype-specific precancerous cervical cell changes.7,8 These major clinical trials have used prevention of HSIL as the efficacy endpoints. Vaccination does not protect women who are already infected with HPV-16 or -18 at the time of vaccination.

Current recommendations include HPV vaccination for boys and girls at age 11 or 12 years. (Vaccination can be started at age 9.) The Advisory Committee on Immunization Practices (ACIP) also recommends vaccination for females aged 13 through 26 years and males aged 13 through 21 years not adequately vaccinated previously. Catch-up vaccination is also recommended through age 26 years for gay, bisexual, and other men who have sex with men, transgender people, and for immunocompromised persons (including those with HIV infection) not adequately vaccinated previously.8 Two doses are given 6 to 12 months apart for patients with an intact immune system, age less than 15 years; three doses are recommended for those ages 15 to 26 years and immunocompromised persons.10 Cervical cancer screening continues to play an important role in detection and treatment of premalignant cervical lesions and prevention of cervical cancer in these high-risk patients and is currently recommended following HPV vaccination.

**Serum Cancer Antigen 125.** Cancer antigen (CA) 125 is a large membrane glycoprotein belonging to the wide mucin family commonly used as a tumor marker in patients known to have ovarian cancer. An elevated CA-125 in the patient without known ovarian cancer should be interpreted in conjunction with patient information and symptoms as well as imaging. In the setting of an adnexal mass, the serum CA-125 test may help with triage of a patient to the appropriate surgical management. The test should be used with caution as it is a nonspecific test and may be elevated with multiple benign conditions including endometriosis, fibroids, infection, and pregnancy and may even vary with the menstrual cycle. For these reasons, the CA-125 test is less useful in the premenopausal woman for triaging an adnexal mass. In the postmenopausal woman, a CA-125 greater than 35 in the setting of a complex adnexal mass merits referral of the patient to a gynecologic oncologist.10

**Common Office Procedures for Diagnosis**

**Vulvar/Vaginal Biopsy.** Any abnormal vulvar or vaginal lesion including skin color changes, raised lesions, or ulcerations should be biopsied. Local infiltration with local anesthetic is followed by a 3- to 5-mm punch biopsy appropriate to the lesion. The specimen is elevated with Adson forceps and cut from its base with scissors. The vaginal biopsy can sometimes be difficult to perform because of the angle of the lesion. After injection with local anesthetic, traction of the area with Allis forceps and direct resection of the lesion with scissors or cervical biopsy instrument (Schubert, Kevorkian, etc) can achieve an adequate biopsy.

**Colposcopy and Cervical Biopsy.** In cases of an abnormal Pap smear cytology or positive HPV testing, a colposcopy is performed for a histologic evaluation. A colposcope is used to achieve 2x to 15x magnification of the cervix. Once the cervix is visualized, cervical mucus, if present, is removed, and then 3% acetic acid is applied to the cervix for one minute. This application dehydrates cells and causes dysplastic cells with dense nuclei to appear white. The lining of the cervix consists of squamous epithelium on the ectocervix, whereas columnar epithelium lines the endocervical canal. The ectocervix therefore appears smooth and pale pink in color while the endocervix forms epithelial fronds or “grape-like” structures visible through the colposcope. The junction between columnar and squamous cell types is called the squamocolumnar junction (SCJ), which in younger women is usually visible on the ectocervix. When columnar epithelium extends onto the ectocervix, it appears as a red zone surrounding the os and is called ectropion or ectopy. The transformation zone (TZ) is the area between mature squamous epithelium distally and columnar epithelium proximally, and it is the site of active squamous metaplasia. For colposcopy to be deemed adequate, the entire SCJ must be visualized during an adequate colposcopy. Areas with acetowhite, punctuation, mosaicism, or atypical blood vessels seen during colposcopy may represent dysplasia or cancer and should be biopsied. A green filter enhances visualization of blood vessels by making them appear darker in contrast to the surrounding epithelium.

An alternative to dilute acetic acid is Lugol’s solution—a concentrated solution of iodine that reacts with the glycogen in normal squamous epithelium to make it appear dark brown. High-grade CIN lesions have low amounts of glycogen because the epithelium is poorly differentiated, and hence they do not turn brown with Lugol’s solution. This is termed Lugol’s nonstaining or Lugol’s negative. Historically, this used to be referred to as the Schiller’s test. Lugol’s can be useful for determining whether a colposcopically equivocal area warrants biopsy: Lugol’s staining areas are most likely normal epithelium, whereas Lugol’s nonstaining areas may be CIN, metaplasia, or inflammation.
Endometrial Biopsy. Endometrial sampling should be performed before planned hysterectomy if there is a history of bleeding between periods, heavy and/or frequent menstrual periods, or postmenopausal bleeding. A patient with the potential for pregnancy should have a pregnancy test before the procedure. A pipelle endometrial biopsy can be performed in the office and is a cost-effective and safe procedure that is generally well tolerated by patients. The pipelle is a flexible polypropylene suction cannula with an outer diameter of 3.1 mm. The pipelle is inserted through the endocervix after cervical cleaning, and the depth of the uterine cavity is noted. If difficulty in entering the endometrium with the pipelle is encountered, a tenaculum may be used to straighten the cervix and/or an OS-finder may be useful in overcoming resistance within the endocervix. The endometrial specimen is obtained by pulling on the plunger within the pipelle, creating a small amount of suction. The pipelle is rotated and pulled back from the fundus to the lower uterine segment within the cavity to access all sides. Additional passes may be needed in order to acquire an adequate amount of tissue. If office biopsy is not possible due to patient discomfort or cervical stenosis, a dilatation and curettage in the operating room may be indicated depending on the clinical circumstances.

Evaluation for Fistula. When a patient presents with copious vaginal discharge, the provider should be concerned about a fistula with the urinary or gastrointestinal tract. A simple office procedure can be performed when there is a concern for a vesicovaginal fistula. A vaginal tampon is placed followed by instillation of sterile blue dye through a transurethral catheter into the bladder; a positive test is blue staining of the tampon. If the test is negative, one can evaluate for a ureterovaginal fistula. The patient is given phenazopyridine, which changes the color of urine to orange. If a tampon placed in the vagina stains orange, the test is positive. Alternatively, the patient can be given an intravenous injection of indigo carmine.

Rectal fistula must be considered when a patient reports stool evacuation per vagina. It can be identified in a similar fashion using a large Foley catheter placed in the distal rectum through which dye may be injected, or with the use of an oral charcoal slurry and timed examination. Common areas for fistulae are at the vaginal apex, at the site of a surgical incision, or around the site of a prior episiotomy or perineal repair after a vaginal delivery.

BENIGN GYNECOLOGIC CONDITIONS

Vulvar Lesions

Patients presenting with vulvar symptoms should be carefully interviewed and examined, and a vulvar biopsy should be obtained whenever the diagnosis is in question, the patient does not respond to treatment, or premalignant and malignant disease is suspected. Vulvar conditions such as contact dermatitis, atrophic vulvovaginitis, lichen sclerosis, lichen planus, lichen chronicus simplex, Paget’s disease, Bowen’s disease, and invasive vulvar cancer are common particularly in postmenopausal women. Systemic diseases like psoriasis, eczema, Crohn’s disease, Behçet’s disease, vitiligo, and seborrheic dermatitis may also involve the vulvar skin.

Leukoplaikias. There are three types of leukoplaikia, a flat white abnormality. Lichen sclerosis is the most common cause of leukoplaikia. There are two peaks of onset: prepubertal girls and perimenopausal or postmenopausal women. Classically, it results in a figure-of-eight pattern of white epithelium around the anus and vulva resulting in variable scarring and itching, and less commonly pain. Diagnosis is confirmed with biopsy, and treatment consists of topical steroids. An established association between lichen sclerosis and vulvar squamous cell carcinoma estimates risk of malignant transformation up to 5%. Lichen planus is a cause of leukoplakia with an onset in the fifth and sixth decade of life. Lichen planus, in contrast to lichen sclerosis which is limited to the vulva and perianal skin, can involve the vagina and oral mucosa, and erosions occur in the majority of patients leading to a variable degree of scarring. Patients usually have a history and dysuria and dyspareunia, and complain of a burning vulvar pain. Histology is not specific, and biopsy is recommended. Treatment is with topical steroids. Systemic steroids are indicated for severe and/or unresponsive cases.

Lichen simplex chronicus is the third cause of leukoplaikia, but is distinguished from the other lichen diseases by epidermal thickening, absence of scarring, and a severe intolerable itch. Intense scratching is common, and contributes to the severity of the symptoms and predisposes the cracked skin to infections. Treatment consists of cessation of the scratching which sometimes requires sedation, elimination of any allergen or irritant, suppression of inflammation with potent steroid ointments, and treatment of any coexisting infections.

Bartholin’s Cyst or Abscess. Bartholin’s glands, great vestibular glands, are located at the vaginal orifice at the four and eight o’clock positions; they are rarely palpable in normal patients. They are lined with cuboidal epithelium and secrete mucus material to keep the vulva moist. Their ducts are lined with transitional epithelium, and their obstruction secondary to inflammation may lead to the development of a Bartholin’s cyst or abscess. Bartholin’s cysts or abscesses are usually asymptomatic and are easily diagnosed on examination. Infections are usually polymicrobial. Treatment consists of incision and drainage and placement of a Word catheter, a small catheter with a balloon tip, for 2 to 3 weeks to allow for formation and epithelialization of a new duct. Recurrent cysts or abscesses may require marsupialization, but on occasion these necessitate excision of the whole gland. Marsupialization is performed by incising the cyst or abscess wall and securing its lining to the skin edges with interrupted sutures. Cysts or abscesses that fail to resolve after drainage and those occurring in patients over 40 years old should be biopsied to exclude malignancy.

Molluscum Contagiosum. Molluscum contagiosum presents with dome-shaped papules and are caused by the pox virus. The papules are usually 2 to 5 mm in diameter and classically have a central umbilication. They are spread by direct skin contact, and present on the vulva, as well as abdomen, trunk, arms, and thighs. Lesions typically clear in several months, but they can be treated with cryotherapy, curettage, or cantharidin, a topical blistering agent.

Genital Ulcers. The frequency of the infectious etiologies of genital ulcers varies by geographic location. The most common causes of sexually transmitted genital ulcers in young adults in the United States are, in descending order of prevalence, herpes simplex virus (HSV), syphilis, and chancroid. Other infectious causes of genital ulcers include lymphogranuloma venerum and granuloma inguinale. Noninfectious etiologies include Behçet’s disease, neoplasms, and trauma. Table 41-3 outlines a rational approach to their evaluation and diagnosis.
### Vulvar Condyloma

Condylomata acuminata (anogenital warts) are viral infections caused by HPV. Genital infection with HPV is the most common sexually transmitted infection in the United States today. HPV 6 and 11 are the most common low-risk types and are implicated in 90% of cases of genital warts. Women with immunosuppression due to HIV or solid organ transplant are at higher risk of vulvar condyloma than immunocompetent women. Genital warts are skin-colored or pink and range from smooth flattened papules to verrucous papilliform lesions. Lesions may be single or multiple and extensive. Diagnosis should be confirmed with biopsy as verrucaous vulvar cancers can be mistaken for condylomata. If small, self-administered topical imiquimod 5% cream or trichloroacetic acid for in-office applications may be tried. Extensive lesions may require surgical modalities that include cryotherapy, laser ablation, cauterization, and surgical excision.

### Paget’s Disease of the Vulva

Paget’s disease of the vulva is an intraepithelial disease of unknown etiology that affects

### Table 41-3

<table>
<thead>
<tr>
<th></th>
<th>HERPES</th>
<th>SYPHILIS</th>
<th>CHANCROID</th>
<th>LYMPOHGRANULOMA VENEREUM</th>
<th>GRANULOMA INGUINALE (DONOVANOSIS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathogen</td>
<td>HSV type 2 and less commonly HSV type 1</td>
<td><em>Treponema pallidum</em></td>
<td><em>Haemophilus ducreyi</em></td>
<td><em>Chlamydia trachomatis</em> L1-L3</td>
<td><em>Calymmatobacterium granulomatis</em></td>
</tr>
<tr>
<td>Incubation period</td>
<td>2–7 days</td>
<td>2–4 weeks (1–12 weeks)</td>
<td>1–14 days</td>
<td>3 days–6 weeks</td>
<td>1–4 weeks (up to 6 months)</td>
</tr>
<tr>
<td>Primary lesion</td>
<td>Vesicle</td>
<td>Papule</td>
<td>Papule or pustule</td>
<td>Papule, pustule, or vesicle</td>
<td>Papule</td>
</tr>
<tr>
<td>Number of lesions</td>
<td>Multiple, may coalesce</td>
<td>Usually one</td>
<td>Usually multiple, may coalesce</td>
<td>Usually one</td>
<td>Variable</td>
</tr>
<tr>
<td>Diameter (mm)</td>
<td>1–2</td>
<td>5–15</td>
<td>2–20</td>
<td>2–10</td>
<td>Variable</td>
</tr>
<tr>
<td>Edges</td>
<td>Erythematous</td>
<td>Sharply demarcated, elevated, round, or oval</td>
<td>Undermined, ragged, irregular</td>
<td>Elevated, round, or oval</td>
<td>Elevated, irregular</td>
</tr>
<tr>
<td>Depth</td>
<td>Superficial</td>
<td>Superficial or deep</td>
<td>Excavated</td>
<td>Superficial or deep</td>
<td>Elevated</td>
</tr>
<tr>
<td>Base</td>
<td>Serous, erythematous</td>
<td>Smooth, nonpurulent</td>
<td>Purulent</td>
<td>Variable</td>
<td>Red and rough (“beefy”)</td>
</tr>
<tr>
<td>Induration</td>
<td>None</td>
<td>Common</td>
<td>Unusual</td>
<td>Usually very tender</td>
<td>Variable</td>
</tr>
<tr>
<td>Pain</td>
<td>None</td>
<td>Firm</td>
<td>Soft</td>
<td>Occasionally firm</td>
<td>Firm</td>
</tr>
<tr>
<td>Lymph-adenopathy</td>
<td>Firm, tender, often bilateral</td>
<td>Firm, nontender, bilateral</td>
<td>Tender, may suppurate, usually unilateral</td>
<td>Tender, may suppurate, loculated, usually unilateral</td>
<td>Pseudo-adenopathy</td>
</tr>
<tr>
<td>Treatment</td>
<td>acyclovir (ACV) 400 mg POI three times a day for 7–10 days for primary infection and 400 mg PO three times a day for 5 days for episodic management</td>
<td>Primary, secondary, and early latent (&lt;1 year): benzathine PCN-G 2.4 million U IM × 1 Late latent (&gt;1 year) and latent of unknown duration: benzathine PCN-G 2.4 million units IM every week × 3</td>
<td>azithromycin 1 g po or ceftriaxone 250 mg IM × 1 OR Ciprofloxacin 500 mg po twice a day for 3 days Erythromycin base 500 mg po three times a day for 7 days</td>
<td>Doxycycline 100 mg po twice a day × 21 days OR Erythromycin base 500 mg po four times a day for 21 days</td>
<td>Doxycycline 100 mg po twice a day for 3 weeks until all lesions have healed</td>
</tr>
<tr>
<td>Suppression</td>
<td>acyclovir 400 mg po twice a day for those with frequent outbreaks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

mostly postmenopausal women in their sixth decade of life. It causes chronic vulvar itching and is sometimes associated with an underlying invasive vulvar adenocarcinoma or invasive cancers of the breast, cervix, or gastrointestinal tract. Grossly, the lesion is variable but usually confluent, raised, erythematous to violet, and waxy in appearance. Biopsy is required for diagnosis; the disease is intraepithelial and characterized by Paget’s cells with large pale cytoplasm. Treatment is assessment for other potential concurrent adenocarcinomas and then surgical removal by wide local resection of the involved area with a 2-cm margin. Free margins are difficult to obtain because the disease usually extends beyond the clinically visible area. Intraoperative frozen section of the margins can be done; however, Paget’s vulvar lesions have a high likelihood of recurrence even after securing negative resection margins.

**Vulvar Intraepithelial Neoplasia.** Two pathologically distinct premalignant lesions of the vulva are currently recognized. Vulvar intraepithelial neoplasia (VIN) of usual type (uVIN) is caused by the HPV virus, tends to occur in younger women, and presents as multifocal disease. VIN of differentiated type (dVIN) develops independently of HPV and is typically unifocal and seen in postmenopausal women. VIN is similar to its cervical intraepithelial neoplasia (CIN) counterpart in the cervix. In 2012, the pathologic terminology of HPV-related disease in the anogenital region was harmonized into a two-tier system where LSIL is equivalent to uVIN 1 and HSIL encompasses uVIN 2 and uVIN 3. Additional risk factors for the development of VIN include HIV infection, immunosuppression, smoking, vulvar dermatoses such as lichen sclerosis, CIN, and a history of cervical cancer. Vulvar pruritus is the most common complaint in women with symptoms. Lesions may be vague or raised, and they may be velvety with sharply demarcated borders. Diagnosis is made with a vulvar skin biopsy and multiple biopsies are sometimes necessary. Evaluation of the perianal and anal area is important as the disease may involve these areas. Once invasive disease is ruled out, treatment usually involves wide surgical excision; however, the treatment approaches may also include 5% imiquimod cream, CO₂ laser ablation, or cavitation ultrasonic surgical aspiration (CUSA), and depends on the number of lesions and their severity. When laser ablation is used, a 1-mm depth in hair-free areas is usually sufficient, while hairy lesions require ablation to a 3-mm depth because the hair follicles’ roots can reach a depth of 2.5 mm. Unfortunately, VIN tends to recur in up to 30% of cases, and high-grade lesions will progress to invasive disease in approximately 10% of patients if left untreated.

**Vaginal Lesions**

**Vaginitis (see Table 41-2).** Vaginal symptoms are extremely common, accounting for over 10 million office visits per year in the United States. The causes of vaginal complaints are commonly infectious in origin, but they include a number of noninfectious causes, such as chemicals or irritants, hormone deficiency, foreign bodies, systemic diseases, and malignancy. Symptoms include abnormal vaginal discharge, pruritus, irritation, burning, odor, dyspareunia, bleeding, and ulcers. A purulent discharge from the cervix should always raise suspicion of upper genital tract infection even in the absence of pelvic pain or other signs.

Normal vaginal discharge is white or transparent, thick, and mostly odorless. It increases during pregnancy, with use of estrogen-progesterin contraceptives, or at mid-cycle around the time of ovulation. Complaints of foul odor and abnormal vaginal discharge should be investigated. Candidiasis, bacterial vaginosis, and trichomoniasis account for 90% of vaginitis cases. The initial workup includes pelvic examination, vaginal pH testing, microscopy, vaginal cultures if microscopy is normal, and gonorrhea/Chlamydia NAAT (see earlier section, “Common Screening and Testing”). The pH of normal vaginal secretions is 3.8 to 4.4, which is hostile to growth of pathogens, and pH greater than or equal to 4.9 is indicative of a bacterial or protozoal infection. Treatment of vaginal infection before anticipated surgery is appropriate, particularly for BV, which may be associated with a higher risk for vaginal cuff infections (Fig. 41-8).

**Bacterial Vaginosis** Bacterial vaginosis (BV) accounts for 50% of vaginal infections. It results from reduction in concentration of the normally dominant lactobacilli and increase in concentration of anaerobic organisms like *Gardnerella vaginalis, M hominis, Bacteroides* species, and others. Diagnosis is made by microscopic demonstration of clue cells. The discharge typically produces a fishy odor upon addition of KOH (amine or Whiff test). Initial treatment is usually a 7-day course of metronidazole.

**Vulvovaginal Candidiasis** Vulvovaginal candidiasis (VVC) is the most common cause of vulvar pruritus. It is generally caused by *C albicans* and occasionally by other *Candida* species. It is common in pregnancy, diabetics, patients taking antibiotics, and in immunocompromised hosts. Initial treatment is usually with topical antifungals, although one dose oral antifungal treatments is also effective.

**Trichomonas Vaginalis** Trichomoniasis is a sexually transmitted infection of a flagellated protozoan and can present with malodorous, purulent discharge. It is typically diagnosed with visualization of the trichomonads during saline wet mount microscopy. Initial treatment is usually a 7-day course of metronidazole.

**Gartner’s Duct Cyst.** A Gartner’s duct cyst is a remnant of the Wolffian tract; it is typically found on the lateral vaginal walls. Patients can be asymptomatic or present with complaints of dyspareunia or difficulty inserting a tampon. If symptomatic, these cysts may be surgically excised or marsupialized. If surgery is planned, preoperative magnetic resonance imaging (MRI) should be obtained to determine the extent of the cyst and verify the diagnosis.

**Vaginal Condyloma.** The etiology and treatment of vaginal condyloma is similar to vulvar condyoma (see earlier section, “Vulvar Condyloma”).

**Vaginal Intraepithelial Neoplasia.** Vaginal intraepithelial neoplasia, or VIN, is similar to VIN and is classified based on the degree of epithelial involvement as mild (I), moderate (II), severe (III), or carcinoma in situ. Upwards of 65% to 80% of VaIN or vaginal cancers are associated with HPV infection. Typically, a patient will have a history of cervical dysplasia and a prior hysterectomy. The majority of lesions are located in the upper one-third of the vagina. Lesions are usually asymptomatic and found incidentally on cytological screening. Biopsy at the time of colposcopy is diagnostic and rules out invasive disease. VaIN is treated with laser ablation, surgical excision, or topical 5-FU therapy.
Cervical Lesions

Benign Cervical Lesions. Benign lesions of the cervix include endocervical polyps, nabothian cysts (clear, fluid filled cysts with smooth surfaces), trauma (such as delivery-related cervical tear or prior cervical surgery), malformation of the cervix, and cervical condyloma. For endocervical polyps, exploration of the base of the polyp with a cotton swab tip to identify that it is cervical and not uterine and to identify the stalk characteristics can help identify the appropriate surgical approach. Small polyps with identifiable base can be removed by grasping the polyp with ring forceps and slowly rotating it until separated from its base. Use of loop electroexcisional procedure (LEEP) is appropriate for larger lesions. Laser or other ablative procedures are appropriate for condyloma proven by biopsy.

Cervical Intraepithelial Neoplasia. Following HPV exposure, dysplastic changes are common. Low grade dysplasia (cervical intraepithelial neoplasia [CIN] I) can be observed and will most often regress to normal within 2 years. However, for girls or women in whom HPV infection is persistent, progression to high-grade cervical dysplasia (CIN II or III) usually require additional treatment due to the high risk of transformation to malignancy. Excisional procedures serve the therapeutic purpose of removal of dysplastic cells, and a diagnostic purpose as histologic review to rule out concomitant early stage cervical cancer can be performed. Either a LEEP or cold knife conization (CKC) may be used for surgical excision of the squamocolumnar junction (SCI) and outer endocervical canal. Risks of both procedures include bleeding, postprocedure infection, cervical stenosis, and risk of preterm delivery with subsequent pregnancies. The benefit of a LEEP is that it can be performed in the office under local anesthesia. A looped wire attachment for a standard monopolar electrosurgical unit is used to perform a LEEP excision. Loops range in a variety of shapes and sizes to accommodate different sizes of cervix. Optimally, one pass of the loop should excise the entire SCI. Hemostasis of the remaining cervix is achieved with the ball electrode and ferrous sulfate paste (Monsel’s solution).

A cervical cold knife conization allows for an excision where the margin status is not obscured by cauterized artifact. This may be particularly useful when the endocervical margin is of interest, or in cases of adenocarcinoma in situ and microinvasive squamous cell carcinoma, where margin status dictates the type and need for future therapy. After injection with dilute vasopressin and the placement of stay sutures at three and nine o’clock on the cervix, a #11 blade is used to circumferentially excise the conical biopsy. Hemostasis is achieved with the cautery or Monsel’s solution.

Uterine Corpus

The average age of menarche, or first menstrual period, in the United States is 12 years and 5 months. Duration of normal menstruation is between 2 to 7 days, with a flow of less than 80 mL, cycling every 21 to 35 days.27 Nonpregnant patients, who present with heavy bleeding and are 35 years of age and older or have risk factors for endometrial cancer, must be ruled out for malignancy as the first step in their management (see earlier section, “Endometrial Biopsy”).

Abnormal Uterine Bleeding. The classification of abnormal uterine bleeding (ABU) has been recently updated.28 Abnormal uterine bleeding may be heavy (AUB/HMB) or intermenstrual (AUB/IMB) and is further divided into acute and chronic categories. Acute AUB is an episode of heavy bleeding that is of sufficient quantity to require immediate intervention to prevent further blood loss. Acute AUB may occur in the setting of chronic AUB. Women with acute AUB should be assessed
rapidly to determine acuity, determine most likely etiology of bleeding, and choose appropriate treatment. Chronic AUB is abnormal uterine bleeding present for most of previous 6 months.

The many causes of AUB are further divided into two categories: structural causes and nonstructural causes. Structural causes include polyps, adenomyosis, leiomyomata, and malignancy. Nonstructural causes can include coagulopathy, ovulatory dysfunction, endometrial effects, and iatrogenic causes. Clinical screening for underlying disorders of hemostasis is recommended in women with heavy menses since menarche, and other risk factors such as bleeding with dental work, epistaxis one or more times per month, or a family history of bleeding symptoms. Poly-, oligo-, and amenorrhea are menstrual cycles of less than 21 days, longer than 35 days, or the absence of uterine bleeding for 6 months or a period equivalent to three missed cycles.

**Endometrial Polyps.** Endometrial polyps are localized hyperplastic growth of endometrial glands and stroma around a vascular core forming sessile or pedunculated projections from the surface of the endometrium. Endometrial polyps are rarely neoplastic (<1%) and may be single or multiple. Many are asymptomatic; however, they are responsible for about 25% of cases of abnormal uterine bleeding, usually metrorrhagia. Polyps are common in patients on tamoxifen therapy and in peri- and postmenopausal women. Up to 2.5% of patients with a polyp may harbor foci of endometrial carcinoma. Diagnosis can be made with saline-infused hysterosonography, hysterosalpingogram, or by direct visualization at the time of hysteroscopy. Definitive treatment, in the absence of malignancy, involves resection with operative hysteroscopy or by sharp curettage.

**Adenomyosis.** Adenomyosis refers to ectopic endometrial glands and stroma situated within the myometrium. When diffuse, it results in globular uterine enlargement secondary to hyperplasia and hypertrophy of the surrounding myometrium. Adenomyosis is very common, tends to occur in parous women, and is frequently an incidental finding at the time of surgery. Symptoms include menorrhagia, dysmenorrhea, and diffuse globular uterine enlargement. MRI typically reveals islands within the myometrium with increased signal intensity. Definitive diagnosis is obtained via hysterectomy and pathologic examination.

**Uterine Leiomyomas.** Leiomyomas, also known colloquially as fibroids, are the most common female pelvic tumor and occurs in response to growth of the uterine smooth muscle cells (myometrium). They are common in the reproductive years, and by age 50. Leiomyomas are described according to their anatomic location (Fig. 41-9) as intramural, subserosal, submucosal, pedunculated, and cervical. Rarely, they can be ectopic. Most are asymptomatic; however, abnormal uterine bleeding caused by leiomyomas is the most common indication for hysterectomy in the United States. Other manifestations include pain, pregnancy complications, and infertility. Pain may result from degenerating myomas that outgrow their blood supply or from compression of other pelvic organs such as the bowel, bladder, and ureters. Hormonal changes during pregnancy can cause significant enlargement of preexisting myomas, which may lead to significant distention of the uterine cavity resulting in recurrent miscarriages, fetal malpresentations, intrauterine growth restriction, obstruction of labor or abnormal placenta- tion, and the subsequent need for cesarean delivery, abortion, preterm labor, and pain from degeneration.

Menorrhagia resulting from leiomyomas can be severe at times, requiring hospitalization or transfusion. Examination typically reveals an enlarged and irregular uterus. Diagnosis is usually made by transvaginal ultrasonography. Other diagnostic modalities, including MRI, computed tomography (CT), and hysterosalpingogram or saline-infused hysterosalpingography, are especially useful in the cases of submucosal and intracervical myomas. Management options of leiomyomas are tailored to the individual patient depending on her age and desire for fertility and the size, location, and symptoms of the myomas. Conservative management options include oral contraceptive pills (OCPs), medroxyprogesterone acetate, GnRH agonists, uterine artery embolization, myomectomy, and hysterectomy. Uterine artery embolization is contraindicated in patients planning future pregnancy and may result in acute degeneration of myomas requiring hospitalization for pain control. Myomectomy is indicated in patients with infertility thought secondary to fibroids and for those with symptomatic fibroids who wish to preserve their reproductive capacity. Hysterectomy is the only definitive therapy. Treatment with GnRH agonists for 3 months prior to surgery may be administered in anemic patients, and it may allow them time to normalize their hematocrit, avoiding transfusions; GnRH also decreases blood loss at hysterectomy and shrinks the myomas by an average of 30%. The latter may make the preferred vaginal surgical approach more feasible.

**Endometrial Hyperplasia.** Endometrial hyperplasia is caused by chronic unopposed hyperestrogenic state (relative absence of progesterone) and is characterized by proliferation of endometrial glands resulting in increased gland-to-stroma ratio. It can be asymptomatic or, more commonly, result in abnormal vaginal bleeding. Hyperplasia can be either simple or complex, based on the architecture of the glands. Of greater importance is the presence or absence of nuclear atypia, described by the WHO classification. A classic retrospective review suggested that untreated endometrial hyperplasia progresses to malignancy in 1%, 3%, 8%, and 29% of cases of simple, complex, simple with atypia, and complex hyperplasia with atypia, respectively. A more modern prospective study noted that of patients who had complex atypical hyperplasia on endometrial biopsy performed prior to hysterectomy, 42.5% had cancer at the time of hysterectomy. Simple and complex hyperplasias can be treated with progesterins, and women should have repeat
endometrial sampling in 3 to 6 months. Atypical hyperplasia is considered a premalignant condition and is treated ideally with simple hysterectomy. If preservation of fertility is desired or surgery is contraindicated, treatment with high-dose progestins such as megestrol acetate 40 to 160 mg per day or with a progestosterone IUD usually reverses these lesions. Close follow-up and repeated sampling are necessary.

The reliability of the pathologic diagnosis of complex atypical hyperplasia is poor, and better and more objective classifications predictive of malignant endometrial behavior are needed. These observations led to the new classification of endometrial intraepithelial neoplasia (EIN). In 2014, the WHO Classification system introduced the diagnosis of EIN into a binary system that aligns with clinical options: hyperplasias are divided into hyperplasia without atypia, and EIN. The new classification is intended to have clinical implications: hyperplasia without atypia may be managed with hormonal therapy, while EIN should be considered a premalignant lesion.

The new classification moves the focus away from cytologic atypia and puts more emphasis on glandular crowding and complexity. While atypia is still important, proliferations can get to EIN without it. For example, the diagnosis of EIN includes cases that lack overt cytologic atypia but show a distinct population from the background epithelium. Morphometric data is utilized to calculate the so-called D-score, which takes into account percentage of stroma, glandular complexity, and gland pleomorphism in an objective manner. A D-score of less than 1 connotes a high rate of progression to endometrial cancer and therefore a diagnosis of EIN. EIN is more predictive than CAH of underlying endometrial malignancy. Most pathology reports are provided with both diagnoses as the transition is made.

Clinicians should be careful to not confuse EIN with endometrial intraepithelial carcinoma (EIC). EIC is a precursor lesion for serous endometrial cancer, and women with a preoperative diagnosis of EIC should always have hysterectomy and appropriate surgical staging performed.

Procedures Performed for Structural Causes of Abnormal Uterine Bleeding

Dilation and Curettage. The patient is placed on the operating table in a lithotomy position, and the vagina and cervix are prepared as for any vaginal operation. The cervix is grasped on the anterior lip with a tenaculum. Some traction on the cervix is necessary to straighten the cervical canal and the uterine cavity. A uterine sound is inserted into the uterine cavity, and the depth of the uterus is noted. The cervical canal is then systematically dilated beginning with a small cervical dilator. Most operations can be performed after the cervix is dilated to accommodate a number 8 or 9 Hegar dilator or its equivalent. Dilatation is accomplished by firm, constant pressure with a dilator directed in the axis of the uterus (Fig. 41-10). The endometrial cavity is then systematically scraped with a uterine curette. Using the largest curette available or suction curettage is a safer choice than a small curette, which tends to cause perforation with less pressure. Uterine perforation is the major complication of dilatation and curettage, diagnosed when the operator finds no resistance to a dilator or curette. Laparoscopy can identify any damage to vessels or bowel if clinically indicated. A uterine perforation through the fundus of the uterus with a dilator or uterine sound is low risk for injury and may be observed without laparoscopy if there is no significant vaginal bleeding noted.
Hysteroscopy. Hysteroscopy, like laparoscopy, has gained widespread support for use both for diagnosis and treatment of intrauterine pathology and for ablation of the endometrium as an alternative to hysterectomy for the treatment of abnormal uterine bleeding. Hysteroscopes can have an objective lens that is offset from the long axis from 0° to 30°.

Diagnostic Hysteroscopy The diagnostic hysteroscope usually has an external diameter of 5 mm. Some diagnostic sheaths allow passage of flexible instruments for biopsy and cutting. Following dilation of the cervix, a diagnostic hysteroscope is placed, and the uterine cavity is distended with the media of choice. Inspection of the cavity includes identifying the uterine fundus, cornua, and any other anomalies to include polyps, leiomyomas, or uterine septum. A dilation and curettage or directed polypectomy with forceps can be performed following identification.

Newer office hysteroscopes can be used to perform hysteroscopy in the office. A paracervical block is placed, and a flexible 3-mm hysteroscope is used. Generally, office hysteroscopy is performed only for diagnostic purposes.

Operative Hysteroscopy An operative hysteroscope is wider than a diagnostic hysteroscope and usually has an integral unipolar or bipolar resecting loop identical to a urologic resectoscope. Electrolyte contacting media are incompatible with conventional monopolar resectoscopic instruments, but electrolyte-free isotonic solutions such as 5% mannitol, 1.5% glycine and 3% sorbitol are acceptable. Large volume deficits have been associated with secondary hyponatremia and hypervolemia due to their metabolism to free water after intravasation. Fluid-management systems are available to monitor the amount of distension media lost during hysteroscopy in order to prevent fluid overload. When fluid deficits reach 1000 to 1500 mL, the procedure should be terminated, and the patient’s serum electrolytes should be assessed. If bipolar instruments are used, resectoscopic instruments can be used without the unique issues related to electrolyte-free hypotonic solutions.

Hysteroscopic Polypectomy Removal of an intrauterine polyp can be performed following diagnostic hysteroscopy through grasping with a polyp forceps. Alternatively, using operative hysteroscopy, the base of the polyp is incised with hysteroscopic scissors. The hysteroscope, sleeve, and polyp are removed simultaneously because most polyps will not fit through the operating channel. Extremely large polyps may have to be removed piecemeal. Any residual base of the polyp may be removed with biopsy forceps.

Endometrial Ablation A common treatment for abnormal uterine bleeding in the absence of endometrial hyperplasia is ablation of the endometrium. Historically, this was performed with an operative hysteroscope using an electrosurgical “roller ball,” where the endometrium was destroyed down to the myometrium in a systematic fashion. Currently, hysteroscopic endometrial ablation has been widely supplanted by various devices, including heated free fluid, cryotherapy, thermal balloon, microwave, and radiofrequency electricity. Most ablation techniques result in amenorrhea in approximately half the patients and decreased menstruation in another third of the patients over the first year of therapy. Subsequent hysterectomy following endometrial ablation is common with rates as high as 40%. Ablation is not recommended in postmenopausal women.

Myomectomy Myomectomy (Fig. 41-11) is the removal of fibroids, and it can be treatment for abnormal uterine bleeding, bulk symptoms, or infertility. Hemostasis during myomectomy can be aided medically by direct injection of dilute vasopressin. Submucosal leiomyoma can be removed safely hysteroscopically. Because myoma tissue is relatively dense, a power cutting instrument is required. The most common method is use of electrosurgery. Both pedunculated and submucosal fibroids are shaven into small pieces with the hysteroscopescope. Stalk resection should only be done to release a pedunculated fibroid if it is 10 mm or less in size; larger fibroids are difficult to remove in one piece without excessive cervical dilatation.

Subserosal, or pedunculated fibroids may require an open or laparoscopic approach depending on the size and location or the leiomyoma. In addition to vasopressin, hemostasis can be further managed through the placement of a Penrose drain around the base of the uterus, pulled through small perforations in the broad ligament lateral to the uterine blood supply on either side and clamped to form a tourniquet for uterine blood flow. An incision is then made through the uterine serosa into the myoma. The pseudocapsule surrounding the tumor is identified, and the tumor is bluntly dissected out with scissors, or bluntly if open. Vessels to the myoma are dissected with the electrosurgical unit. Several myomas may be removed through a single incision, depending upon size. The uterine incisions are then closed with absorbable sutures to obliterate the dead space and provide hemostasis. The uterine serosa is closed with a 3-0 absorbable suture, placed subserosally if possible. Because myomectomies are associated with considerable postoperative adhesion formation, barrier techniques are used to decrease adhesion formation.

During a laparoscopic myomectomy, hemostasis is assisted by intrauterine injection of dilute vasopressin (10 U in 50 mL) at the site of incision, similar to an open procedure. This is usually performed percutaneously with a spinal needle. Pedunculated leiomyomas can be excised at the base using scissors or a power instrument. Intramural leiomyomas require deep dissection into the uterine tissue, which must be closed subsequently with laparoscopic suturing techniques. Removing the specimen may require morcellation; this should be performed after placement of the specimen in a bag. Although power morcellators were previously used for this purpose, an FDA warning in 2014 has virtually eliminated their use. Severe complications including damage to surrounding bowels and vascular structures caused by the spinning blade of the morcellator were reported. Multiple reports of benign tissues such as leiomyoma and endometriosis scattering and dispersing onto abdominal organ surfaces leading to inflammation, infection, and intestinal obstruction often requiring additional surgical interventions and treatments were made. The unintentional dissemination of malignant cells worsens prognosis if an undiagnosed malignancy (most frequently leiomyosarcoma) was morcellated. Although contained morcellation (in a bag) may reduce these risks, informed consent to the patient is prudent.

Total Abdominal Hysterectomy (Fig. 41-12) After the abdomen is entered, the upper abdomen is examined for evidence of extrapelvic disease, and a suitable retractor is placed in the abdominal incision. The uterus is grasped at either cornu with clamps and pulled up into the incision. The round ligament is identified and divided. The peritoneal incision is extended from the round ligament to just past the ovarian hilum, lateral the infundibulopelvic ligament, if the ovaries are to be removed. The retroperitoneal space is bluntly opened, the ureter identified on the medial leaf of the broad ligament, and the
infundibulopelvic ligament isolated, clamped, cut, and suture-ligated; a similar procedure is carried out on the opposite side. If the ovaries are to be left in situ, the ureter is identified and an opening below the utero-ovarian ligament and fallopian tube created. The fallopian tube and utero-ovarian ligament are clamped, cut, and ligated. The bladder is mobilized by sharply dissecting it free of the anterior surface of the uterus and cervix. Clamps are placed on the uterine vessels at the cervicouterine junction, and the vessels are cut and suture-ligated. The cardinal ligaments are then serially clamped, cut, and ligated. Following division of the remaining cardinal ligaments, the uterus is elevated and the vagina clamped. The cervix is amputated from the vagina with scissors or a knife. Sutures are placed at each lateral angle of the vagina, and the remainder of the vagina is closed with a running or interrupted absorbable suture. Pelvic reperitonealization is not necessary.

**Transvaginal Hysterectomy** (Fig. 41-13) Vaginal hysterectomy is the preferred approach in patients in whom the uterus descends and the pubic arch allows enough space for a vaginal operation. A bladder catheter may be placed before the procedure and the patient is placed in a lithotomy position. A weighted vaginal speculum is placed in the vagina, and the cervix is grasped with a tenaculum and pulled in the axis of the vagina. Injection of the cervix and paracervical tissue with analgesic with epinephrine may be helpful in defining planes and decreasing obscuring bleeding. A circumferential incision may be made with a scalpel or scissors. The posterior cul-de-sac is identified and entered with scissors. A long, weighted speculum is then placed through this opening into the peritoneal cavity. Metzenbaum scissors are used to dissect anteriorly on the cervix down to the pubocervical-vesical fascia, reflecting the bladder off the lower uterine segment. When the peritoneum of the anterior cul-de-sac is identified, it is entered with the scissors, and a retractor is placed in the defect. The uterosacral ligaments are identified, doubly clamped, cut, and ligated. Serial clamps are placed on the parametrial structures above the uterosacral ligament; these pedicles are cut and ligated. At the cornu of the uterus, the tube, round ligament, and utero-ovarian ligament of the ovary are doubly clamped and cut. The procedure is carried out usually concurrently on the opposite side, and the uterus is removed. The pelvis is inspected for hemostasis; all bleeding must be meticulously controlled at this point.

The pelvic peritoneum is closed with a running purse-string suture incorporating the uterosacral and ovarian pedicles, those that were held. This exteriorizes those areas that might tend to bleed. The sutures attached to the ovarian pedicles are cut. The vagina may be closed with interrupted mattress stitches,
Figure 41-12. Hysterectomy.
incorporating the uterosacral ligaments into the corner of the vagina with each lateral stitch. On occasion, the uterus, which is initially too large to remove vaginally, may be reduced in size by morcellation (Fig. 41-14). After the uterine vessels have been clamped and ligated, serial wedges are taken from the central portion of the uterus in order to reduce the uterine mass. This procedure will allow the vaginal delivery of even very large uterine leiomyomas.

**Laparoscopic Hysterectomy**

The advantages of laparoscopy over laparotomy include decreased postoperative pain, shorter hospital stays, and reduced blood loss. Laparoscopy has been used to augment vaginal hysterectomy to avoid laparotomy in patients with known pelvic adhesions, endometriosis, or to ensure removal of the entire ovary if oophorectomy is planned or an adnexal mass is present. Over 20% of benign hysterectomies performed in the United States are estimated to be performed laparoscopically.46

Although multiple variations in technique exist, there are three basic laparoscopic approaches for hysterectomy: laparoscopic-assisted vaginal hysterectomy (LAVH), total laparoscopic hysterectomy (TLH), and laparoscopic supracervical hysterectomy (LSH). The technically simplest is the LAVH. A multiple-port approach is used to survey the peritoneal cavity, and any pelvic adhesions are lysed. The round ligaments are then occluded and divided, and the utero-ovarian ligament is divided. The course of the ureter and any adhesions or implants, such as endometriosis that might place the ureter in the way of the surgical dissection, are carefully dissected. Next, the proximal uterine blood supply is dissected for identification and then occluded with a laparoscopic energy device. When the ovaries are removed, the infundibulopelvic ligaments containing the ovarian vessels are divided. If the ovaries are conserved, the utero-ovarian ligament and blood vessels are divided and occluded. In many cases, the posterior cul-de-sac is also incised laparoscopically and the uterosacral ligaments separated with an energy device. The amount of dissection that is done prior to the vaginal portion depends on individual patient characteristics and operator comfort with the vaginal approach, and it may include as little as ovarian and adhesion management to full dissection, including bladder dissection, with only the last vaginal incision done by the vaginal approach. During a TLH, the vaginal incision is performed laparoscopically, and the vaginal incision may be closed with laparoscopic suturing. This procedure is used for the indications listed earlier and also when lack of uterine descent makes the vaginal approach impossible.

**Figure 41-12. (Continued)**
During an LSH, the uterine vessels are divided after the bladder is dissected from the anterior uterus. The ascending branches of the uterine arteries are occluded, and the entire uterine fundus is amputated from the cervix. The endocervix is either cauterized or cored out. The fundus is then morcellated and removed an abdominal port. The end result is an intact cervix, with no surgical dissection performed below the uterine artery. This approach avoids both a large abdominal incision and a vaginal incision. The risks of LSH including subsequent bothersome bleeding from the remaining endometrium or endocervix and cancer risk from the residual cervical stump combining with concerns about power morcellation (see earlier section, “Myomectomy”) have made this procedure less attractive.

Benign Ovarian and Fallopian Tube Lesions
The most common ovarian benign findings include functional follicular cysts, endometriomas (due to ovarian endometriosis), and serous cystadenomas or cystadenofibromas. These can present with varying degrees of pelvic pain, or sometimes be completely asymptomatic. Ultrasound is the best initial imaging modality for evaluating ovarian abnormalities.

Ovarian Cystectomy. When a cystic lesion persists or causes pelvic pain, surgical intervention is usually justified. Performing a cystectomy with ovarian preservation is recommended in women who desire future fertility. Whether the cystectomy is performed laparoscopically or by laparotomy, the procedure is
initiated with inspection of the peritoneal cavity, peritoneum, diaphragm, liver, and pelvis. In the absence of signs of malignancy, pelvic washings are obtained, and the ovarian capsule is incised superficially sharply or with the electrosurgical unit. The cyst is shelled out carefully through the incision. During laparoscopy, it is placed in a bag, intact if possible, and the bag opening is brought through a 10-mm port. If a cyst should rupture before removal, contents are aspirated thoroughly, and the cyst wall is removed and sent for pathologic evaluation. The peritoneal cavity is copiously rinsed with Ringer’s lactate solution. This is especially important when a dermoid cyst is ruptured because the sebaceous material can cause a chemical peritonitis unless all the visible oily substance is carefully removed. A cyst may need to be drained to facilitate removal, but only after bag edges are completely out of the abdomen assuring no leakage within the abdomen. Hemostasis of the ovary is achieved with bipolar electrocoagulation, but the ovary is usually not closed. If there are solid growths within the cyst, it should be sent for frozen section to verify the absence of the malignancy. If malignancy is detected, immediate definitive surgery is recommended.

**Removal of Adnexa.** Indications for removal of adnexae include persistent ovarian cyst, pelvic pain, concern for malignancy, and risk reduction surgery in women with genetic predisposition for ovarian or endometrial cancers (BRCA1/2 mutation carrier, Lynch syndrome). In general, the peritoneum lateral to the infundibulopelvic (IP) ligament is incised in a parallel fashion to allow retroperitoneal dissection and identification of the ureter. Once this has been accomplished, the IP ligament is ligated with suture or an energy source (ultrasonic or bipolar). The remaining posterior leaf of the broad ligament is incised toward the uterus in a direction parallel to the utero-ovarian ligament to avoid ureteral injury. The fallopian tube and utero-ovarian ligaments are then ligated with either suture or an energy source. If performed laparoscopically, the specimen(s) is/are removed in a bag as described earlier.

**Tubal Sterilization.** As in diagnostic laparoscopy, a one- or two-port technique can be used. Fallopian tubes are occluded in the mid-isthmic section, approximately 3 cm from the cornua, using clips, elastic bands, or bipolar electrosurgery. With electrosurgery, approximately 2 cm of tube should be desiccated. Pregnancy rates after any of these techniques have been reported in the range of 3 per 1000 women. Complete removal of the fallopian tube (salpingectomy) at the time of tubal sterilization for the purposes of ovarian cancer prevention has recently become more common.27

A transvaginal tubal occlusion technique may also be used for tubal sterilization. A routine hysteroscopy is first performed to inspect the cavity and identify the tubal ostia. The tubal insert introducer sheath is then placed into the working channel of the hysteroscope. The insert is then threaded into the fallopian tube. Following this procedure, the patient must undergo a hysterosalpingogram to confirm tubal occlusion at 3 months post procedure. Prior to the hysterosalpingogram, the patient is counseled to use a reliable birth control method. Transvaginal tubal sterilization has been associated with perforation of the uterus and/or fallopian tubes, identification of inserts in the abdominal or pelvic cavity, persistent pain, and suspected allergic or hypersensitivity reactions.

**Other Benign Pelvic Pathology**

**Chronic Pelvic Pain.** Chronic pelvic pain is defined as pain below the umbilicus that has lasted at least 6 months or causes functional disability, requiring treatment. While there can be gastrointestinal and urologic causes of chronic pelvic pain, gynecologic causes are frequently identified. Odentimes, a surgical evaluation is needed for diagnosis and/or intervention. The most common gynecologic causes of chronic pelvic pain include endometriosis, adenomyosis, uterine leiomyomas, and adhesive disease.

**Endometriosis** Endometriosis is the finding of ectopic endometrial glands and stroma outside the uterus. It affects 10% of the general population, and it is an incidental finding at the time of laparoscopy in more than 20% of asymptomatic women. Chronic pelvic pain (80%) and infertility (20–50%) are the two most common symptoms.27 The pathophysiology of endometriosis is poorly understood; etiologic theories explaining dissemination of endometrial glands include retrograde menstruation, lymphatic and vascular spread of endometrial glands, and coelomic metaplasia. Endometriosis commonly involves the ovaries, pelvic peritoneal surfaces, and uterosacral ligaments. Other possible sites include the rectovaginal septum, sigmoid colon, intraperitoneal organs, retroperitoneal space, ureters, incisonal scars, umbilicus, and even the thoracic cavity. Involvement of the fallopian tubes may lead to scarring, blockage, and subsequent infertility. Ovarian involvement varies from superficial implants to large complex ovarian masses called endometriomas or “chocolate cysts.” Endometriomas are found in approximately one-third of women with endometriosis and are often bilateral.

While endometriosis can be totally asymptomatic, complaints vary from mild dyspareunia and cyclic dysmenorrhea, to debilitating chronic pelvic pain with dysmenorrhea. Less common manifestations include painful defecation, hematochezia, and hematuria if there is bowel and/or bladder involvement. Catamarnial pneumothorax has been reported from endometriosis implanted in the pleura. Pelvic examination in symptomatic patients typically demonstrates generalized pelvic tenderness, nodularity of the uterosacral ligaments, and at times a pelvic mass may be appreciated if an endometrioma is present. The severity of symptoms does not correlate with the degree of clinical disease present. Endometriosis commonly causes of elevations in serum CA-125. Definitive diagnosis usually requires laparoscopy and visualization of the pathognomonic endometriotic implants. These appear as blue, brown, black, white, or yellow lesions that can be raised and at times puckered giving...
Table 41-4
Centers for Disease Control and Prevention recommended treatment of pelvic inflammatory disease (2015)

<table>
<thead>
<tr>
<th>RECOMMENDED INTRAMUSCULAR/ORAL REGIMENS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ceftriaxone</strong> 250 mg IM in a single dose plus</td>
</tr>
<tr>
<td><strong>Doxycycline</strong> 100 mg orally twice a day for 14 days with* or without</td>
</tr>
<tr>
<td><strong>Metronidazole</strong> 500 mg orally twice a day for 14 days or</td>
</tr>
<tr>
<td><strong>Cefoxitin</strong> 2 g IM in a single dose and <strong>Probenecid</strong>, 1 g orally administered concurrently in a single dose plus</td>
</tr>
<tr>
<td><strong>Doxycycline</strong> 100 mg orally twice a day for 14 days with or without</td>
</tr>
<tr>
<td><strong>Metronidazole</strong> 500 mg orally twice a day for 14 days or</td>
</tr>
<tr>
<td>Other parenteral third-generation <strong>cephalosporin</strong> (e.g., ceftizoxime or cefotaxime) plus</td>
</tr>
<tr>
<td><strong>Doxycycline</strong> 100 mg orally twice a day for 14 days with* or without</td>
</tr>
<tr>
<td><strong>Metronidazole</strong> 500 mg orally twice a day for 14 days</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RECOMMENDED PARENTERAL REGIMENS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cefotetan</strong> 2 g IV every 12 hours plus</td>
</tr>
<tr>
<td><strong>Doxycycline</strong> 100 mg orally or IV every 12 hours or</td>
</tr>
<tr>
<td><strong>Cefoxitin</strong> 2 g IV every 6 hours plus</td>
</tr>
<tr>
<td><strong>Doxycycline</strong> 100 mg orally or IV every 12 hours or</td>
</tr>
<tr>
<td><strong>Clindamycin</strong> 900 mg IV every 8 hours plus</td>
</tr>
<tr>
<td><strong>Gentamicin</strong> loading dose IV or IM (2 mg/kg), followed by a maintenance dose (1.5 mg/kg) every 8 hours. Single daily dosing (3–5 mg/kg) can be substituted.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ALTERNATIVE PARENTERAL REGIMEN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ampicillin/Subbactam</strong> 3 g IV every 6 hours plus</td>
</tr>
</tbody>
</table>
| **Doxycycline** 100 mg orally or IV every 12 hours *

*The addition of metronidazole to treatment regimens with third-generation cephalosporins should be considered until the need for extended anaerobic coverage is ruled out.

Data from Centers for Disease Control and Prevention. 2015 Sexually Transmitted Diseases Treatment Guidelines: Pelvic Inflammatory Disease.

Surgical management for endometriosis varies depending on the age and fertility desires of the patient. A diagnostic laparoscopy with biopsies may be indicated to confirm the diagnosis of endometriosis. If endometriosis is suspected, an operative laparoscopy with ablation of endometriotic implants usually decreases the severity of pelvic pain. Ablation of endometriotic implants can be performed with CO₂ laser or electrosurgery, and/or resection of deep endometriotic implants. Endometriomas can cause pain and if found should be treated by ovarian cystectomy. Complete resection of the cyst wall is required as recurrence of the endometrioma is common after partial removal. Unfortunately, endometriosis is a chronic disease, and conservative therapy, medical or surgical, provides only temporary relief, with the majority of patients relapsing with 1 to 2 years. For patients with severe debilitating symptoms who do not desire future fertility and have not responded to conservative management extirpative surgery to remove the uterus, ovaries, and fallopian tubes; this intervention is curative and should be considered.

Although endometriosis is not generally thought to be a premalignant lesion, there is an increased risk of type I ovarian cancer in women with a history of endometriosis. Molecular evidence that endometriosis is likely a precursor lesion to clear cell carcinoma and endometrioid carcinomas includes the presence of mutations in both PIK3CA and ARID1A in benign endometriotic lesions in close proximity, suggesting that loss of expression of these genes likely occurs early in the development of endometrioid carcinomas.

Pelvic Adhesive Disease
Pelvic adhesions usually are related to previous surgery, endometriosis, or infection, the latter of which can be either genital (i.e., pelvic inflammatory disease) or extragenital (e.g., ruptured appendix) in origin. Adhesions can be lysed mechanically and preferably with minimal cautery.

Pelvic Inflammatory Disease. Pelvic inflammatory disease (PID) is an inflammatory disorder of the upper female genital tract, including any combination of endometritis, salpingitis, tubo-ovarian abscess, and pelvic peritonitis. Sexually transmitted organisms, especially N gonorrhoeae and C trachomatis, are implicated in many cases although microorganisms that comprise the vaginal flora (e.g., anaerobes, G vaginalis, Haemophilus influenzae, enteric Gram-negative rods, and Streptococcus agalactiae) have been implicated as well. PID can additionally result from extension of other pelvic and abdominal infections, such as appendicitis and diverticulitis, or may be precipitated by medical procedure, such as hysterosalpingography, endometrial biopsy, or dilation and curettage.

The presentation of PID can be subtle. Differential diagnosis includes appendicitis, cholecystitis, inflammatory bowel disease, pyelonephritis, nephrolithiasis, ectopic pregnancy, and ovarian torsion. Long-term sequelae can include infertility, chronic pelvic pain, and increased risk of ectopic pregnancy. Because of the severity of these sequelae, presumptive treatment is recommended in young, sexually active women experiencing pelvic or lower abdominal pain, when no cause for the illness other than PID can be identified and if cervical motion tenderness, uterine tenderness, or adnexal tenderness is present on examination. Because of the psychosocial complexity associated with a diagnosis of PID, additional criteria should be used to enhance the specificity of the minimum clinical criteria when possible. These include the following: oral temperature >101°F (>38.3°C); abnormal cervical mucopurulent discharge or cervical friability; presence

them a “gunpowder” appearance. Biopsy is not routinely done but should be obtained if the diagnosis is in doubt.

Treatment is guided by severity of the symptoms and whether preservation of fertility is desired and varies from expectant, to medical, to surgical. Expectant management is appropriate in asymptomatic patients. Those with mild symptoms can be managed with oral contraceptive pills and/or non-steroidal anti-inflammatory analgesia; moderate symptoms are treated with medroxyprogesterone acetate. Severe symptoms are treated with gonadotropin releasing hormone (GnRH) agonists to induce medical pseudomenopause.
of abundant numbers of white blood cells on saline microscopy of vaginal fluid; elevated erythrocyte sedimentation rate; elevated C-reactive protein; and laboratory documentation of cervical infection with *N gonorrhoeae* or *C trachomatis*. Laparoscopy can be used to obtain a more accurate diagnosis of salpingitis and a more complete bacteriologic diagnosis and is often useful in ruling out other causes of peritonitis. Laparoscopic findings may include swollen erythematous tubes with purulent exudates.55

Several outpatient parenteral and oral antimicrobial regimens have been effective in achieving clinical and microbiologic cure. Hospitalization for intravenous antibiotics may be necessitated in cases of where surgical emergencies cannot be ruled out, tubo-ovarian abscess is identified, pregnancy, severe illness (nausea and vomiting, or high fever), inability to follow or tolerate an outpatient oral regimen; or failure of outpatient oral antimicrobial therapy. Treatment of a tubo-ovarian abscess may include placement of a percutaneous drain in addition to intravenous antibiotics.55

Surgical intervention becomes necessary if medical therapy fails or if the patient becomes unstable. Hysterectomy and bilateral salpingo-oophorectomy is the procedure of choice; however, conservative surgery must be considered in young patients desiring future fertility. The abdomen should be explored for metastatic abscesses, and special attention must be paid to bowel, bladder, and ureteral safety due to the friability of the infected tissue and the adhesions commonly encountered at the time of surgery. Placement of an intraperitoneal drain and mass closure of the peritoneum, muscle, and fascia with delayed-absorbable sutures is advised. Conservative surgery, when feasible, may be attempted by laparoscopy and may involve unilateral salpingo-oophorectomy or drainage of the abscess and liberal irrigation of the abdomen and pelvis.53

### PREGNANCY-RELATED SURGICAL CONDITIONS

Many pregnant women will undergo invasive diagnostic procedures for prenatal diagnosis, and in the United States, nearly one-third of all births are cesarean deliveries.56 About 1 in 500 pregnant women will require surgery for nonobstetrical issues.57,58 Diagnostic challenges and physiologic changes due to pregnancy, as well as the unique anesthesia risks and potential risks to the pregnancy, should be kept in mind whether the primary surgeon is an obstetrician, gynecologist, or a general surgeon (Table 41-5).58

Trauma in the obstetric patient requires stabilization of the mother while considering the fetal compartment.58,59 Trauma-related hypovolemia may be compounded by pregnancy-induced decreases in systemic vascular resistance, and when supine, the weight of the gravid uterus on the vena cava. When feasible, a left lateral tilt should be instituted to improve venous return to the right heart. Later in pregnancy, the small bowel is displaced into the upper abdomen, making it vulnerable to complex injury from penetrating upper abdominal trauma. Though small bowel is displaced from the pelvis, the dramatic increase in pelvic blood flow can lead to rapid blood loss due to penetrating pelvic trauma, fractures, or avulsion of pelvic vessels. Gastric motility is decreased increasing the risk of aspiration. Peritoneal signs may be attenuated by the stretching of the abdominal wall. Several coagulation factors are also increased in pregnancy, increasing the likelihood for thromboembolic events, but also giving the unsuspecting surgeon false security when low-normal levels are observed during resuscitative efforts. Only the third trimester fetus has any ability to autoregulate in the context of decreased uterine blood flow and oxygen delivery. In the third trimester, perimortem cesarean delivery should be considered as part of maternal resuscitation in cases of maternal hemodynamic collapse. Though treating the maternal compartment is the primary concern, it should also be recognized that the fetus will be impacted significantly by maternal hypotension, as blood may be shunted away from the uterus.

### Conditions and Procedures Performed Before Viability

#### Amniocentesis/Chorionic Villus Sampling. Noninvasive prenatal testing has for the most part replaced invasive fetal testing. Amniocentesis is a procedure in which amniotic fluid is aspirated from the uterine cavity and sent for genetic or laboratory testing typically under ultrasound guidance with a 20- to 22-gauge needle. This procedure may be used to confirm abnormal noninvasive testing.

#### Miscarriage and Pregnancy Terminations. Spontaneous pregnancy loss is common. Although the miscarriage rate among women who know they are pregnant is roughly 10% to 20%, if the start of pregnancy is set to fertilization, rates are as high as 50%. Chromosomal abnormalities are the underlying cause of miscarriage and are present in over half of cases. Patient may report cramping, bleeding and passage of tissue. If products of conception are not passed, diagnosis can be made by transvaginal ultrasound if an empty gestational sac is identified or an embryo is noted not to have a heartbeat. Treatment can include expectant management, medical management with misoprostol, or surgical management with dilation and curettage.50

Half of all pregnancies in the United States are unintended, and many of these are undesired. Additional reasons for termination of pregnancy include fetal anomalies such as trisomies, fetal infections, and maternal health. Medical terminations are

### Table 41-5

<table>
<thead>
<tr>
<th>Physiologic changes due to pregnancy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cardiovascular changes</strong></td>
</tr>
<tr>
<td>Increased cardiac output</td>
</tr>
<tr>
<td>Increased blood volume</td>
</tr>
<tr>
<td>Increased heart rate</td>
</tr>
<tr>
<td>Decreased blood pressure</td>
</tr>
<tr>
<td>Decreased systemic vascular resistance</td>
</tr>
<tr>
<td>Decreased venous return from lower extremities</td>
</tr>
<tr>
<td><strong>Respiratory changes</strong></td>
</tr>
<tr>
<td>Increased minute ventilation</td>
</tr>
<tr>
<td>Decreased functional residual capacity</td>
</tr>
<tr>
<td><strong>Gastrointestinal changes</strong></td>
</tr>
<tr>
<td>Decreased gastric motility</td>
</tr>
<tr>
<td>Delayed gastric emptying</td>
</tr>
<tr>
<td><strong>Coagulation changes</strong></td>
</tr>
<tr>
<td>Increased clotting factors (II, VII, VIII, IX, X)</td>
</tr>
<tr>
<td>Increased fibrinogen</td>
</tr>
<tr>
<td>Increased risk for venous thromboembolism</td>
</tr>
<tr>
<td><strong>Renal changes</strong></td>
</tr>
<tr>
<td>Increased renal plasma flow and GFR</td>
</tr>
<tr>
<td>Ureteral dilation</td>
</tr>
</tbody>
</table>

available up to 10 weeks of gestation, and surgical terminations can be performed to viability. Rates of pregnancy termination have been declining due decreasing access to abortion services and widespread availability of long-acting contraceptives (LARC). LARCs are safe, effective, easy to use and protect against unintended pregnancy for up to 10 years.61

Up to 15 weeks’ gestation, manual vacuum aspiration can be used following cervical dilation to mechanically evacuate the fetus or embryo, placenta, and membranes by suction using a manual syringe. Alternatively, cervical dilation and suction curettage can be performed. The uterine cervix is grasped with a tenaculum, then mechanically dilated occasionally using adjunctive prostaglandins, and an appropriately sized vacuum cannula is inserted into the uterus and rotated on its axis to remove the products of conception. Dilation and extraction is performed for pregnancies in the second trimester. The additional cervical dilation required at greater gestational ages is usually a two-step (often over 2 days) process. Osmotic dilators are placed within the cervix a day prior to the procedure and expand as water is absorbed, passively dilating the endocervical canal. These are removed immediately prior to the procedure and mechanical dilation is then performed as needed. Forceps are then used to remove fetal parts. Curettage of the postabortal uterus must be approached carefully because the uterus is extremely soft and perforation can occur with very little warning. Complications are rare (particularly when contrasted to the risks of pregnancy and term delivery) but include infection, hemorrhage due to uterine atony, cervical lacerations, uterine perforations, and inadvertent bowel injury from the vacuum cannula or forceps.

**Cerclage.** Cervical insufficiency is defined as painless cervical dilation leading to recurrent second trimester pregnancy loss, or shortened cervical length as determined by transvaginal ultrasound, or advanced cervical change before 24 weeks’ gestation in a woman with either prior preterm birth/loss or significant risk factors for insufficiency. A cervical cerclage refers to a procedure in which suture or synthetic tape is used to circumferentially reinforce the cervix to improve pregnancy outcome in at-risk patients.62 Shirodkar and McDonald techniques have been described63,64; both involve transvaginally placing a non-absorbable suture at the uterocervical junction to lengthen and close the cervix. An abdominal cerclage of the lower uterine segment performed laparor by laparotomy can be considered for a patient with a severely shortened or absent cervix who has previously failed a transvaginal cerclage.

**Ectopic Pregnancies.** Extraterine pregnancies are most commonly located along the fallopian tubes but can also implant on the ovary. Rarely, implantation can occur primarily on other abdominal organs or peritoneal surfaces. A high index of suspicion and early diagnosis typically includes an abnormal rise in β-hCG assays and presence of an adnexal mass on transvaginal ultrasound. Early ectopic pregnancies can be managed medically with a methotrexate injection; however, close follow-up with twice-weekly β-hCG testing is required. Laparoscopy is the definitive management and can be used either as primary treatment or when medical management fails. The tube should be removed (salpingectomy) in its entirety if the ectopic is identified within the fallopian tube. This can be performed using a vessel sealing device or even an endo-loop and endo-shears. Laparotomy is reserved for unstable patients with a known hemoperitoneum where Kelly clamps can be placed along the mesosalpinx to control bleeding. Cornual ectopic pregnancies may require wedge resection of the uterine serosa and myometrium, which is then closed in two layers.65 Linear salpingostomy along the antimesenteric border and removal of the products of conception is now rarely used due to low rates of postoperative tubal function and high recurrent ectopic pregnancies presumably due to scarring.

**Conditions and Procedures Performed After Viability**

**Obstetric Lacerations and Repair.** At the time of vaginal delivery, perineal lacerations are common. These lacerations involve, in varying degrees, the vaginal mucosa, the muscular elements inserting onto the perineal body, the levator ani, and in 4% to 5% of vaginal deliveries, the anal sphincter or anorectal mucosa. Although episiotomies were historically cut prophylactically to prevent unstructured tearing of the perineum, this practice has fallen out of favor as the benefit of episiotomy has not been demonstrated.

**Perineal Laceration** First-degree tears involve only the perineal skin and may or may not need to be reapproximated. Second-degree tears involve the perineal body and can generally be repaired with some variation using a single continuous, nonlocking suture technique, typically a 2-0 or 3-0 synthetic absorbable suture. The apex of the vaginal epithelium is approximated first including epithelium and underlying tissue to build up the rectovaginal septum. Upon reaching the hymenal ring, the perineal body and bulbocavernous muscle are approximated, and a transition stitch is placed from the vaginal mucosa, which was repaired along a horizontal plane, to the deep perineal layer, which lies in a vertically-oriented plane. A running closure is then completed incorporating the deep perineal tissues from the introitus to the extent of the perineal defect.

At this point, the perineal skin is closed from inferior to superior in a subcuticular fashion and tied just inside the introitus.

Third-degree lacerations extend through the perineal body and involve the external anal sphincter, while fourth-degree lacerations involve the internal anal sphincter and rectal mucosa. When present, third- and fourth-degree lacerations should be repaired first before proceeding with the second-degree repair. This is accomplished by first closing the anal mucosa, and then identifying and closing the internal anal sphincter in a second layer. The external anal sphincter is then identified, and the muscular cylinder is reconstructed by suturing the severed ends together using either an end-to-end or overlapping technique. Although these are typically straightforward layered closures, knowledge of the anatomy is important. Incomplete reconstruction, particularly of third- or fourth-degree lacerations, can contribute to future pelvic floor disorders, as well as the development of fistulae or incontinence.

**Cervical and Vaginal Lacerations** Significant lacerations to the cervix or vagina may also occur during childbirth, particularly with instrumented deliveries or macrosomic infants. These lacerations may present as persistent bleeding, not readily recognized due to their location, and often in association with a firmly contracted uterus. Vaginal lacerations may be repaired primarily but should only be closed after deeper tissues are inspected to insure no active bleeding. Cervical lacerations can be repaired in a running, locking fashion, insuring that the apex of the laceration is incorporated in the closure. If the apex is challenging to reach, the closure can be started more distally using the suture to apply traction so that the apex may be closed.
Puerperal Hematoma. Trauma during childbirth can occasionally result in significant hematoma formation with or without a visible laceration. These hematomas may hide significant blood loss and most commonly occur in the vulva, paravaginal, and pelvic retroperitoneum. Typical presentation is pain and mass effect. Small hematomas can be managed conservatively with close observation and patient monitoring. Though there are no evidence-based size criteria, an unstable patient or expanding hematomas should prompt surgical intervention. After the hematoma is incised and drained, diffuse venous oozing is usually encountered rather than a single bleeding vessel. Hemostasis can be achieved using electrocautery or fine absorbable suture, though caution must be used due to the proximity of bowel, bladder, and ureters to some hematomas. Pressure on the vulva or packing the vagina, rather than the hematoma cavity, may prevent further bleeding.

Cesarean Deliveries. Typical indications for cesarean delivery include nonreassuring fetal status, breech or other malpresentations, triplet and higher order gestations, cephalopelvic disproportion, failure to progress in labor, placenta previa, and active genital herpes. Previous low transverse cesarean delivery is not a contraindication to subsequent vaginal birth after cesarean; however, much of the increase in cesarean delivery in the past two decades is attributable to planned repeat cesareans. Cesarean deliveries typically are performed via a lower anterior (caudal) uterine transverse incision because there is decreased blood loss, and the uterine rupture rate with future pregnancies is about 0.5% (Fig. 41-15). A prior classical cesarean delivery is an absolute indication for a planned repeat cesarean delivery because of a high rate of uterine rupture during labor, unlike with the lower anterior uterine transverse incision. Abdominal access is obtained by a Pfannenstiel, Maylard or vertical incision. Once the abdomen is entered, a vesicouterine reflection is created if a low transverse uterine incision is planned. The uterine incision is then made and extended laterally, avoiding the uterine vessels. After amniotomy, the baby is delivered, and the uterus is closed. Approximately 1000 mL of blood is typically lost during a cesarean delivery. Along with rapid closure of the uterine incision, uterotonics, such as intravenous oxytocin, are administered. A classical, vertical, uterine incision is made in certain very early viable gestations, or in the case of certain transverse lies or abnormal placentation. Infection, excessive blood loss due to uterine atony, and urinary tract and bowel injuries are potential complications at the time of cesarean delivery. The risk of those injuries, as well as abnormal placentation (placenta accreta, increta, and percreta) rises with each subsequent cesarean delivery. Bleeding can only be controlled in some instances by performing a cesarean hysterectomy.

Postpartum Hemorrhage. Postpartum hemorrhage is an obstetrical emergency that can follow either vaginal or cesarean delivery. Hemorrhage is usually caused by uterine atony, trauma to the genital tract, or rarely, coagulation disorders. Hemorrhage may also be caused by abnormal placentation (also called morbidly adherent placenta). Management consists of mitigating potential obstetric causes while simultaneously acting to avert or treat hypovolemic shock. In the absence of atony, the genital tract should be thoroughly evaluated for trauma. Atony is the most common cause of postpartum hemorrhage. It is typically treated with fundal massage and uterotonics such as oxytocin, methylergonovine, carboprost tromethamin, and misoprostol. When aggressive medical management fails, surgical management may be necessary and life-saving.

Uterine Curettage. Retained products of conception may result in uterine atony. It may be possible to remove retained products via manual extraction or with ring forceps. Bedside ultrasound may be helpful in localization. When clinical suspicion is high, uterine curettage is indicated. A blunt, large curette, banjo curette, is introduced and removal of retained tissue typically results in contraction of the myometrium and cessation of bleeding.

Procedures Short of Hysterectomy. As bleeding from postpartum hemorrhage becomes increasingly acute, interventions short of hysterectomy should be carried out expeditiously while supporting the hemodynamic status of the patient and preparing for possible definitive surgery. A number of techniques for packing and tamponade of the uterus have been described, including a balloon device reported by Bakri and colleagues.

These are typically left in place for 24 to 36 hours and appear to be safe and often effective conservative measures short of laparotomy and hysterectomy. The B-Lynch compression suture may control bleeding of atony at the time of cesarean section. A suture is placed through the hysterotomy, around the fundus of the uterus anterior to posterior, and then through the posterior lower uterine segment, to the contralateral side. At this point, the steps are reversed with the suture brought around the fundus posterior to anterior, through the contralateral side of the hysterotomy, and then tied in the midline to compress the uterus. Additional procedures described include the O’Leary uterine artery ligation and the hypogastric artery ligation. “O’Leary stitches” are a series of sutures placed around the branches of the uterine artery and through the myometrium, resulting in compression of the vessels against the uterus. Hypogastric artery ligation entails the isolation of the internal iliac artery at its bifurcation with the external iliac artery. The hypogastric artery is ligated at least 3 cm distal to the bifurcation to avoid compromising the posterior division.

Postpartum/Cesarean Hysterectomy. A cesarean or postpartum (absent a prior cesarean delivery) hysterectomy involves the same steps as in a nonpregnant patient, but it is distinctly different due to the engorged vessels and the pliability of the tissues. If a cesarean section has been performed, occasionally the
incision can be used for traction to keep the vessels and tissues attenuated. Vascular pedicles should be secured with clamps, but not ligated until both uterine arteries have been secured, to fully control bleeding. Lack of typical anatomic landmarks requires careful identification of the ureters and the dilated cervix visually or by palpation, to separate from the bladder and vagina (Fig. 41-16). This procedure is often done for life-threatening hemorrhage, thus appropriate blood products, including packed red blood cells, fresh frozen plasma, platelets, and fibrinogen should be on call and are usually required. Fibrinogen is typically elevated in a pregnant woman, such that a low-normal fibrinogen level can be cause for alarm, and further fibrinogen may be required before consumptive coagulopathy reverses. A massive transfusion protocol is helpful.

Abnormal Placenta...tion rate in the United States. When cytotrophoblasts invade decidualized endometrium and encounter a uterine scar, they do not encounter the normal myometrial signals to stop invasion. In the setting of a placenta previa, the presence of a uterine scar is a particular risk for placenta accreta with rates of 11%, 40%, and 61% for one, two, or three prior cesarean deliveries, respectively. Ultrasound or MRI can assist in the diagnosis, depending on the experience and comfort of the imager.

Women at risk for abnormal placenta should ideally be identified during pregnancy and be prepared for cesarean section followed by cesarean hysterectomy. Since the blood supply to the gravid uterus is 500 cc per minute, these surgeries have the potential to have very high blood loss, which can then lead to the development of disseminated intravascular coagulation. Over 50% of cases require more than 4 units of blood transfused.

Unintentional bladder or ureteral injuries are common as well due to impaired visualization and poor dissection planes. For these reasons, patients with suspected placenta accreta should be delivered in a tertiary care center with a multidisciplinary team that has the capacity for massive blood transfusion protocol. While some sites have implemented protocols involving interventional radiology with placement of occlusive balloons in the uterine arteries prior to delivery, these protocols have not been shown to decrease morbidity or overall blood loss. Postoperative embolization should be available. Even with scheduled delivery in a well-resourced setting with a highly experienced and prepared multidisciplinary team, the morbidity of abnormal placentation is high. ICU stays are common, and maternal mortality as high as 7% has been reported.

Delayed hysterectomy where the placenta is left in situ after delivery of the baby if there is not significant bleeding and the mother is stable is advocated by certain centers but remains controversial. The risks of leaving the placenta in utero include later hemorrhage, infection, and sepsis. Planned hysterectomy at 6 to 12 weeks postpartum is recommended unless subsequent fertility is strongly desire.

PELVIC FLOOR DYSFUNCTION

Pelvic floor disorders can be categorized, from a urogynecologic perspective, into three main topics: female urinary incontinence and voiding dysfunction, pelvic organ prolapse, and disorders of defecation. Approximately 11% of women will undergo surgery for incontinence or prolapse. The normal functions of support, storage, and evacuation can be altered by derangements in neuromuscular function both centrally and peripherally and through acquired changes in connective tissue. Reconstructive surgeons aim to repair or compensate for many of these losses.

Evaluation

Diagnostic evaluations, in addition to the history and examinations previously described, can aid in the diagnosis of many pelvic floor disorders. Cystoscopy, multichannel urodynamics, and/or fluoroscopic evaluation of the urinary tract can be obtained for patients with urinary incontinence or voiding dysfunction. Defecography, anal manometry, and endorectal ultrasound may be useful for diagnosis of defecatory dysfunction. A standardized examination called the pelvic organ prolapse quantification (POP-Q) helps to clarify which vaginal compartment, and therefore which specific structure, has lost its anatomic integrity in women with uterovaginal prolapse. Finally, dynamic MRI and pelvic floor electromyography has growing utility for all three disorders.

Surgery for Pelvic Organ Prolapse

Many factors are important in determining which reconstructive operation is optimal for a given patient with pelvic organ prolapse. Surgical decisions are often based on case series and expert opinions that may not have universal applicability. However, the few reports with the highest level of evidence suggest that failure rates for prolapse reconstruction may be twice as high using the vaginal approach when compared with the abdominal route.

Colporrhaphy. Anterior colporrhaphy, also known as an “anterior repair,” is performed for a symptomatic cystocele. The procedure begins with incision of the anterior vaginal epithelium...
in a midline sagittal direction. The epithelium is dissected away from the underlying vaginal muscularis. The vaginal muscularis is plicated with interrupted delayed absorbable stitches, after which the epithelium is trimmed and reapproximated. The vaginal canal is therefore shortened and narrowed proportionate to the amount of removed epithelium. Posterior colporrhaphy is performed for a symptomatic rectocele. This procedure is performed in a similar manner, often including the distal pubococcygeus muscles in the plication. Recently, in attempts to decrease surgical failures alluded to previously, many surgeons have opted to utilize grafts and meshes to augment these vaginally performed procedures. Unfortunately, the apparent number of postoperative complications, including mesh erosion, pelvic pain, and dyspareunia, prompted the FDA to publish a warning encouraging a much more limited use of vaginal mesh for prolapse repair until greater surveillance and more rigorous studies could be completed.77

**Sacropinous and Uterosacral Ligament Fixations.** Both the sacropinous ligament fixation (SSLF) and uterosacral ligament fixation (USLF) procedures are vaginal procedures that suspend the apex of the vagina using native tissue for treatment of apical prolapse. The sacropinous ligament is found embedded in and continuous with the coccygeus muscle, which extends from the ischial spine to the lateral surface of the sacrum. The procedure begins with entry into the rectovaginal space, usually by incising the posterior vaginal wall at its attachment to the perineal body. The space is developed to the level of the vaginal apex and the rectal pillar is penetrated to gain access to the pararectal space. A long-ligature carrier is used to place sutures medial to the ischial spine, through the substance of the ligament-muscle complex. Structures at risk in this procedure include the pudendal neurovascular bundle, the inferior gluteal neurovascular bundle, the lumbosacral plexus, and sciatic nerve. After the stitches are placed, the free ends are sewn to the undersurface of the vaginal cuff. The sacropinous stitches are tied to firmly approximate the vagina to the ligament without suture bridging.

When using the uterosacral ligaments for repair of prolapse, it is important to recall that these structures are not “ligaments” in the true sense of the word, but rather condensations of smooth muscle, collagen, and elastin. Several support sutures are placed from the lateral-most portion of the vaginal cuff to the distal-most part of the ligament, and the medial vaginal cuff to the proximal ligament. Intraoperative evaluation of the lower urinary tract is important to confirm the absence of ureteral compromise.

**Colpocleisis.** Colpocleisis is reserved for patients who are elderly, who do not wish to retain coital ability, and for whom there is good reason not to perform a more extensive reconstructive operation. A colpocleisis removes part or all of the vaginal epithelium, obliterating the vaginal vault and leaving the external genitalia unchanged. The procedure can be performed with or without a hysterectomy. Successive purse-string sutures through the vaginal muscularis are used to reduce the prolapsed organs to above the level of the levator plate.

**Sacrocolpopexy.** The procedure with the lowest risk of recurrence for patients with prolapse of the vaginal apex is an abdominal sacral colpopexy. In these patients, the natural apical support structure, the cardinal–uterosacral ligament complex, is often damaged and attenuated. The abdominal placement, as opposed to vaginal placement, of graft material to compensate for defective vaginal support structures is well described.78 Apical support defects rarely exist in isolation, and the sacrocolpopexy may be modified to include the anterior and posterior vaginal walls as well as the perineal body in the suspension. Sacrocolpopexies can be performed via laparotomy as well as via laparoscopy or robotically. Like rectopexies and low anterior resections, deep pelvic access is needed. Significant suturing at varied angles is required. The advent of the DaVinci robotic laparoscopic system has made visualization and adequate placement of the mesh and sutures easier to perform when using the minimally invasive approach.

During a sacrocolpopexy, a rigid stent (usually an EEA sizer) is placed into the vagina to facilitate its dissection from the overlying bladder and rectum and to allow the graft material to be spread evenly over its surface. A strip of synthetic mesh is fixed to the anterior and posterior vaginal walls. The peritoneum overlaying the presacral area is opened, extending to the posterior cul-de-sac. The sigmoid colon is retracted medially, and the anterior surface of the sacrum is skeletonized. Two to four permanent sutures are placed through the anterior longitudinal ligament in the midline, starting at the S2 level and proceeding distally. The sutures are passed through the graft at an appropriate location to support the vaginal vault without tension. The peritoneum is then closed with an absorbable running suture. The most dangerous potential complication of sacrocolpopexy is sacral hemorrhage.

**Surgery for Stress Urinary Incontinence**

Stress incontinence is believed to be caused by lack of urethrovaginal support (urethral hypermobility) or intrinsic sphincter deficiency (ISD). ISD is a term applied to a subset of stress-incontinent patients who have particularly severe symptoms, including urine leakage with minimal exertion. This condition is often recognized clinically as the low pressure or “drainpipe” urethra. The urethral sphincter mechanism in these patients is severely damaged, limiting coaptation of the urethra. Standard surgical procedures used to correct stress incontinence share a common feature: partial urethral obstruction that achieves urethral closure under stress.

**Burch Procedure.** Despite the wide acceptance of midurethral sling procedures, a retropubic urethropexy procedure called the Burch procedure is still performed for stress incontinence.79 The space of Retzius is approached extraperitoneally, from an abdominal approach, allowing the bladder to be mobilized from the surrounding adipose tissue and lateral pelvis. Two pairs of large-caliber nonabsorbable sutures are placed through the periurethral vaginal wall, one pair at the midurethra and one at the urethrovesical junction. Each stitch is then anchored to the ipsilateral Cooper’s (iliopubic) ligament. The sutures are tied to give preferential support to the urethrovesical junction relative to the anterior vaginal wall without overcorrection. Long-term outcome studies up to 10 years have shown the Burch procedure yields cure rates of 80% to 85%.

**Tensionless Sling.** The tension-free vaginal tape (TVT) is a modified sling that uses a strip of polypropylene mesh. Unlike traditional sling procedures, the mesh is positioned at the midurethra, not the urethrovesical junction, and it is not sutured or otherwise fixed into place. Advantages of TVT include the ability to perform the procedure under local anesthesia on an outpatient basis. Small subepithelial tunnels are made bilaterally to the descending pubic rami through an anterior vaginal wall incision. A specialized conical metal needle coupled to a handle is used to drive one end of the sling through the perineal membrane, space of Retzius, and through one of two small suprapubic stab incisions. The tape is set in place without any
tension after bringing up the other end of the tape through the other side. Recently, multiple modifications have been made to carry the tape through the bilateral medial portions of the obturator space (TVT-O). Risks of the procedure include visceral injury from blind introduction of the needle, bleeding, and nerve and muscle injury in the obturator space. Additionally, voiding dysfunction and delayed erosion of mesh into the bladder or urethra has been seen.

**Urethral Bulking Injections.** A transurethral or periurethral injection of bulking agents is indicated for patients with intrinsic sphincter deficiency. Several synthetic injectable agents, such as polydimethylsiloxane and calcium hydroxyapatite are now used, as glutaraldehyde cross-linked (GAX) bovine dermal collagen is no longer commercially available.\(^8^0\) Anesthesia is easily obtained by using intraurethral 2\% lidocaine jelly and/or transvaginal injection of the perirectal tissues with 5 mL of 1\% lidocaine. The material is injected underneath the urethral mucosa at the bladder neck and proximal urethra at multiple positions, until mucosal bulk has improved. Patients must demonstrate a negative reaction to a collagen skin test prior to injection. The long-term cure rate is 20\% to 30\%, with an additional 50\% to 60\% of patients demonstrating improvement.\(^7^2\) Repeat injections are frequently necessary because of migration and dissolution of the collagen material.

**Mesh in Reconstructive Pelvic Surgery.** As noted earlier, pelvic reconstructive surgery frequently uses polypropylene mesh to augment procedures in the hopes of providing lasting repair. However, use of permanent mesh is associated with complications, most notably mesh erosion. In 2011, the FDA issued an updated statement to stipulate the risks when using transvaginally inserted mesh for prolapse.\(^8^1\) Ultimately, this has led to categorizing transvaginal mesh products as class III devices in 2016. In addition to appropriate patient selection, and extensive informed consent, the American Urogynecologic Society recommends appropriate training to perform the procedures and manage the complications.\(^8^2,8^3\)

**GYNECOLOGIC CANCER**

**Vulvar Cancer**

Vulvar cancer is the fourth most common gynecologic cancer. The mean age at diagnosis is 65, though this has trended down over the last several decades.\(^8^4\) Evidence supports an HPV-dependent pathway of carcinogenesis with risk factors similar to VIN in approximately 60\% of cases. A second pathway independent of HPV is associated with chronic inflammation, vulvar dystrophy.\(^8^5\) Patients usually present with a vulvar ulcer or mass. Pruritis is a common complaint, and vulvar bleeding or enlarged inguinal lymph nodes are signs of advanced disease. Careful evaluation of the patient is necessary to rule out concurrent lesions of the vagina and cervix. Biopsy is required and should be sufficiently deep to allow evaluation of the extent of stromal invasion. Vulvar carcinomas are squamous in 90\% of cases. Other less common histologies include melanoma (5\%), basal cell carcinoma (2\%), and soft tissue sarcomas (1–2\%).

Spread of vulvar carcinoma is by direct local extension and via lymphatic microembolization. Hematogenous spread is uncommon except for vulvar melanoma. Lymphatic spread seems to follow a stepwise, predictable pattern traveling from superficial, above the cribiform fascia, to deep inguino-femoral nodes and ultimately the pelvic, external iliac, nodal basin
radical local excision are adequate. Patients with IB tumors have deeper invasion but negative nodes and therefore carry an excellent prognosis. Stage II includes patients with local extension and negative nodes and therefore carry a prognosis similar to other node-negative patients.

Stage III disease includes patients with lymph node metastases, and stage IV disease is either locally advanced or distant metastasis. Treatment options for stage III and stage IV disease include (a) chemoradiation followed by limited resection if needed; (b) radical vulvectomy; and (c) radical vulvectomy coupled with pelvic exenteration. External beam radiotherapy combined with radiosensitizing chemotherapy of cisplatin and 5-fluorouracil (5-FU) is emerging as the preferred initial management of advanced disease, followed by limited surgical resection of residual disease. Reconstruction of the vulva and groin, if needed, can be accomplished using grafts and rotational or myocutaneous flaps depending on the size and type of defect.

Inguinoemoral lymphadenectomy is indicated beyond clinical stage IA. Unilateral lymphadenectomy is recommended for lateralized lesions or bilateral for central lesions that cross the midline, or those involving the perirectal area (Figs. 41-19 and 41-20). Complications of complete inguinoemoral lymphadenectomy include wound dehiscence or infection and lymphedema. Sentinel lymph node biopsy (SLNB) is an alternative to inguinoemoral lymphadenectomy for selected patients with stage I or II disease and no palpable inguinoemoral nodes. SLNB appears to be effective in detecting inguinoemoral lymph node metastases without increasing the risk of groin recurrence while avoiding the morbidities associated with complete inguinoemoral lymphadenectomy. Several prospective studies support this approach. However, it is recognized that successful SLNB depends on operator experience. Surgeons with limited experience in SLNB (have performed fewer than 10 of these procedures) may choose to perform complete groin node dissection or use this procedure only for tumors that are less than 2 cm in size.

Nodal failure in the groin and pelvis is difficult to treat successfully, and attention to primary management of these areas is key. Postoperative adjuvant inguinal and pelvic radiotherapy is indicated when inguinal lymph nodes are positive and is superior to pelvic lymphadenectomy, which has been largely abandoned. It is also indicated when the vulvectomy margins are positive or close positive for disease and further surgical management is not anatomically feasible.

**Vaginal Cancer**

Vaginal carcinoma is a rare gynecologic malignancy and accounts for about 3% of cancers affecting the female reproductive system. Squamous cell carcinomas account for 85% to 90% of cases; more than two-thirds of vaginal cancers are diagnosed in women 60 years of age or older. Risk factors are similar to other HPV-related cervical and vulvar cancers. Rare clear cell carcinoma of the vagina is associated to in utero exposure to diethylstilbestrol (DES), which is now largely of historical interest due to aging of the exposed cohort. Patients with vaginal cancer usually present with postmenopausal and/or postcoital bleeding and may also complain of vaginal discharge, vaginal mass, dysuria, hematuria, rectal bleeding, or pelvic pain, which may be indicative of advanced disease. Diagnosis is made via biopsy of suspicious lesions, which may require colposcopic guidance.
Vaginal cancer is staged clinically by pelvic exam, chest X-ray, cystoscopy, and proctoscopy (Table 41-7). Vaginal cancer spreads by local extension to adjacent pelvic structures, by lymphatic embolization to regional lymph nodes, and, less commonly, via the hematogenous route. Lymphatic drainage is complex, but in general, lesions in the upper vagina drain to the pelvic lymph nodes while lesions involving the lower third drain to the inguinofemoral lymph nodes.

Stage I disease, involving the upper vagina, may be treated surgically or with intracavitary radiation therapy. Surgery consists of a radical hysterectomy, upper vaginectomy, and bilateral pelvic lymphadenectomy. Stage I disease in the mid to lower vagina is treated with radiation and concurrent chemotherapy. External beam pelvic radiation is the mainstay of treatment for stages II to IV and may be followed by intracavitary radiation fields for women with locally advanced disease. Staging factors in this disease, and positron emission tomography scans are useful in pretreatment planning and determination of radiation fields for women with locally advanced disease. Staging and management options are outlined in Table 41-8.

### Table 41-8
#### 2009 FIGO cervical cancer staging and management options

<table>
<thead>
<tr>
<th>STAGE</th>
<th>DESCRIPTION</th>
<th>OPTIONS FOR MANAGEMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Carcinoma in situ</td>
<td>Adenocarcinoma in situ: simple hysterectomy, may be followed for fertility preservation if all margins negative on cone Squamous cell carcinoma in situ: local excision withLEEP or cone or laser ablation</td>
</tr>
<tr>
<td>I</td>
<td>Confined to the cervix A1: Confined to the cervix, diagnosed only by microscopy with invasion of ≤3 mm in depth and lateral spread ≤7 mm A2: Confined to the cervix, diagnosed with microscopy with invasion of &gt;3 mm and ≤5 mm with lateral spread ≤7 mm B1: Clinically visible lesion or greater than A2, ≤4 cm in greatest dimension B2: Clinically visible lesion, &gt;4 cm in greatest dimension</td>
<td>A1 and some A2: fertility preservation through large cone followed by close monitoring, followed by hysterectomy B1 and B2: radical hysterectomy or chemoradiation; radical trachelectomy with uterine preservation for childbearing is under investigation for highly selected patients with small lesions</td>
</tr>
<tr>
<td>II</td>
<td>A1: Involvement of the upper two-thirds of the vagina, without parametral invasion, ≤4 cm in greatest dimension A2: &gt;4 cm in greatest dimension B: Parametrial involvement</td>
<td>For some IIA radical hysterectomy may be considered IIA and B: chemoradiation is preferred</td>
</tr>
<tr>
<td>III</td>
<td>A. Involvement of the lower third of the vagina B. Involvement of a parametria to the sidewall or obstruction of one or both ureters on imaging</td>
<td>Chemoradiation</td>
</tr>
<tr>
<td>IV</td>
<td>A. Local involvement of the bladder or rectum B. Distant metastases</td>
<td>A. Chemoradiation B. Chemotherapy with palliative radiation as indicated</td>
</tr>
</tbody>
</table>

Data from Pecorelli S: Revised FIGO staging for carcinoma of the vulva, cervix, and endometrium, Int J Gynaecol Obstet. 2009 May;105(2):103-104.
Procedures for Cervical Cancer Treatment. Certain cervical cancers that are confined to the cervix may be treated surgically. Very small lesions (less than 7 mm wide, less than 3 mm deep) with no LVSI may be treated with simple hysterectomy. In a woman who desires future fertility, a cone biopsy with negative surgical margins may be an acceptable alternative. Any tumor larger than this (larger than stage IA1) should be treated with radical hysterectomy or in special cases radical tracheectomy for fertility preservation. Some authors advocate a large cone biopsy with lymph node dissection for stage IA2 tumors in patients who desire future fertility, though this recommendation is somewhat controversial. Tumors that are greater than 4 cm in size are most often treated with chemoRT even if they are confined to the cervix, given the high likelihood of need for postoperative radiotherapy due to cervical risk factors.

Radical Hysterectomy This procedure may be performed via laparotomy, or increasingly via a minimally invasive (laparoscopic or robotic) approach. The key elements are dissection of the pelvic and paraaortic nodes and the dissection of the parametrium from the pelvic sidewall to allow en bloc removal with the uterus. The principle steps of an open procedure are demonstrated in Fig. 41-21. In contrast to a typical simple hysterectomy, the radical hysterectomy involves dissection much closer to the bowel, bladder, ureters, and great vessels, resulting in a higher complication rate to these organs. Additionally, disruption of the
nerves supplying the bladder and the rectum, which traverse the cardinal and uterosacral ligaments, may result in temporary or long-term bladder and bowel dysfunction. Radical hysterectomies allow for the maintenance of the ovaries since the incidence of metastases to this area is very low, providing a clear advantage of surgery over radiation therapy in the younger patient.

**Radical Trachelectomy** Interest in fertility preservation with stages IA1 and 2, and stage IB1 lesions has led to the development of methods of radical trachelectomy with uterine preservation. This procedure depends on an adequate blood supply to the uterus from the ovarian anastamoses, as the cervical portion is removed. The lower uterine segment closed with a cerclage and attached directly to the vaginal cuff. The rates of recurrence, pregnancy outcomes, and the best surgical candidates for this surgery are still under study, but there are sufficient numbers and experience, both obstetric and surgical, to suggest that this procedure is oncologically safe and allows live births.

**Pelvic Exenteration for Recurrent Disease (Fig. 41-22)** Cervical cancer recurrences after primary surgical management are treated with radiation. Surgery may be a consideration in selected patients with recurrent cervical cancer who have received maximal radiation therapy. If the recurrence is locally confined with no evidence of spread or metastatic disease, then pelvic exenteration may be considered. Attempted exenteration procedures are aborted intraoperatively if metastatic disease is found. Exenteration is tailored for the disease size and location and may be supravelvator or extend below the levator ani muscle and require vulvar resection. Reconstruction of the pelvis may require a continent urinary pouch (if radiation enteritis is limited) or ileal conduit and colostomy, as well as rebuilding of the pelvic floor and vagina with grafts or myocutaneous flaps.

**Uterine Cancer**

**Endometrial Cancer.** Endometrial cancer is the most common gynecologic malignancy and fourth most common cancer in women. It is most common in menopausal women in the fifth decade of life; up to 15% to 25% of cases occur prior to menopause, and 1% to 5% occur before age 40. Risk factors for the most common type of endometrial cancer include increased exposure to estrogen without adequate opposition by progesterone, either endogenous (obesity, chronic anovulation) or exogenous (hormone replacement). Additional risk factors include diabetes, Lynch II syndrome (hereditary nonpolyposis coli syndrome), and prolonged use of tamoxifen. Tamoxifen is a mixed agonist/antagonist ligand for the estrogen receptor. It is an agonistic in the uterus and an antagonistic to the breast and ovary. Protective factors for endometrial cancer include smoking and use of combination oral contraceptive pills. Adenocarcinomas are the most prevalent histologic type.

Endometrial adenocarcinomas have historically been divided into type I and type II tumors with five classic histologic subtypes. Type I tumors are estrogen-dependent endometrioid...
histology and have a relatively favorable prognosis; they can be broken down further by presence or absence of microsatellite instability. Type II endometrial cancers are estrogen-independent, aggressive, and characterized by nonendometrioid, serous or clear cell, histology, or carcinosarcoma. Emerging data, however, suggest that the molecular features could provide reproducible subtypes that have the potential to guide and refine treatment. The most comprehensive molecular study of endometrial cancer to date has been The Cancer Genome Atlas, which included a combination of whole genome sequencing, exome sequencing, microsatellite instability assays, copy number analysis, and proteomics. Molecular information was used to classify 232 endometrial cancer patients into four groups: POLE ultramutated, MSI hypermutated, copy number low, and copy number high that correlated with progression-free survival. Two practical pared-down classification systems to identify four molecular subgroups with distinct prognostic outcomes have been described.

Postmenopausal bleeding is the most common presentation of endometrial cancer and often permits early stage diagnosis, resulting in a favorable prognosis. Abnormal bleeding should prompt endometrial evaluation and sampling, which is usually done with an office endometrial biopsy, though at times requires operative curettage or diagnostic hysteroscopy. Transvaginal ultrasonography (TVUS) often reveals a thickened endometrial stripe. An endometrial stripe measuring 5 mm or more in a postmenopausal patient with vaginal bleeding raises concern and should be followed by endometrial sampling; patients with stripe of 4 mm or less rarely have occult malignancy, and TVUS may thus be used to triage patients before invasive endometrial sampling. Even with a normal endometrial stripe, endometrial sampling should be performed for persistent postmenopausal bleeding. Uterine cancer is surgically staged and is graded based on the degree of histologic differentiation of the glandular components (Table 41-9). Grade is an important prognostic factor, independent of stage.

Treatment is surgical, and most commonly involves hysterectomy, bilateral salpingo-oophorectomy, peritoneal cytology, and resection of any gross disease. Evidence supports equivalent oncologic outcomes with minimally invasive approaches. The inclusion and utility of lymphadenectomy remains an area of controversy. If a lymph node dissection is performed, it may be performed via laparotomy or laparoscopy. Generally, the bilateral pelvic and para-aortic lymph nodes are removed. The pelvic node dissection includes: bilateral removal of nodal tissue from the distal one-half of each common iliac artery, the anterior and medial aspect of the proximal half of the external iliac artery and vein, and the distal half of the obturator fat pad anterior to the obturator nerve. Most of the pelvic lymph nodes lie anterior, medially, and posteriorly to the external and internal iliac vessels and the obturator nerve. There are a few nodes that lie lateral to these structures, between the vessels and the pelvic sidewall, and are generally removed in a complete dissection. The para-aortic lymph nodes include resection of nodal tissue over the distal vena cava from the level of the inferior mesenteric artery to the mid right common iliac artery and between the aorta and the left ureter from the inferior mesenteric artery to the left mid common iliac artery. Some also advocate resection of lymph nodes between the IMA and the gonadal vessels, as some uterine fundal tumors may drain directly into these lymph nodes.

The need for postoperative intervention is individualized based on the histology, stage, and risk factors such as age, lymphvascular space invasion, and histology. Early-stage patients...
are typically cured with surgery alone, while patients with high-intermediate risk factors, as defined by collaborative trials groups, commonly receive intracavitary brachytherapy to decrease local recurrence.\textsuperscript{108,109} Patients with advanced disease and high-grade histologies commonly receive platinum-based chemotherapy with or without radiation.

Similar to the case with vulvar cancer described earlier, sentinel node biopsy is becoming more prevalent in endometrial cancer. A sentinel lymph node biopsy may be considered in apparent uterine-confined malignancy when there is no metastasis demonstrated by imaging studies or no obvious extratumor disease at exploration. For this procedure, most frequently the cervix is injected with ICG dye, and the immunofluorescence detecting camera is used eitherrobotically or laparoscopically to identify the sentinel node. If no node is mapped, a full lymphadenectomy is generally advised.\textsuperscript{110}

**Lynch Syndrome.** Lynch syndrome, a cancer family syndrome also known as hereditary nonpolyposis colorectal cancer (HNPCC), is an autosomal dominant inherited predisposition to develop colorectal carcinoma and extracolonic cancers, predominantly including tumors of the uterus and ovaries, now also including breast cancer.\textsuperscript{111} Genes involved in HNPCC are those required for proper single-strand DNA repair via the mismatch repair pathway; most commonly involved are \textit{MLH1}, \textit{MSH2}, \textit{MSH6}, and \textit{PMS2}. The risk of colorectal carcinoma is as high as 75% by age 75 years. Affected women have a 40% and 10% lifetime risk of developing uterine and ovarian cancers, respectively. Surveillance has not been proven to identify disease in early stage for these patients, though it is recommended and should include annual cervical cytology, mammography, transvaginal ultrasonography, CA-125 measurements, and an endometrial biopsy. Risk-reducing salpingo-oophorectomy with hysterectomy is now being recommended for women who have completed childbearing, ideally 5 to 10 years earlier than the first case of endometrial or ovarian cancer in the family. Dysregulation of the mismatch repair pathway leads to the microsatellite instability phenotype, now known to be associated with susceptibility to select immunotherapies.

**Uterine Sarcomas.** Uterine sarcomas arise from the uterine muscle and connective tissue elements and are typically aggressive tumors with a poorer prognosis compared to the more common endometrial carcinomas. The most common histopathologic types are endometrial stromal sarcomas, undifferentiated endometrial sarcomas, and leiomyosarcomas. Risk factors are challenging to assess but may include prior pelvic radiation and tamoxifen exposure. Patients typically present with bleeding or mass effects, although some are discovered incidentally at the time of hysterectomy for other indications. Leiomyosarcoma is the most common uterine sarcoma, and hysterectomy with salpingooophorectomy is the treatment of choice. Lymph node metastases are rare in sarcomas in general, and in the absence of palpable nodes or extrauterine disease. There are limited data to support cytoreduction when extraterine disease is present. The benefits of adjuvant therapy are unknown. Advanced disease is typically treated with systemic chemotherapy.\textsuperscript{112}

**Ovarian Cancer**

**Epithelial Ovarian, Tubal, and Primary Peritoneal Cancer.** Ovarian cancer is a rare disease affecting 1 in 70 women with a median age at diagnosis of 62 years.\textsuperscript{90} Epithelial malignancies make up the vast majority of ovarian cancers. The majority of women (70%) are diagnosed with advanced staged disease leading to the poor survival associated with this malignancy. Survival in advanced disease is due both to late diagnosis and lack of effective second-line cytotoxic therapy for the majority of patients who relapse following initial clinical complete response to platinum-based chemotherapy. Despite multiple prospective population based trials evaluating the use of CA-125, ultrasound, or combinations of these tests for early detection of disease, a mortality benefit to screening programs has not been demonstrated.\textsuperscript{113-116} Symptoms for either benign or malignant ovarian tumors are nonspecific but frequent, and they include bloating, pelvic or abdominal pain, difficulty eating or feeling full quickly, and urinary symptoms of urgency or frequency,\textsuperscript{117} which form the basis of an ovarian cancer symptom index (Table 41-10). When newly developed and persistent, these symptoms should prompt an evaluation specifically targeted for identification of gynecologic malignancy.

The histologic heterogeneity of ovarian cancer has long been recognized, but with the emergence of more robust clinicopathologic, molecular, and genetic data over the past decade these distinctions have become more clearly defined. Type I tumors consist of low-grade serous (LGS), low-grade endometrioid, clear cell carcinomas (CCC), and mucinous carcinomas and are characterized by mutations in \textit{KRAS}, \textit{BRAF}, \textit{PTEN}, \textit{PIK3CA}, \textit{CTNNB1}, \textit{ARID1A}, and \textit{PPP2R1A}. Type II ovarian cancers are the most common of the ovarian cancer histotypes, consisting of high-grade serous (70%), high-grade endometrioid, carcinosarcoma, and undifferentiated carcinomas. Type II tumors are defined by \textit{TP53} mutations, which are rare in type I cancers.\textsuperscript{118-121} Each of these types have distinct risk factors and potential precursor lesions.\textsuperscript{121}

Risk factors for development of ovarian cancer include hormonal factors such as early menarche, late menopause, and nulliparity. The use of oral contraceptives reduces risk of ovarian carcinoma—this risk reduction persists for up to 30 years after cessation of use.\textsuperscript{122} Additionally, tubal ligation and hysterectomy decrease population level epithelial ovarian cancer risk. Genetic predisposition to breast or ovarian cancer is the most important risk factor for the development of ovarian cancer, and 18% to 24% of ovarian carcinomas may arise in conjunction with a hereditary predisposition.\textsuperscript{123-126} Germline genetic mutations are far more common among type II ovarian cancers, while endometriosis and hormonal factors predispose to type I ovarian malignancies.\textsuperscript{121,126,129}

Since 2007, the National Comprehensive Cancer Network guidelines began recommending that all women diagnosed with ovarian cancer receive genetic testing as up to 20% of ovarian cancer patients are \textit{BRCA1}/\textit{2} mutation carriers.\textsuperscript{127-130,134} Although family history of breast and/or epithelial ovarian cancer is one of the strongest factors for lifetime risk of having breast or epithelial ovarian cancer, up to 50% of women with ovarian cancer who test positive for a \textit{BRCA} mutation have no family history of either malignancy, supporting the importance of testing all women with a personal diagnosis of ovarian cancer, regardless of family history. The identification of deleterious mutations allows for cascade testing. Relatives of the affected patient are referred for genetic testing limited to the identified mutation. The lifetime risk for the development of ovarian cancer for carriers of mutations in the \textit{BRCA1} and \textit{BRCA2} genes
is estimated to be between 20% and 45% and 10% and 20%, respectively.\textsuperscript{123,130,135}

One of the challenges associated with early detection of ovarian cancer has historically been the lack of an identifiable precursor lesion. In 2001, however, “dysplastic changes” in the fallopian tubes removed from women with increased risk of developing ovarian carcinoma were first described.\textsuperscript{136} Subsequent careful microscopic examination using a newly developed “sectioning and extensively examining of the fimbriated end” protocol (SEE-FIM) of the grossly normal fallopian tubes and ovaries from women with \textit{BRCA1/2} mutations revealed occult tubal cancer and precancers designated as serous tubal intraepithelial carcinoma. The relationship between serous tubal intraepithelial carcinomas and high-grade serous and endometrioid cancers is supported by the ubiquitous presence of \textit{TP53} mutations and their typical location within the fimbriated end of the fallopian tube.\textsuperscript{118,121,137} High-grade, serous epithelial cancers of the ovary, fallopian tube, and peritoneum are now recognized to have a common fallopian tubal precursor lesion and often combined under the rubric of epithelial ovarian cancer (HGSOC).

For women at increased risk of ovarian cancer, the only confirmed prevention strategy is risk-reducing salpingo-oophorectomy.\textsuperscript{138,139} The lifetime risk of HGSOC is reduced to under 3\% with risk-reducing salpingo-oophorectomy. A modern understanding of the fallopian tube as the site of origin for many ovarian cancers has led to the suggestion that opportunistic salpingectomy could be implemented as a potential cancer prevention strategy in the general population. Scandinavian population-based cohort studies have demonstrated a significant decrease in epithelial ovarian cancer following salpingectomy.\textsuperscript{140,141} Opportunistic salpingectomy is feasible among women undergoing tubal ligation, hysterectomy, or other pelvic surgery.\textsuperscript{142}

\textbf{Early Staged Ovarian Cancer.} Early stage epithelial ovarian cancer has an excellent outcome. Low grade, stages IA and B disease can be cured in up to 90\% to 95\% of cases by a complete surgical procedure. The prevailing position in the United States is that such patients do not benefit from chemotherapy.\textsuperscript{143}

The standard of care for women with stages IC and II, and all women with grade 3 or clear cell histology, is adjuvant chemotherapy with 3 to 6 cycles of platinum- and taxane-based chemotherapy.\textsuperscript{144}

\textbf{Advanced Ovarian Cancer.} A pelvic mass with ascites, an omental cake, and an elevated CA-125 is pathognomonic for advanced ovarian cancer. CT scan is the imaging modality of choice to evaluate the upper abdomen and potential resectability of disease. Concerning physical or radiographic exam findings should prompt referral to a gynecologic oncologist (Table 41-10), as studies demonstrate inferior patient outcome for women who have had primary surgery by nongynecologic oncologists.

The objectives of surgery in ovarian cancer are threefold. The first is to make the histologic diagnosis. The second is to assess the extent of disease through complete surgical staging (Tables 41-11 and 41-12). When epithelial ovarian cancer is identified on frozen section and disease is grossly limited to the pelvis, complete staging with node dissection will upstage nearly one-third of patients.\textsuperscript{145} The third objective is (when feasible) surgical cytoreduction or debulking. The extent of disease upon entering the abdomen and the residual disease upon completion of the debulking surgery are independent prognostic variables for patient outcome. The Gynecologic Oncology Group has defined optimal residual disease as residual tumor ≤1 cm in the largest diameter. However, more contemporary data suggest that the most favorable survival outcomes are associated with complete cytoreduction to no gross residual disease.\textsuperscript{146} Decisions about the benefits and risks of radical debulking for individual presentations and diverse pathology depend on the age and medical stability of the patient, as well as the pathologic type of the cancer.

The publication of two randomized prospective trials of neoadjuvant chemotherapy (NACT) for ovarian cancer has led to a questioning of the dogma of maximum surgical effort. Both trials revealed no survival difference compared to primary debulking.\textsuperscript{147,148} In a patient who is medically compromised or in whom complete primary cytoreduction is unlikely, neoadjuvant...
Table 41-11

2014 International Federation of Gynecology and Obstetrics staging of epithelial ovarian cancer

<table>
<thead>
<tr>
<th>Stage</th>
<th>Description</th>
<th>Tumor stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Tumor confined to ovaries or fallopian tube(s)</td>
<td>T1</td>
</tr>
<tr>
<td>IA</td>
<td>Tumor limited to one ovary (capsule intact) or fallopian tube</td>
<td>T1a</td>
</tr>
<tr>
<td></td>
<td>No tumor on ovarian or fallopian tube surface</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No malignant cells in the ascites or peritoneal washings</td>
<td></td>
</tr>
<tr>
<td>IB</td>
<td>Tumor limited to both ovaries (capsules intact) or fallopian tubes</td>
<td>T1b</td>
</tr>
<tr>
<td></td>
<td>No tumor on ovarian or fallopian tube surface</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No malignant cells in the ascites or peritoneal washings</td>
<td></td>
</tr>
<tr>
<td>IC</td>
<td>Tumor limited to one or both ovaries or fallopian tubes, with any of the following:</td>
<td>T1c</td>
</tr>
<tr>
<td></td>
<td>IC1 Surgical spill intraoperatively</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IC2 Capsule ruptured before surgery or tumor on ovarian or fallopian tube surface</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IC3 Malignant cells present in the ascites or peritoneal washings</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>Tumor involves one or both ovaries or fallopian tubes with pelvic extension (below pelvic brim) or peritoneal cancer (Tp)</td>
<td>T2</td>
</tr>
<tr>
<td>IIA</td>
<td>Extension and/or implants on the uterus and/or fallopian tubes and/or ovaries</td>
<td>T2a</td>
</tr>
<tr>
<td>IIB</td>
<td>Extension to other pelvic intraperitoneal tissues</td>
<td>T2b</td>
</tr>
<tr>
<td>III</td>
<td>Tumor involves one or both ovaries, or fallopian tubes, or primary peritoneal cancer, with cytologically or histologically confirmed spread to the peritoneum outside the pelvis and/or metastasis to the retroperitoneal lymph nodes</td>
<td>T3</td>
</tr>
<tr>
<td>IIIA</td>
<td>Metastasis to the retroperitoneal lymph nodes with or without microscopic peritoneal involvement beyond the pelvis</td>
<td>T1, T2, T3aN1</td>
</tr>
<tr>
<td>IIIA1</td>
<td>Positive retroperitoneal lymph nodes only (cytologically or histologically proven)</td>
<td></td>
</tr>
<tr>
<td>IIIA1(i)</td>
<td>Metastasis ≤10 mm in greatest dimension (note this is tumor dimension and not lymph node dimension)</td>
<td>T3a/T3aN1</td>
</tr>
<tr>
<td>IIIA1(ii)</td>
<td>Metastasis &gt;10 mm in greatest dimension</td>
<td></td>
</tr>
<tr>
<td>IIIA2</td>
<td>Microscopic extrapelvic (above the pelvic brim) peritoneal involvement with or without positive retroperitoneal lymph nodes</td>
<td>T3a/T3aN1</td>
</tr>
<tr>
<td>IIIB</td>
<td>Macroscopic peritoneal metastases beyond the pelvic brim ≤2 cm in greatest dimension, with or without metastasis to the retroperitoneal lymph nodes</td>
<td>T3b/T3bN1</td>
</tr>
<tr>
<td>III C</td>
<td>Macroscopic peritoneal metastases beyond the pelvic brim &gt;2 cm in greatest dimension, with or without metastases to the retroperitoneal nodes (Note 1)</td>
<td>T3c/T3cN1</td>
</tr>
<tr>
<td>IV</td>
<td>Distant metastasis excluding peritoneal metastases</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stage IV A: Pleural effusion with positive cytology</td>
<td>Any T, any N, M1</td>
</tr>
<tr>
<td></td>
<td>Stage IV B: Metastases to extra-abdominal organs (including inguinal lymph nodes and lymph nodes outside of abdominal cavity) (Note 2)</td>
<td></td>
</tr>
</tbody>
</table>


Table 41-12

Components of comprehensive surgical staging and debulking of epithelial ovarian cancer

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertical abdominal incision</td>
<td>Adequate to visualize the diaphragms</td>
</tr>
<tr>
<td>Evacuation of ascites</td>
<td></td>
</tr>
<tr>
<td>Peritoneal washings</td>
<td>Of each pelvic gutter and diaphragm</td>
</tr>
<tr>
<td>En bloc hysterectomy and bilateral salpingo-oopherectomy</td>
<td></td>
</tr>
<tr>
<td>Infragastric omentectomy</td>
<td></td>
</tr>
<tr>
<td>Retroperitoneal and pelvic lymph node dissection</td>
<td></td>
</tr>
<tr>
<td>Examination of the entire bowel</td>
<td></td>
</tr>
<tr>
<td>Random biopsies</td>
<td>Of apparently uninvolved areas of peritoneum, pericolic gutters, diaphragm</td>
</tr>
</tbody>
</table>

Chemotherapy followed by interval debulking may be more appropriate and is supported by recent randomized controlled trials. Typically, treatment with NACT includes three cycles of platinum-based chemotherapy prior to open debulking, then three additional cycles after surgery. Diagnostic laparoscopic evaluation prior to cytoreductive surgery has been suggested as a means to avoid unnecessary laparotomy, resulting in suboptimal cytoreduction. Patients deemed not to be candidates for cytoreduction could proceed immediately to NACT at the time of tissue collection for definitive diagnosis. A Fagotti predictive index ≥8 (Table 41-13) is a predictor of suboptimal cytoreduction in advanced ovarian cancer with reasonable sensitivity and high specificity.149 These recommendations currently apply to HGSO, clear cell cancer, and high-grade endometrioid ovarian...
cancers. Low-grade tumors are less chemotherapy sensitive, and primary surgical resection is recommended when feasible. Standard of care adjuvant therapy of advanced stage epithelial ovarian cancer remains intravenous platinum- and taxane-based chemotherapy. In 2006, the National Cancer Institute issued a clinical alert indicating that combination intravenous/intraperitoneal platinum/taxane postoperative chemotherapy should be considered first line for women with optimally cytoreduced EOC. This was the result of completion and analysis of three independent randomized clinical trials showing a significant survival advantage for intraperitoneal therapy. Intraperitoneal (IP) therapy is administered via an implanted 9.6 French venous port catheter with the port placed over the right or left costal margin. The catheter is tunneled caudad with insertion through the fascia in the lower abdomen and the tip in the pelvis. The IP catheter may be placed at the time of surgical debulking via an open laparotomy approach or prior to initiating chemotherapy via a laparoscopic approach. In some centers, the IP catheter may be placed by interventional radiology with CT guidance.

Patients who have suboptimally debulked advanced stage disease and/or who are not candidates for intraperitoneal therapy should receive intravenous adjuvant chemotherapy. Interest has increased in both dose dense IV chemotherapy dosing as well as incorporation of biologic agents.

Secondary cytoreduction upon recurrence can be considered (Table 41-14). Patients who have had a disease-free

### Table 41-13

<table>
<thead>
<tr>
<th>LAPAROSCOPIC FEATURE</th>
<th>SCORE 0</th>
<th>SCORE 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peritoneal carcinomatosis</td>
<td>Carcinomatosis involving a limited area (along the paracolic gutter or the pelvic peritoneum) and surgically removable by peritonectomy</td>
<td>Unresectable massive peritoneal involvement as well as with a miliary pattern of distribution</td>
</tr>
<tr>
<td>Diaphragmatic disease</td>
<td>No infiltrating carcinomatosis and no nodules confluent with the most part of the diaphragmatic surface</td>
<td>Widespread infiltrating carcinomatosis or nodules confluent with the most part of the diaphragmatic surface</td>
</tr>
<tr>
<td>Mesenteric disease</td>
<td>No large infiltrating nodules and no involvement of the root of the mesentery as would be indicated by limited movement of the various intestinal segments</td>
<td>Large infiltrating nodules or involvement of the root of the mesentery indicated by limited movement of the various intestinal segments</td>
</tr>
<tr>
<td>Omental disease</td>
<td>No tumor diffusion observed along the omentum up to the large stomach curvature</td>
<td>Tumor diffusion observed along the omentum up to the large stomach curvature</td>
</tr>
<tr>
<td>Bowel infiltration</td>
<td>No bowel resection was assumed and no miliary carcinomatosis on the ansae observed</td>
<td>Bowel resection assumed or miliary carcinomatosis on the ansae observed</td>
</tr>
<tr>
<td>Stomach infiltration</td>
<td>No obvious neoplastic involvement of the gastric wall</td>
<td>Obvious neoplastic involvement of the gastric wall</td>
</tr>
<tr>
<td>Liver metastases</td>
<td>No surface lesions</td>
<td>Any surface lesion</td>
</tr>
</tbody>
</table>

### Table 41-14

<table>
<thead>
<tr>
<th>TIME FROM COMPLETION OF PRIMARY THERAPY</th>
<th>DEFINITION</th>
<th>INTERVENTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Progression on therapy</td>
<td>Platinum-refractory</td>
<td>No value of secondary debulking unless remediating complication such as bowel obstruction Non–platinum-based chemotherapy Consider clinical trial</td>
</tr>
<tr>
<td>Progression within 6 months of completion of primary therapy</td>
<td>Platinum-resistant</td>
<td>No value of secondary debulking unless remediating complication such as bowel obstruction Non–platinum-based chemotherapy consider adding bevacizumab Consider clinical trial</td>
</tr>
<tr>
<td>Progression after 6 months post completion of primary therapy</td>
<td>Platinum-sensitive</td>
<td>Consider secondary debulking if greater than 12 months interval Consider platinum +/- taxane +/- bevacizumab, +/- pegylated liposomal doxorubicin, +/- gemcitabine Consider maintenance PARP inhibitor Consider clinical trial</td>
</tr>
</tbody>
</table>
period of at least 12 months following an initial complete clinical response to surgery and initial chemotherapy, who have no evidence of carcinomatosis on imaging, and who have disease that can be completely resected are considered optimal candidates. A randomized controlled trial reported in abstract form demonstrated a benefit of secondary cytoreduction under strict entry criteria (DESKTOP3); the GOG-0213 study of secondary cytoreduction is maturing. Debunking surgery done after subsequent relapses or in women with early recurrence has not been shown to result in an outcome benefit and should be used only to palliate disease complications.

The most common cause of palliative surgery is bypass of bowel obstruction. The majority of women with advanced ovarian cancer will eventually develop and potentially die from malignant bowel obstruction. While management of these cases is controversial, in some cases surgical correction has been shown to prolong life and improve quality of life.153 Nonsurgical options include placement of a venting gastrostomy tube, performed endoscopically or surgically. Management of malignant bowel obstruction in women with recurrent advanced disease should be individualized.

Chemotherapy is the mainstay of therapy for recurrent EOC. Treatment approaches are based upon platinum sensitivity.154 Referral to an oncologist with specific expertise in chemotherapeutic treatment of ovarian cancer and access to clinical trials is important. In determining secondary and subsequent therapy, consideration of prior therapies, sites of disease, organs at risk from cancer, organs sustaining injury from prior therapy, and quality of life desires of patient should be taken into consideration.

**Ovarian Germ Cell Tumors.** Ovarian germ cell tumors occur most commonly in women under age 30. The most common benign germ cell neoplasm is the mature cystic teratoma; approximately 1% of teratomas contain a secondary malignancy arising from one of the components, most commonly squamous cell cancer and most commonly in postmenopausal women. Malignant germ cell tumors often grow and disseminate rapidly and are symptomatic. The rapid growth may be accompanied by torsion or rupture, producing an acute abdomen and the need for emergent intervention. Because they are derived from primordial germ cells, many produce characteristic tumor markers. Immature teratomas comprise a significant proportion of malignant germ cell tumors and may be associated with elevated lactate dehydrogenase (LDH) or α-fetoprotein (AFP). Excluding teratomas, the most common malignant germ cell tumor is dysgerminoma, made up of pure undifferentiated germ cells. Bilaterality occurs in up to 15% of patients; lactate dehydrogenase is commonly elevated, and elevated b-hCG may occur.

Less common malignant germ cell tumors include endodermal sinus or yolk sac tumors, embryonal carcinomas, mixed germ cell neoplasms, polyembryomas, and choriocarcinomas. Endodermal sinus tumors may have elevated AFP levels in the blood while embryonal and mixed germ cell tumors may have elevated β-hCG, LDH, or AFP. Tumor markers are useful to follow during surveillance and definitive therapy. Other than completely resected stage I, grade I immature teratoma, adjuvant chemotherapy with a platinum-containing regimen has been historically recommended.155 Because of the high response rates to chemotherapy and the long-term toxicity of treatment, a “watch and wait” approach with treatment only upon recurrence has been suggested as safe for selected, well-staged patients with germ cell tumors.156 The cure rate remains high, near 90% even when metastatic disease is present; recurrent disease is more difficult to eradicate.155

Fertility preservation is the standard surgical approach for ovarian germ cell tumors as disease tends to be diagnosed at stage I, and salvage chemotherapy is overall extremely successful. Staging should include removal of the involved ovary, biopsy of any suspicious areas, pelvic and para-aortic node dissection, and omentectomy. Hysterectomy or removal of the second ovary is rarely indicated.

Growing teratoma syndrome is a rare sequela of germ cell malignancies. Characteristically, during or after chemotherapy slow-growing tumors will increase in size and may even compress surrounding organs. Malignant transformation within these masses has been described. Treatment is with surgical resection.157

**Ovarian Sex Cord-Stromal Tumors.** Sex cord-stromal cell tumors, rare tumors, are derived from cells that support and surround the oocyte and can present with symptoms referable to endocrine activity of the tumor. These include granulosa cell tumors (female differentiated), fibroma-thecomas, and Sertoli-Leydig cell tumors (male differentiated). Granulosa cell tumors are the most common in this group and are a low-grade malignancy with fewer than 3% bilaterality. They are treated with conservative surgery, similar to germ cell tumors in young women.153 Hysterectomy and bilateral salpingo-oophorectomy is recommended for women who have completed childbearing. Nodal staging can be safely omitted in the absence of grossly involve nodes and fertility preservation is possible in disease limited to one ovary, the most common presentation. Debunking surgery is recommended for more extensive disease. These tumors and the thecomas in the same class often stimulate estrogen production and can be found in association with endometrial hyperplasia and cancer (5%). Granulosa cell tumors can recur over a prolonged period given their low rate of proliferation and tendency for local or intraperitoneal recurrence. Inhibin has been shown to be elaborated by these tumors and often is followed to identify recurrence of the disease. The Sertoli/Leydig cell tumors can present with virilization as a primary symptom. Evaluation of the ovary when this symptom is found is always of value.

**Gestational Trophoblastic Disease.** Gestational trophoblastic disease (GTD) is a spectrum of abnormal pregnancy–related trophoblastic proliferations. Premalignant histologic types include partial and complete hydatidiform moles. Primary surgery for diagnosis and initial therapy is a suction dilatation and curettage. Clinically, partial moles present as missed abortions and usually resolve with observation. Partial moles are triploid, usually XXX, which can result from dispermic fertilization of an egg. A previously described classical presentation of hyperemesis gravidarum, hyperthyroidism, preeclampsia, pulmonary trophoblastic embolization, and uterine size larger than dates is rarely seen today because of routine ultrasound assessments during early pregnancy. Even in the first trimester, however, a characteristic “snow storm” appearance may be seen on ultrasound. Pathologic examination will demonstrate no fetal tissue and have a diploid karyotype resulting from paternal duplication occurring after loss of maternal genetic material, or occasionally
with dispermic fertilization of an empty egg. Often associated theca lutein ovarian cysts, which can be greater than 6 cm in diameter, are seen on ultrasound. They should be followed without surgical intervention as they resolve with removal or treatment of the GTD. Following uterine evacuation, patients with molar pregnancies must be followed closely with weekly β-hCGs until normal for 3 weeks and then monthly for at least 6 months. Contraception should be provided to allow for surveillance. Any increase in β-hCG should trigger further evaluation and consideration of chemotherapy.

Invasive moles, choriocarcinoma, and placental site trophoblastic tumors are malignant disorders. Invasive moles are diagnosed following the diagnosis of a molar pregnancy if any of the following are demonstrated: (a) a plateau of β-hCG lasts for four measurements over a period of 3 weeks or longer; (b) a rise in β-hCG for three consecutive weekly measurements over at least a period of 2 weeks or more; or (c) β-hCG level remains elevated for 6 months or more. Metastatic GTD can present on the cervix, vagina, liver, lung, or brain and should not be managed surgically. In a woman of reproductive age, a diagnosis of metastatic GTN can be made without biopsy if a β-hCG is found to be elevated in the setting of widespread metastatic disease. In fact, given the incidence of bleeding complications biopsy is not recommended.

Chemotherapy is the primary recommended therapy. Per 2000 FIGO staging and classification, a risk score of 6 and below is classified as low risk and above 6 is considered high risk (Table 41-15). Low-risk patients are treated with single agent chemotherapy (methotrexate or actinomycin-D); high-risk patients receive multiagent chemotherapy. In either case, chemotherapy continues until β-hCG levels have normalized. Modern salvage and cure rates are high, with 5-year survival of high-risk patients reported as high as 90%. Twelve months of surveillance with contraception is recommended following treatment in order to allow complete surveillance for relapse.

Beyond dilation and curettage, surgery may have a role in the management of GTD. Hysterectomy is recommended for placental site trophoblastic tumors for which metastasis is rare. Laparotomy may be indicated in the cases of uncontrolled intra-abdominal or uterine bleeding. Neurosurgery may be required if there is intracranial bleeding or increased intracranial pressure due to metastatic disease.

**MINIMALLY INVASIVE GYNECOLOGIC SURGERY**

**Hysteroscopy**

See earlier section, “Hysteroscopy” under “Procedures Performed for Structural Causes of Abnormal Uterine Bleeding.”

**Laparoscopy**

The standard method for gynecologic laparoscopy follows the same methods as all minimally invasive surgery. In general, a camera port is placed near the umbilicus. Sometimes it must be placed more cephalad if the patient has a larger fibroid uterus. Two additional ports are placed laterally, usually just superior and medial to the anterior superior iliac spines. Single site laparoscopic procedures may improve cosmesis and reduce postoperative pain, but challenges including lack of triangulation and instrument crowding at the umbilicus make this technique challenging to apply to more complex procedures.

**Robotic Surgery**

Over the last decade, there has been increased use of robotics for gynecologic surgery. With the DaVinci robotic system, the surgeon sits at a console and visualizes the operative field with three-dimensional optics. The use of robotic surgery has been described for virtually every gynecologic procedure that has been performed abdominally or laparoscopically. The laparoscopic instruments are “wristed” and move as the surgeon’s hands/fingers move the actuators at the console. Robotic surgery

---

**Table 41-15**

**International Federation of Gynecology and Obstetrics/World Health Organization scoring system for gestational trophoblastic disease based on prognostic factors**

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td>&lt;40</td>
<td>&gt;40</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><strong>Antecedent pregnancy</strong></td>
<td>Mole</td>
<td>Abortion</td>
<td>Term</td>
<td></td>
</tr>
<tr>
<td><strong>Interval from index pregnancy, months</strong></td>
<td>&lt;4</td>
<td>4–6</td>
<td>7–12</td>
<td>&gt;12</td>
</tr>
<tr>
<td><strong>Pretreatment hCG mL/mL</strong></td>
<td>&lt;10³</td>
<td>&gt;10³–10⁴</td>
<td>&gt;10⁴–10⁵</td>
<td>&gt;10⁵</td>
</tr>
<tr>
<td><strong>Largest tumor size including uterus, cm</strong></td>
<td>–</td>
<td>3–4</td>
<td>≥5</td>
<td>–</td>
</tr>
<tr>
<td><strong>Site of metastases including uterus</strong></td>
<td>Lung</td>
<td>Spleen, kidney</td>
<td>Gastrointestinal tract</td>
<td>Brain, liver</td>
</tr>
<tr>
<td><strong>Number of metastases identified</strong></td>
<td>–</td>
<td>1–4</td>
<td>5–8</td>
<td>&gt;8</td>
</tr>
<tr>
<td><strong>Previous failed chemotherapy</strong></td>
<td>–</td>
<td>–</td>
<td>Single drug</td>
<td>Two or more drugs</td>
</tr>
</tbody>
</table>
uses a camera port, two to three robotic ports, and an accessory port. More meticulous dissection, improved visualization, and ability to operate with lower intra-abdominal pressures make the robotic platform advantageous, especially in obese patients. Longer set-up time and increased cost, however, are distinct disadvantages. The robotic unit costs up to $2.3 million and is associated with annual maintenance costs of $180,000 a year.\(^{162}\)

There is significant data to support robotic surgery in gynecologic malignancy; however, most procedures can be performed successfully with either robotic or laparoscopic platform depending on operator comfort and skill set. One large study suggested a lower conversion to laparotomy rate for robotic versus laparoscopic hysterectomy, but this was not statistically significant: conversion to laparotomy for laparoscopic hysterectomy was 9.9% compared with 4.9% for robotic cases (\(P = .06\)).\(^{163}\)

Complications Pertinent to Gynecologic Surgery

Abdominal Wall Vessels. The vessel at greatest risk of injury during the lateral trocar placement is the inferior epigastric artery. The superficial epigastric vessels and the superficial circumflex iliac vessels can be injured as well (Fig. 41-23). The primary methods to avoid vessel injury are knowledge of the vessels at risk and their visualization prior to trocar placement, when possible. The superficial vessels often can be seen and avoided by transillumination of the abdominal wall with the laparoscope. In contrast, the larger inferior epigastric vessels cannot be seen by transillumination because of their deeper location; these vessels often can be seen laparoscopically and avoided as they course along the peritoneum between the lateral umbilical fold of the bladder and the insertion of the round ligament into the inguinal canal. Anatomic variation and anastomoses between vessels make it impossible to know the exact location of all the abdominal wall vessels. For this reason, other strategies also should be used to avoid vessel injury, including the use of trocars with conical tips rather than pyramid tips and the use of the smallest trocars possible lateral to the midline.

Intestinal Injury. Another potentially serious complication of laparoscopic surgery is injury to either small or large intestines. An estimated incidence of bowel injury during laparoscopic gynecologic surgery is estimated to be 0.13%, 41% of which had a delayed diagnosis.\(^{164}\) Bowel injury can occur at the time of trocar insertion, especially if the patient has had previous abdominal procedures that often result in bowel adhesions to the anterior abdominal wall peritoneum, but rates appear similar regardless of entry technique. Due to the proximity of surgery to the bowel, thermal injury due to electrosurgery is also frequently implicated in intestinal injury. Time to diagnosis in these cases is typically several days postoperatively as a thermal injury takes time to mature and necrose.

Urologic Injuries. A risk of injury to the urogenital tract is inherent to gynecologic surgery due to proximity. Prevention of injury and intraoperative recognition and repair are crucial to avoiding long-term sequelae. Most urogenital fistulae are the result of unrecognized injuries to the urogenital tract at the time of surgery.

Bladder Injury. Placement of a Foley catheter prior to gynecologic surgery is critical to reducing risk of bladder injuries. Bladder injury during open or laparoscopic surgery results from retroperitoneal perforation during lower trocar placement or during sharp dissection of the bladder from the lower uterine segment during hysterectomy. The latter of these two situations is usually recognized intraoperatively; the first sign of the former may be postoperative hematuria, lower-port incisional drainage, or pneumoturia during laparoscopy. Once diagnosed, large defects require layered closure, whereas smaller defects usually close spontaneously within days or weeks with the aid of transurethral catheter drainage.

Ureteral Injury. Although ureteral injury is rare, occurring in less than 1% of gynecologic procedures, it is the most serious of the complications related to gynecologic surgery, particularly if unrecognized.\(^{165,166}\) There are three anatomic locations where the ureter is at risk during gynecologic procedures (see Fig. 41-5): (a) the ureter descends over the pelvic brim as it courses over the bifurcation of the common iliac artery into the external and internal iliac arteries just below the ovarian vessels; (b) in the pelvis, the ureter courses along the lateral aspect of the broad ligament to enter the base of the broad ligament; and (c) the ureter is found less than 2 cm lateral to the cervix, passing under the uterine artery and then medially over the anterior vaginal fornix before entering the trigone of the bladder—this is the most common location of ureteral injury. Ureteral injuries, including complete ligation, partial resection, or thermal injuries, usually will manifest within hours to days of surgery. Complete obstruction most often manifests as flank pain, whereas the first sign of partial or complete transection may be symptoms of intra-abdominal irritation caused by urine leakage. Transperitoneal thermal injuries resulting from fulguration of endometriosis may be similar to those after transection, but the appearance of symptoms may be delayed several days until tissue necrosis occurs.

Routine cystoscopy following hysterectomy is advocated by some gynecologists. For procedures performed for prolapse or incontinence where injury to the urinary tract is highest, routine cystoscopy is recommended. Consideration of a surgeon’s individual complication rate and the difficulty of an individual procedure are considerations for the provision of routine cystoscopy.\(^{166}\)

Vaginal Vault Dehiscence. This complication of hysterectomy seems to be more common in laparoscopic and robotic

\[\text{Figure 41-23. Location of anterior abdominal wall blood vessels.}\]
surgery. This may be due to the use of cautery in dividing the vaginal cuff or in the method of vaginal closure when done minimally invasively. Vaginal closure of the cuff appears to decrease the rate of vaginal cuff dehiscence in MIS hysterectomy.

Hemodynamically stable women without bowel evisceration may be candidates for transvaginal repair without abdominal exploration. Vaginal approach may also be appropriate in select cases of evisceration in which the bowel can be completely evaluated vaginally. Since bowel evisceration can lead to peritonitis and sepsis, all women with bowel eviscerations are considered to have a surgical emergency, and surgery should not be delayed for imaging. In most cases of bowel evisceration, evaluation of the bowel by laparoscopy or laparotomy is indicated to ensure bowel integrity.

REFERENCES

SPECIFIC CONSIDERATIONS

PART II


156. Mangili G, Sigismondo C, Lorusso D, et al. The role of staging and adjuvant chemotherapy in stage I malignant ovarian


OVERVIEW

Neurologic surgery provides the operative and nonoperative management (i.e., prevention, diagnosis, evaluation, treatment, critical care, and rehabilitation) of disorders of the central, peripheral, and autonomic nervous systems (ANSs). Such disorders include those of the brain, meninges, skull and skull base, and their blood supply, including surgical and endovascular treatment of disorders of the intracranial and extracranial vasculature supplying the brain and spinal cord; disorders of the pituitary gland; disorders of the spinal cord, meninges, and vertebral column, including those that may require treatment by fusion, instrumentation, or endovascular techniques; and disorders of the cranial and spinal nerves throughout their distribution.

An accurate history is the first step toward neurologic diagnosis. A history of trauma or of neurologic symptoms is of obvious interest, but general constitutional symptoms are also important. Neurologic disease may have systemic effects, while diseases of other systems may affect neurologic function. The patient’s general medical ability to withstand the physiologic stress of anesthesia and surgery should be understood. A detailed history from the patient and/or family, along with a reliable physical examination, will clarify these issues.

NEUROANATOMY

An understanding of neuroanatomy is the foundation of comprehensive neurologic examination and diagnosis. Salient features will be considered, from cephalad to caudal. The cerebral hemispheres (or telencephalon) consist of the cerebral cortex, underlying white matter, the basal ganglia, hippocampus, and amygdala. The cerebral cortex is the most recently evolved part of the nervous system. Its functions are mapped to discrete anatomic areas. The frontal areas are involved in executive function, decision making, and restraint of emotions. The motor strip, or precentral gyrus, is the most posterior component of the frontal lobes, and is arranged along a homunculus with the head inferior and lateral to the lower extremities superiorly and medially. The motor speech area (Broca’s area) lies in the left posterior inferior frontal lobe in almost all right-handed people and in up to 90% of left-handed people. The parietal lobe lies between the central sulcus anteriorly and the occipital lobe posteriorly. The postcentral gyrus is the sensory strip, also arranged along a homunculus. The rest of the parietal lobe is involved with awareness of one’s body in space and relative to the immediate environment, body orientation, and spatial relationships. The occipital lobes are most posterior. The visual cortex is arrayed along the apposing
medial surfaces of the occipital lobes. The left occipital lobe receives and integrates data from the left half of each retina. A left occipital lesion would therefore result in an inability to see objects right of center. The temporal lobes lie below the Sylvian fissures. The hippocampus, amygdala, and lower optic radiations (Meyer’s loops) are important components of the temporal lobe and are involved in memory, emotion, and vision, respectively. The receptive speech area (Wernicke’s area) typically is found in the area of the left posterior superior temporal lobe and inferior parietal lobe. The basal ganglia include the caudate, putamen, globus pallidus, subthalamic nucleus, substantia nigra, and nucleus accumbens. These structures are involved in the selection, activation and termination of movement, and facilitate learning of appropriate context-dependent motor behaviors.

Lying deep to the cerebral hemispheres is the diencephalon, which includes the thalamus and hypothalamus. The thalamus is a key processor and relay circuit for most motor and sensory information traveling to or from cortex. The hypothalamus regulates homeostasis via the autonomic and neuroendocrine systems.

The brain stem consists of the midbrain (mesencephalon), pons (metencephalon), and medulla (myelencephalon). Longitudinal fibers run through the brain stem, carrying motor and sensory information between the cerebral hemispheres and spinal cord. The corticospinal tract is the major motor tract, while the medial lemniscus and spinothalamic tracts are the major sensory tracts. The nuclei of cranial nerves III through XII are also located within the brain stem. These nerves relay the motor, sensory, and special sense functions of the eye, face, mouth, and throat.

The cerebellum arises from the dorsal aspect of the brain stem. It integrates somatosensory, vestibular, and motor information for coordination and timing of movement. Midline, or vermian, lesions lead to truncal ataxia. Lateral, or hemispheric, lesions lead to tremor and dyscoordination in the extremities.

The ventricular system is the cerebrospinal fluid (CSF)—containing contiguous space inside the brain, continuous with the subarachnoid space outside the brain. The paired lateral ventricles consist of temporal, occipital, and frontal horns, as well as the main body. CSF travels from each lateral ventricle through the foramina of Monroe to the third ventricle, located between the left and right thalami. CSF then drains through the cerebral aqueduct to the fourth ventricle within the brain stem. The foramen of Magendie (midline) and paired foramina of Luschka (lateral) drain to the subarachnoid space. The approximate CSF volume in an average adult is 150 mL, and the choroid plexus produces approximately 500 mL of CSF per day.

The spinal cord starts at the bottom of the medulla and extends caudally through the spinal canal to the first lumbar vertebra, approximately. Motor tracts (efferent pathways) continue from the brain stem down via the lateral and anterior corticospinal tracts to anterior horn cells, and then exit via ventral nerve roots. Sensory information (afferent pathways) enters via dorsal nerve roots, travels cranially via the dorsal columns ( proprioception and fine touch) or spinothalamic tract (pain and temperature), and into the brain stem. Paired nerves exit the spinal cord at each level. There are 31 pairs: 8 cervical, 12 thoracic, 5 lumbar, 5 sacral, and 1 coccygeal.

The dorsal and ventral nerve roots at each level fuse to form mixed motor-sensory spinal nerves and spread through the body to provide innervation to muscles and sensory organs. The C5–T1 spinal nerves intersect in the brachial plexus and divide to form the main nerve branches to the arm, including the median, ulnar, and radial nerves. The L2–S4 spinal nerves intersect in the lumbar sacral plexus, and divide to form the main nerve branches to the leg, including the femoral and sciatic nerves.

The principal motor tract of the spinal cord is the corticospinal tract. It is a two-neuron path, including an upper motor neuron and a lower motor neuron. The upper motor neuron cell body is located within the motor strip of the cerebral cortex. The axon travels through the internal capsule to the brain stem, decussates at the brain stem–spinal cord junction, and travels down the contralateral corticospinal tract to the lower motor

Key Points

1. Neurologic surgery specializes in primarily surgical management of central, peripheral, and autonomic nervous system disorders.
2. Although clinical examination is paramount, neurosurgical diagnosis and treatment are aided largely by a variety of modalities, such as magnetic resonance imaging and intracranial pressure monitoring.
3. The common treatment goals for traumatic brain and spinal injury are aimed at preventing secondary insults of hypoxia and hypotension.
4. Aneurysmal subarachnoid hemorrhage remains one of the most morbid and intensive neurosurgical diseases. Endovascular therapy is a growing technology that allows for safer securing of ruptured aneurysms.
5. Brain tumors can arise from primary or metastatic tissues. Treatment typically involves resection, followed by radiation and/or chemotherapy, depending on the type and grade of tumor.
6. Spinal instrumentation is used for surgical stabilization of many types of spinal instability, including traumatic, infectious, oncologic, and degenerative.
7. Infection of the nervous system is a serious and prevalent medical problem. Operative management is indicated for most conditions in which there is symptomatic compression of neural structures.
8. Functional neurosurgery via device implantation is a rapidly evolving discipline that has already become the standard of care in treating medically refractory Parkinson’s disease and essential tremor. A wider variety of deep brain stimulation targets will treat additional neuropsychiatric diseases.
9. Stereotactic radiosurgery is a powerful treatment option for intracranial disease, whether it is primary or adjunct. Gamma knife surgery can be used to treat tumors, vascular malformations, and cranial neuralgias.
The lower motor neuron axon then travels via peripheral nerves to its target muscle. Damage to upper motor neurons typically results in hyperreflexia and mild atrophy. Damage to lower motor neurons results in flaccidity and significant atrophy.

The two major sensory tracts are three-neuron pathways. Fine touch and proprioceptive signals enter the spinal cord via the dorsal root ganglia and then ascend ipsilaterally via the dorsal columns. Then they synapse and decussate in the lower medulla, travel up the contralateral medial lemniscus to make a second synapse in the thalamus, and then finally ascend to the sensory cortex. Pain and temperature fibers first synapse in the dorsal horn of the spinal cord at their entry level, decussate, and then travel up the contralateral spinothalamic tracts to the thalamus. The second synapse occurs in the thalamus, and the output axons ascend to the sensory cortex.

The aforementioned motor and sensory tracts together constitute the somatic nervous system. In addition to this system, the ANS is the other constituent of the nervous system. The ANS carries messages for homeostasis and visceral regulation from the central nervous system (CNS) to target structures such as arteries, veins, the heart, sweat glands, and the digestive tract. CNS control of the ANS arises particularly from the hypothalamus and the nucleus of the tractus solitarius. The ANS is divided into the sympathetic, parasympathetic, and enteric systems. The sympathetic system drives the “fight or flight” response, using epinephrine to increase heart rate, blood pressure, blood glucose, and temperature, as well as to dilate the pupils. It arises from the thoracolumbar spinal segments. The parasympathetic system promotes the “rest and digest” state and uses acetylcholine to maintain basal metabolic function under nonstressful conditions. Parasympathetic fibers arise from cranial nerves III, VII, IX, and X, and from the second to fourth sacral segments. The enteric nervous system controls the complex synchronization of the digestive tract, especially the pancreas, gallbladder, and small and large bowels. It can run autonomously but is regulated by the sympathetic and parasympathetic systems.

**NEUROLOGIC EXAMINATION**

The neurologic examination is divided into several components and generally is done from head to toe. First, one must assess mental status. A patient may be awake, lethargic (will follow commands and answer questions, but then returns to sleep), stuporous (difficult to arouse), or comatose (no purposeful response to voice or pain). Cranial nerves may be thoroughly tested in the awake patient, but pupil reactivity, eye movement, facial symmetry, and gag are the most relevant measures when mental status is impaired. Motor testing is based on maximal effort of major muscle groups in those able to follow commands, while assessing for amplitude and symmetry of movement to deep central pain may be all that is possible for stuporous patients. Table 42-1 details scoring for motor assessment tests. Characteristic motor reactions to pain in patients with depressed mental status include withdrawal from the central nervous system (CNS) to target structures. T1 sequences made before and after gadolinium administration are useful for detecting neoplastic and infectious processes. T2 sequences facilitate assessment of

<table>
<thead>
<tr>
<th>GRADE</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No muscle contraction</td>
</tr>
<tr>
<td>1</td>
<td>Visible muscle contraction without movement across the joint</td>
</tr>
<tr>
<td>2</td>
<td>Movement in the horizontal plane, unable to overcome gravity</td>
</tr>
<tr>
<td>3</td>
<td>Movement against gravity</td>
</tr>
<tr>
<td>4</td>
<td>Movement against some resistance</td>
</tr>
<tr>
<td>5</td>
<td>Normal strength</td>
</tr>
</tbody>
</table>

It is critical to document sensory patterns in spinal cord injury (SCI) patients. Muscle stretch reflexes should be examined. Often comparing left to right or upper extremity to lower extremity reflexes for symmetry is the most useful for localizing a lesion. Check for ankle-jerk clonus or up-going toes (Babinski’s test). Presence of either is pathologic and signifies upper motor neuron disease.

**Diagnostic Studies**

**Plain Films.** Plain X-rays of the skull may demonstrate fractures, osteolytic or osteoblastic lesions, radiolucent foreign bodies, or pneumocephaly (air in the head). Plain films of the cervical, thoracic, and lumbar spine are used to assess for evidence of bony trauma or soft tissue swelling suggesting fracture. Spinal deformities and osteolytic or osteoblastic pathologic processes also will be apparent. However, the use of plain films has decreased given the rapid availability and significantly increased detail of computed tomography (CT) scans. They are typically used for assessing alignment in patients with known fractures, for intraoperative localization, and postoperative assessment of spinal instrumentation.

**Computed Tomography.** The noncontrast CT scan of the head is an extremely useful diagnostic tool in the setting of new focal neurologic deficit, decreased mental status, or trauma. It is rapid and almost universally available in hospitals in the United States. Its sensitivity allows for the detection of acute hemorrhage. Fine-slice CT scanning of the spine is helpful for defining bony anatomy and pathology and is the method of choice for identifying fractures of the spine. By providing an assessment of spinal alignment, CT scans can provide an indirect assessment of ligamentous injury, for example, “Rule of Spence” for assessing transverse ligament injury during Jefferson fractures (see “Spine Trauma” section later in this chapter). Conventional contrast-enhanced CT scan will help show neoplastic or infectious processes. In the current era, contrast CT generally is used for those patients who cannot undergo magnetic resonance imaging (MRI) scanning due to pacemakers or metal in the orbits (see following section for discussion of CT angiography, venography, and perfusion).

**Magnetic Resonance Imaging.** Magnetic resonance imaging (MRI) provides excellent imaging of soft tissue structures in the head and spine. It is a complex and evolving science. Several of the most clinically useful MRI sequences are worth describing. T1 sequences made before and after gadolinium administration are useful for detecting neoplastic and infectious processes. T2 sequences facilitate assessment of
Table 42-2

The Glasgow Coma Scale score

<table>
<thead>
<tr>
<th>MOTOR RESPONSE</th>
<th>VERBAL RESPONSE</th>
<th>EYE-OPENING RESPONSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obey commands</td>
<td>6</td>
<td>Oriented</td>
</tr>
<tr>
<td>Localizes to pain</td>
<td>5</td>
<td>Confused</td>
</tr>
<tr>
<td>Withdraws from pain</td>
<td>4</td>
<td>Inappropriate words</td>
</tr>
<tr>
<td>Flexor posturing</td>
<td>3</td>
<td>Unintelligible sounds</td>
</tr>
<tr>
<td>Extensor posturing</td>
<td>2</td>
<td>No sounds</td>
</tr>
<tr>
<td>No movement</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

*aAdd the three scores to obtain the Glasgow Coma Scale (GCS) score, which can range from 3 to 15. Add “T” after the GCS if intubated and no verbal score is possible. For these patients, the GCS can range from 3T to 10T.*
noninvasive alternative for the initial screening assessment and follow-up of patients with suspected or known vascular lesions, as well as the evaluation of vasospasm. Similarly, fine-slice time-of-flight axial images can be reformatted in three dimensions to build MRI angiograms and MRI venograms. MRI angiograms can detect stenosis of the cervical carotid arteries or intracranial aneurysms >3 mm in diameter. MRI venograms can assess the dural venous sinuses for patency or thrombosis. Two-dimensional time of flight imaging performs vascular reconstructions purely based on flow and does not require gado
dium contrast administration.

**CT and MR Perfusion.** Perfusion scans have recently emerged as a method to a global assessment of the vascular integrity in the cerebral hemispheres, which is very important in the assessment of ischemic stroke (see “Stroke”). CT perfusion scans generate quantitative color maps that indicate various physiologic parameters such as cerebral blood flow (CBF), cerebral blood volume (CBV), and mean transit time (MTT) through quantitative analysis of rapidly acquired image sequences during intravenous contrast administration. Similar to perfusion CT, perfusion MRI can be used to generate quantitative color maps of relative cerebral CBV and MTT. These perfusion-based measures can be used along with diffusion-weighted imaging in the evaluation of ischemic stroke, particularly to identify an ischemic “penumbra” or tissue that is ischemic but not yet infarcted, and may be salvageable with intervention.3,4

**Angiography.** Transarterial catheter-based angiography remains the gold standard for evaluation of vascular pathology of the brain and spine. The current state of the art is biplane imaging to reduce dye load and facilitate interventional procedures. Digital subtraction technologies minimize bony interference in the resultant images. Bilateral carotid arteries and bilateral vertebral arteries may be injected and followed through arterial, capillary, and venous phases for a complete cerebral angiogram.

**Electroencephalography.** Electroencephalography (EEG) involves measuring weak electrical signals from the brain that are transmitted through the skull through electrodes that are applied to the scalp. The voltage fluctuations detected by EEG are thought to reflect summed membrane potentials from underlying brain tissue. Clinically, EEG is useful for detecting seizures, interictal markers of epileptogenic tissue, and widespread abnormalities in brain function, such as diffuse encephalopathy. EEG is also used in concert with electrical stimulation to detect sensory evoked potentials that can be useful for intraoperative mapping during cranial and spine surgery.

**Electromyography and Nerve Conduction Studies.** Electromyography and nerve conduction studies (EMG/NCS) are useful for assessing the function of peripheral nerves. EMG records muscle activity in response to a proximal stimulation of the motor nerve. NCS record the velocity and amplitude of the nerve action potential. EMG/NCS typically is performed approximately 3 to 4 weeks after an acute injury, as nerves distal to the injury continue to transmit electrical impulses normally until degeneration of the distal nerve progresses.

**Invasive Monitoring.** The most reliable monitor, always, is an alert patient with a reliable neurologic examination. If a reliable neurologic examination is not possible due to the presence of brain injury, sedatives, or paralytics, or if there is active and unstable intracranial pathology, invasive monitoring is required. There are several methods of monitoring intracranial physiology. The methods described in the following sections are bedside intensive care unit (ICU) procedures that allow for continuous monitoring. Both procedures involve making a small hole in the skull with a hand-held drill. They generally are placed in the right frontal region to minimize the neurologic impact of possible complications such as hemorrhage.

**External Ventricular Drain.** An external ventricular drain is also known as a ventriculostomy. A perforated plastic catheter is inserted into the frontal horn of the lateral ventricle. An uninterrupted fluid column through a rigid tube allows transduction of intracranial pressure (ICP). CSF also can be drained to reduce ICP or sampled for laboratory studies.

**Intraparenchymal Physiologic Monitoring.** Intraparenchymal monitors can be inserted into the brain through a threaded post locked securely into a burr hole, commonly referred to as a bolt. A bolt allows ICP monitoring with a fiber-optic pressure transducer, but it is smaller and less invasive than a ventriculostomy and may be associated with fewer complications, although the data do not clearly support this. Furthermore, a bolt can also be used to introduce probes to measure brain tissue oxygenation, brain temperature, and to perform microdialysis of parenchymal samples; however, the utility of these latter measures in clinical practice is still under investigation. Patients with severe brain injury due to trauma or aneurysmal hemorrhage may benefit from placement of these sensors in addition to a ventriculostomy to drain CSF for control of ICP. Such monitoring requires two twist-drill holes, which may be placed on adjacent or opposite sides of the head.

**NEUROLOGIC AND NEUROSURGICAL EMERGENCIES**

**Raised Intracranial Pressure**

ICP normally varies between 4 and 14 mmHg. Sustained ICP levels above 20 mmHg can injure the brain. The Monro-Kellie doctrine states that the cranial vault is a rigid structure, and therefore, the total volume of the contents determines ICP. The three normal contents of the cranial vault are brain tissue, blood, and CSF. The brain’s contents can expand due to swelling from traumatic brain injury (TBI), stroke, or reactive edema. Blood volume can increase by extravasation to form a hematoma, or by reactive vasodilation in a hypoventilating, hypercarbic patient. CSF volume increases in the setting of hydrocephalus. Figure 42-2 demonstrates the classic CT findings of hydrocephalus. The addition of a lesion, such as a tumor or abscess, also will increase ICP. The pressure-volume curve depicted in Fig. 42-3 demonstrates a compensated region with a small ΔP/ΔV, and an uncompensated region with large ΔP/ΔV.

In the compensated region, increased volume is offset by decreased volume of CSF and blood.

Increased ICP can injure the brain in several ways. Focal mass lesions cause shift and herniation. Temporal lesions push the uncus medially and compress the midbrain. This phenomenon is known as uncal herniation. The posterior cerebral artery (PCA) passes between the uncus and midbrain and may be occluded, leading to an occipital infarct. Masses higher up in the hemisphere can push the cingulate gyrus under the falx cerebri. This process is known as subfalcine herniation. The anterior cerebral artery (ACA) branches run along the medial surface of the cingulate gyrus and may be occluded in this case, leading to medial frontal and parietal infarcts. Diffuse increases in pressure
in the cerebral hemispheres can lead to central, or transtentorial, herniation. Increased pressure in the posterior fossa can lead to upward central herniation or downward tonsillar herniation through the foramen magnum. Uncal, transtentorial, and tonsillar herniation can cause direct damage to the brain stem. Figure 42-4 diagrams patterns of herniation.

Patients with increased ICP, or intracranial hypertension, often will present with headache, nausea, vomiting, and progressive mental status decline. Cushing’s triad is the classic presentation of hypertension, bradycardia, and irregular respirations. Focal neurologic deficits such as hemiparesis may be present if there is a focal mass lesion causing the problem. Patients with these symptoms should undergo an immediate head CT and rapid neurosurgical evaluation.

Initial management of intracranial hypertension includes airway protection and adequate ventilation. A bolus of mannitol up to 1 g/kg causes free water diuresis, increased serum osmolality, and extraction of water from the brain. The effect is delayed by about 20 minutes and has a transient benefit. Driving serum osmolality above 300 mOsm/L is of indeterminate benefit and can have deleterious cardiovascular side effects, such as hypovolemia that leads to hypotension and decreased brain perfusion. A ventriculostomy and/or craniectomy may be needed for definitive decompression.

It is critical to note that lethargic or obtunded patients often have decreased respiratory drive. This causes the partial pressure of arterial carbon dioxide (Paco₂) to increase, resulting in cerebral vasodilation and worsening of intracranial hypertension. This cycle causes a characteristic “crashing patient,” who rapidly loses airway protection, becomes apneic, and herniates. Emergent intubation and ventilation to reduce Paco₂ to roughly 35 mmHg can reverse this process.

Figure 42-2. Head computed tomography scan demonstrating hydrocephalus. The third ventricle (3rd) is widened and rounded, the anterior horns of the lateral ventricles are plump, and pressure-driven flow of cerebrospinal fluid into brain parenchyma adjacent to the ventricles is seen (arrowhead). This is known as transependymal flow of cerebrospinal fluid.

Figure 42-3. Pressure-volume curve demonstrating the effect of changing the volume of intracranial contents on intracranial pressure. Note the compensated zone, with little change of pressure with change of volume, and the uncompensated zone, with significant change of pressure with change of volume. (Adapted with permission from Ellenbogen RG, Abdulrauf SI, Sekhar LN: Principles of Neurosurgery, 3rd ed. Philadelphia, PA: Elsevier/Saunders; 2012.)
Brain Stem Compression

The posterior fossa (brain stem and cerebellum) requires special consideration because the volume of the posterior fossa within the cranial vault is small. Posterior fossa lesions such as tumors, hemorrhage, or stroke can cause mass effect that can rapidly kill the patient in two ways. Occlusion of the fourth ventricle can lead to acute obstructive hydrocephalus, raised ICP, herniation, and eventually death. This mass effect can also lead directly to brain stem compression (Fig. 42-5). Symptoms of brain stem compression include hypertension, agitation, and progressive obtundation, followed rapidly by brain death. A patient exhibiting any of these symptoms needs an emergent neurosurgical evaluation for possible ventriculostomy or suboccipital craniectomy (removal of the bone covering the cerebellum). This situation is especially critical, as expeditious decompression can lead to significant functional recovery.

Stroke

Patients presenting with acute focal neurologic deficits at a clearly defined time of onset (i.e., when the patient was last seen in a normal state of health) must be evaluated as rapidly as possible. An emergent head CT scan should be done. The study is often normal because CT changes from ischemic stroke may take up to 24 hours to appear (Fig. 42-6). A patient with a clinical diagnosis of acute stroke <4.5 hours old, without hemorrhage on CT, may be a candidate for thrombolytic therapy with tissue plasminogen activator (tPA). When a proximal large-vessel obstruction is suspected, patients should be evaluated for endovascular mechanical thrombectomy if therapy can be initiated within 6 to 8 hours of symptom onset. Intravenous tPA should be given regardless, but a noninvasive intracranial vascular study such as CT angiography should also be obtained in these cases. An emergent MRI is helpful but not always diagnostically necessary.

Seizure

A seizure is defined as an uncontrolled synchronous organization of neuronal electrical activity. A new-onset seizure often signifies an irritative mass lesion in the brain, particularly in adults, in whom tumors commonly present with seizure. Patients with traumatic intracranial hemorrhage are at risk for seizure. In addition to airway and ventilatory problems, a seizing patient is also at risk for neural excitotoxicity if the activity is prolonged, such as in status epilepticus. Any patient with a new-onset seizure should have imaging of the brain after the seizure is controlled and the patient is resuscitated.

TRAUMA

Trauma is the leading cause of death in children and young adults; however, the incidence of death and disability from trauma has been slowly decreasing. This decline is partly attributable to increased awareness of safety devices such as seat belts and motorist helmets. Nonetheless, trauma remains a major cause of morbidity and mortality, and it can affect every major organ system in the body. The three main areas of neurosurgical focus are: traumatic brain injury (TBI), spinal cord injury (SCI), and peripheral nerve injury.

Head Trauma

Glasgow Coma Scale Score. The initial assessment of the trauma patient includes the primary survey, resuscitation,
secondary survey, and definitive care. Neurosurgical evaluation begins during the primary survey with the determination of the GCS score (usually referred to simply as the GCS) for the patient. The GCS is determined by adding the scores of the best responses of the patient in each of three categories. The motor score ranges from 1 to 6, verbal from 1 to 5, and eyes from 1 to 4. The GCS therefore ranges from 3 to 15, as detailed in Table 42-2. Tracheal intubation or severe facial or eye swelling can impede verbal and eye responses. In these circumstances, the patient is given the score of 1 with a modifier, such as verbal “1T” where T = tube.

**Scalp Injury.** Blunt or penetrating trauma to the head can cause injury to the densely vascularized scalp, and significant blood loss can result. Direct pressure initially controls the bleeding, allowing close inspection of the injury. If a simple laceration is found, it should be copiously irrigated and closed primarily. If the laceration is short, a single-layer, percutaneous suture closure will suffice. If the laceration is long or has multiple arms, the patient may need debridement and closure in the operating room, with its superior lighting and wider selection of instruments and suture materials. Careful reapproximation of the galea will provide a more secure closure and better hemostasis.
Blunt trauma also can cause crush injury with subsequent tissue necrosis. These wounds require debridement and consideration of advancement flaps to cover the defect.

**Skull Fractures.** The usual classification system for bony fractures may be applied to the skull. The fracture may be characterized by skull X-rays or head CT. A closed fracture is covered by intact skin. An open, or compound, fracture is associated with disrupted overlying skin. The fracture lines may be single (linear); multiple and radiating from a point (stellate); or multiple, creating fragments of bone (comminuted). Closed skull fractures do not normally require specific treatment. Open fractures require repair of the scalp and operative debridement. Indications for craniotomy include depression greater than the cranial thickness, intracranial hematoma, and frontal sinus involvement. Skull fractures generally indicate that a significant amount of force was transmitted to the head and should increase the suspicion for intracranial injury. Fractures that cross meningeal arteries can cause rupture of the underlying vessels and subsequent epidural hematoma (EDH) formation.

Depressed skull fractures may result from a focal injury of significant force. The inner and outer cortices of the skull are disrupted, and a fragment of bone is pressed in toward the brain in relation to adjacent intact skull. The fragment may overlap the edge of intact bone, or it may plunge completely below the level of adjacent normal skull. The inner cortex of the bone fragments often has multiple sharp edges that can lacerate dura, brain, and vessels. Craniotomy is required to elevate the fracture, repair dural disruption, and obtain hemostasis in these cases (Fig. 42-7). However, fractures overlying dural venous sinuses require restraint. Surgical exploration can lead to life-threatening hemorrhage from the lacerated sinus.

Fractures of the skull base are common in head-injured patients, and they indicate significant impact. They are generally apparent on routine head CT, but they should be evaluated with dedicated fine-slice coronal-section CT scan to document and delineate the extent of the fracture and involved structures. If asymptomatic, they require no treatment. Skull base fractures requiring intervention include those with an associated cranial nerve deficit or CSF leak. A fracture of the temporal bone, for instance, can damage the facial or vestibulocochlear nerve, resulting in vertigo, ipsilateral deafness, or facial paralysis. A communication may be formed between the subarachnoid space and the middle ear, allowing CSF drainage into the pharynx via the Eustachian tube or from the ear (otorrhea). Extravasation of blood results in ecchymosis behind the ear, known as Battle’s sign. A fracture of the anterior skull base can result in anosmia (loss of smell from damage to the olfactory nerve), CSF drainage from the nose (rhinorrhea), or periorbital ecchymosis, known as raccoon eyes.

Copious clear drainage from the nose or ear makes the diagnosis of CSF leakage obvious. Often, however, the drainage may be discolored with blood or small in volume if some drains into the throat. In indeterminate cases, it is important to consider radiographic findings on the CT scan near the fracture that suggest CSF leak, such as pneumocephalus, subarachnoid, or intraparenchymal blood at the fracture site. The “halo” test assesses for a double ring when a drop of the fluid is allowed to fall on an absorbent surface, but it has been shown to have poor clinical utility. The fluid can be sent for β-2 transferrin testing, a carbohydrate-free isoform of transferrin exclusively found in the CSF; however, these tests often take 1 to 2 weeks to result and also can be difficult to incorporate into clinical practice.

![Figure 42-7](image-url)  
**Figure 42-7.** A. Bone-window axial head computed tomography (CT) of a patient who presented aphasic after being struck with the bottom of a beer bottle. CT demonstrates a depressed skull fracture in the left posterior temporoparietal area. B. Brain-window axial head CT demonstrating intraparenchymal hematoma caused by laceration of cortical vessels by the edge of the fractured bone. Arrowhead indicates traumatic subarachnoid hemorrhage in the sylvanian fissure.
Many CSF leaks will heal with elevation of the head of the bed for several days. An elevation of the head of the bed reduces the hydrostatic pressure of the CSF fluid column in the cranial vault, near the site of the defect. As such, when the CSF leak is in the lumbar thecal sac, the head of the bed should be flat so as to maximize hydrostatic pressure of the CSF fluid column at the cranial vault, away from the site of the defect. In addition, lumbar drain can be used to reduce CSF pressure. When there is a contraindication, to lumbar drain placement (such as an intracranial mass lesion or hematoma), an extraventricular drain should be used for CSF diversion. Although persistent CSF leaks have been shown to increase the risk of meningitis, there is no evidence supporting the use of prophylactic antibiotic use for preventing meningitis in patients with CSF leaks.

Traumatic cranial neuropathies generally can be managed conservatively, with documentation of the extent of impairment and signs of recovery. Patients with traumatic facial nerve palsies may benefit from a course of steroids, although their benefit is unproven. Patients with facial nerve palsy of abrupt onset, who do not respond to steroids within 48 to 72 hours, may be considered for surgical decompression of the petrous portion of the facial nerve. Patients also may present with delayed-onset facial nerve palsy. Again, steroids are used and surgery can be considered, with mixed results.

Closed Head Injury. Closed head injury (CHI) is the most common type of TBI and a significant cause of morbidity and mortality in the United States. There are two important factors that affect the outcome of CHI in general. The initial impact causes the primary injury, defined as the immediate injury to neurons from transmission of the force of impact. The long, delicate axons of the neurons can shear as they undergo differential acceleration or deceleration along their projecting pathways. Prevention strategies, such as wearing helmets, remain the best means to decrease disability from primary injury. Subsequent neuronal damage due to the sequelae of trauma is referred to as secondary injury. Hypoxia, hypotension, hydrocephalus, intracranial hypertension, thrombosis, and intracranial hemorrhage may all be mechanisms of secondary injury.

One focus of basic research in TBI, critical care medicine, and neurosurgical intervention is to decrease the effects of secondary injury.

The Brain Trauma Foundation’s most recent summary of management recommendations for TBI patients was published in 2016 and is endorsed by the American Association of Neurological Surgeons, Congress of Neurological Surgeons, and the World Health Organization. The guidelines standardize the care of these patients with the hope of improving outcomes. Level I recommendations are based on a body of high-quality evidence, such as large, well-received randomized controlled trials. Level II and III recommendations are based on moderate and low quality evidence, respectively. Some of the common patterns of CHI, including concussion, contusion, and diffuse axonal injury, are discussed in “Types of Closed Head Injury.”

Initial Assessment The initial evaluation of a trauma patient remains the same whether or not the primary surveyor suspects head injury. The first three elements of the ABCDs of resuscitation—airway, breathing, and circulation—must be assessed and stabilized. Hypoxia and hypotension are known to worsen outcome in TBI (due to secondary injury), making cardiopulmonary stabilization critical. Patients who cannot follow commands require intubation for airway protection and ventilatory control. The fourth element, assessment of “D,” for disability, is undertaken next. Motor activity, speech, and eye opening can be assessed in a few seconds and a GCS score assigned.

The following is an example of how a primary surveyor may efficiently assess disability and GCS: Approach the patient and enter his or her field of view. Observe whether the patient is visually attentive. Clearly command: “Tell me your name.” Then ask the patient to lift up two fingers on each side sequentially, and wiggle the toes. A visually or verbally unresponsive patient should be assessed for response to peripheral stimuli such as nail-bed pressure, or deep central painful stimulation, such as a firm, twisting pinch of the sensitive supraclavicular skin. Watch for eye opening and movement of the extremities, whether purposeful or reflexive. Assess the verbal response. The motor, verbal, and eye-opening scores may be correctly assigned using this rapid examination. An initial assessment of the probability of significant head injury can be made, assuming that pharmacologic and toxic elements have not obscured the examination. The surveyor must also take note of any external signs of head injury, including bleeding from the scalp, nose, or ear, or deformation of the skull or face.

Classification TBI can be classified as mild, moderate, or severe. For patients with a history of head trauma, classification is as follows: severe head injury if the GCS score is 3 to 8, moderate head injury if the GCS score is 9 to 12, and mild head injury if the GCS score is 13 to 15. Many patients present to emergency rooms and trauma bays with a history of TBI. A triage system must be used to maximize resource utilization while minimizing the chance of missing occult or progressing injuries.

TBI patients who are asymptomatic, who have only headache, dizziness, or scalp lacerations, and who did not lose consciousness, have a low risk for intracranial injury and may be discharged home without a head CT scan. Head-injured patients who are discharged should be sent home with reliable family or friends who can observe the patient for the first post-injury day. Printed discharge instructions, which describe monitoring for confusion, persistent nausea, weakness, or speech difficulty, should be provided to the caretaker. The patient should return to the emergency department for evaluation of such symptoms.

Patients with a history of altered consciousness, amnesia, progressive headache, skull or facial fracture, vomiting, or seizure have a moderate risk for intracranial injury and should undergo a prompt head CT. If the CT is normal, and the neurologic examination has returned to baseline (excluding amnesia of the event), then the patient can be discharged to the care of a responsible adult, again with printed criteria for returning to the emergency room. Otherwise the patient must be admitted for a 24-hour observation period.

Patients with depressed consciousness, focal neurologic deficits, penetrating injury, depressed skull fracture, or changing neurologic examination have a high risk for intracranial injury. These patients should undergo immediate head CT and admission for observation or intervention as needed.

Types of Closed Head Injury

Concussion A concussion is defined as temporary neuronal dysfunction following nonpenetrating head trauma. The head CT is normal, and deficits resolve over minutes to hours. Definitions vary; some require transient loss of consciousness, while others include patients with any alteration of mental status. Memory difficulties, especially amnesia of the event, are very
common. Concussions may be graded. One method is the Colorado grading system. Head trauma patients with confusion only are grade 1, patients with amnesia are grade 2, and patients who lose consciousness are grade 3. Studies have shown that the brain remains in a hypermetabolic state for up to a week after injury. The brain is also much more susceptible to injury from even minor head trauma in the first 1 to 2 weeks after concussion. This is known as second-impact syndrome, and patients should be informed that, even after mild head injury, they might experience memory difficulties or persistent headaches. Return to play guidelines after sports-related concussions are controversial and are under active debate.

**Contusion** A contusion is a bruise of the brain, and occurs when the force from trauma is sufficient to cause breakdown of small vessels and extravasation of blood into the brain. The contused areas appear bright on CT scan, as seen in Fig. 42-8. The frontal, occipital, and temporal poles are most often involved. The brain sustains injury as it collides with rough, bony surfaces. Contusions themselves rarely cause significant mass effect as they represent small amounts of blood in injured parenchyma rather than coherent blood clots. Edema may develop around a contusion, causing mass effect. Contusions may enlarge or progress to frank hematoma, particularly during the first 24 hours. Contusions also may occur in brain tissue opposite the site of impact. This is known as a contre-coup injury. These contusions result from deceleration of the brain against the skull.

**Diffuse Axonal Injury** Diffuse axonal injury (DAI) is caused by damage to axons throughout the brain, due to rotational acceleration and then deceleration. Axons may be completely disrupted and then retract, forming axon balls. Small hemorrhages can be seen in more severe cases, especially on MRI.

Hemorrhage is classically seen in the corpus callosum and the dorsolateral midbrain. DAI can be considered to be a severe form of a concussion, often with irreversible consequence. It can often explain a poor neurological examination (such as impaired arousal) in cases without clear radiographic signs of global brain injury, particularly when there is damage in structures, such as the pontine reticular activating system or bilateral thalami, that are necessary for arousal. In these cases, alternative explanations of poor arousal, such as a basilar thrombus, must also be investigated.

**Penetrating Injury** These injuries are complex and must be evaluated individually. The two main subtypes are missile (e.g., due to bullets or fragmentation devices) and nonmissile (e.g., due to knives or ice picks). Some general principles apply. If available, skull X-rays and CT scans are useful in assessing the nature of the injury. Cerebral angiography must be considered if the object passes near a major artery or dural venous sinus. Operative exploration is necessary to remove any object extending out of the cranium, as well as for debridement, irrigation, hemostasis, and definitive closure. Small objects contained within brain parenchyma are often left in place to avoid iatrogenic secondary brain injury. High-velocity missile injuries (from high-powered hunting rifles or military weapons) are especially deadly, because the associated shock wave causes cavitory tissue destruction of an area that is much larger than the projectile itself. Projectiles that penetrate both hemispheres or traverse the ventricles are almost universally fatal. Antibiotics are given to decrease the chances of meningitis or abscess formation; however, the evidence supporting the use of antibiotics following missile injury is weak and largely comes from retrospective case studies and expert opinion. Recent guidelines published in regard to preventing combat-related infections recommend antimicrobial therapy for 5 days or until resolution of the associated CSF leak, albeit with limited supporting evidence.

**Traumatic Intracranial Hematomas.** The various traumatic intracranial hematomas contribute to death and disability secondary to head injury. Hematomas can expand rapidly and cause brain shift and subsequent herniation. Emergent neurosurgical evaluation and intervention often are necessary.

**Epidural Hematoma** EDH is the accumulation of blood between the skull and the dura. EDH usually results from arterial disruption, especially of the middle meningeal artery. The dura is adherent to bone, and some pressure is required to dissect between the two. On head CT, the blood clot is bright, biconvex in shape (lentiform), and has a well-defined border that usually respects cranial suture lines. An EDH is typically found over the convexities but may rarely occur in the posterior fossa as well. EDH has a classic, three-stage clinical presentation that is probably seen in only 20% of cases. The patient is initially unconscious from the concussive aspect of the head trauma. The patient then awakens and has a “lucid interval,” while the hematoma subclinically expands. As the volume of the hematoma grows, the decompensated region of the pressure-volume curve is reached, ICP increases, and the patient rapidly becomes lethargic and herniates. Uncal herniation from an EDH classically causes ipsilateral third nerve palsy and contralateral hemiparesis.

Open craniectomy for evacuation of the congealed clot and hemostasis generally is indicated for EDH. In some cases, EDH can be caused from bony venous bleeding that is self-limited.

---

Figure 42-8. Severe bilateral contusions in the basal aspect of the frontal lobes, caused by the brain moving over the rough, irregular skull base during sudden cranial acceleration.
and may not require surgical intervention. Generally, patients who meet all of the following criteria may be managed conservatively: clot volume <30 cm\(^3\), maximum thickness <1.5 cm, and GCS score >8.\(^9\) Prognosis after successful evacuation is better for EDH than subdural hematoma (SDH). EDHs are associated with lower-energy trauma with less resultant primary brain injury. Good outcomes may be seen in 85% to 90% of patients, with rapid CT scan and intervention.\(^1\) In some cases, EDH can also be caused by dural venous sinus tears that rapidly expand and are typically associated with a high degree of morbidity when treated surgically.

**Acute Subdural Hematoma** An acute SDH is the result of an accumulation of blood between the arachnoid membrane and the dura. Acute SDH usually results from venous bleeding, typically from tearing of a bridging vein running from the cerebral cortex to the dural sinuses. The bridging veins are subject to stretching and tearing during acceleration/deceleration of the head because the brain shifts in relation to the dura, which firmly adheres to the skull. Elderly and alcoholic patients are at higher risk for acute SDH formation after head trauma due to brain atrophy.

On head CT scan, the clot is bright or mixed-density, crescent-shaped (lunate), may have a less distinct border, and does not cross the midline due to the presence of the falk. Most SDHs occur over the cerebral hemispheres, but they may also occur between the hemispheres or layer over the tentorium.

Open craniotomy for evacuation of acute SDH is indicated for any of the following: thickness >1 cm, midline shift >5 mm, or GCS drop by two or more points from the time of injury to hospitalization. Nonoperatively managed hematomas may stabilize and eventually reabsorb, or evolve into chronic SDHs.\(^1\) This management requires frequent neurologic examinations until the clot stabilizes based on serial head CT scans.

The prognosis for functional recovery is significantly worse for acute SDH than EDH because it is associated with greater primary injury to brain parenchyma from high-energy impacts. Prompt recognition and intervention minimizes secondary injury. The elderly patients with low admission GCS, or GCS drop by two or more points from the time of injury to hospitalization. Nonoperatively managed hematomas may stabilize and eventually reabsorb, or evolve into chronic SDHs.\(^1\) This management requires frequent neurologic examinations until the clot stabilizes based on serial head CT scans.

The prognosis for functional recovery is significantly worse for acute SDH than EDH because it is associated with greater primary injury to brain parenchyma from high-energy impacts. Prompt recognition and intervention minimizes secondary injury. The elderly patients with low admission GCS, or GCS drop by two or more points from the time of injury to hospitalization. Nonoperatively managed hematomas may stabilize and eventually reabsorb, or evolve into chronic SDHs.\(^1\) This management requires frequent neurologic examinations until the clot stabilizes based on serial head CT scans.

The prognosis for functional recovery is significantly worse for acute SDH than EDH because it is associated with greater primary injury to brain parenchyma from high-energy impacts. Prompt recognition and intervention minimizes secondary injury. The elderly patients with low admission GCS, or GCS drop by two or more points from the time of injury to hospitalization. Nonoperatively managed hematomas may stabilize and eventually reabsorb, or evolve into chronic SDHs.\(^1\) This management requires frequent neurologic examinations until the clot stabilizes based on serial head CT scans.

**Chronic Subdural Hematoma** Chronic SDH is a collection of blood breakdown products that is at least 2 to 3 weeks old. Acute hematomas are bright white (hyperdense) on CT scan for approximately 3 days, after which they fade to isodensity with brain, and then to hypodensity after 2 to 3 weeks. A true chronic SDH will be nearly as dark as CSF on CT. Traces of white blood are often seen due to small, recurrent hemorrhages into the collection. These small bleeds may expand the collection enough to make it symptomatic. This phenomenon is referred to as an acute-on-chronic SDH. Figure 42-9 demonstrates the CT appearance of an acute-on-chronic SDH. Vascularized membranes form within the hematoma as it matures. These membranes may be the source of acute hemorrhage.

Chronic SDHs often occur in patients without a clear history of head trauma as they may arise from minor head injury. Alcoholics, the elderly, and patients on anticoagulation are at higher risk for developing chronic SDH. Patients may present with headache, seizure, confusion, contralateral hemiparesis, or coma.

A chronic SDH >1 cm or any symptomatic SDH should be surgically drained. Unlike acute SDH, which consists of a thick, congealed clot, chronic SDH typically consists of a viscous fluid with the texture and dark brown color reminiscent of motor oil. A simple burr hole can effectively drain most chronic SDHs. However, the optimal treatment of chronic SDH remains controversial.\(^19\) Recent data suggest that open craniotomy is effective at reducing recurrence, but may be associated with more short-term complications.\(^20\) Most authorities agree that burr hole drainage should be attempted first to obviate the risks of formal craniotomy.\(^21\) A single burr hole placed over the dependent edge of the collection can be made, and the space is copiously irrigated until the fluid is clear. A second, more anterior burr hole can then be placed if the collection does not drain satisfactorily due to containment by membranes. The procedure is converted to open craniotomy if the SDH is too congealed for irrigation drainage, the complex of membranes prevents effective drainage, or persistent hemorrhage occurs that cannot be reached with bipolar cautery through the burr hole. The required surgical prepping and draping are always performed to allow simple conversion to craniotomy, and the scalp incision and burr holes are placed to allow easy incorporation into larger skin flaps.

There are various strategies to prevent reaccumulation of blood. Subdural or subgaleal drains may be left in place for 1 to 2 days. Subdural drains have been shown to reduce the risk of recurrence, whereas corticosteroid use in this patient population has been associated with higher morbidity without benefit.\(^20\) Mild hydration and bedrest with the head of the bed flat may

![Figure 42-9. Head computed tomography scan of an elderly patient with progressing left hemiplegia and lethargy, demonstrating an acute-on-chronic subdural hematoma.](image-url)
encourage brain expansion. High levels of inspired oxygen may help draw nitrogen out of the cavity. Regardless of the strategy used, follow-up head CT scans are required postoperatively and approximately 1 month later to document resolution.

**Intraparenchymal Hemorrhage** Isolated hematomas within the brain parenchyma are most often associated with hypertensive hemorrhage or arteriovenous malformations (AVMs). Bleeding may occur in a contused area of brain. Mass effect from developing hematomas may present as a delayed neurologic deficit. Delayed traumatic intracerebral hemorrhage is most likely to occur within the first 24 hours. Patients with contusion on the initial head CT scan should be reimaged 24 hours after the trauma to document stable pathology. Indications for craniotomy include: any clot volume >50 cm³ or a clot volume >20 cm³ with referable neurologic deterioration (GCS 6–8) and associated midline shift >5 mm or basal cistern compression.

**Pneumocephalus** Pneumocephalus, or air in the intracranial cavity, is commonly seen in neurosurgical patients following head trauma or following intracranial surgery. Pneumocephalus requires a defect in the skull that allows air to enter the intracranial cavity. This may occur following may represent an iatrogenic defect created following cranial surgery or following head trauma. Approximately 66% of postcraniotomy CT scans demonstrate some extent of pneumocephalus. The identification of pneumocephalus following head trauma can offer important clues about the extent of injury, such as the presence of skull based fractures or a CSF leak. In rare cases, pneumocephalus can also be seen in association with skull based tumors or infections.

A tension pneumocephalus occurs when the intracranial air pocket is under tension which can result in life threatening herniation if left untreated. This is a neurosurgical emergency and requires an urgent neurosurgical consultation. Two radiographic features have been associated with a tension pneumocephalus (Fig. 42-10). First, the “Mount Fuji” sign, where the air pocket separates the frontal lobes and widens the inter-hemispheric fissure, mimicking the silhouette of Mount Fuji. Second, the “air bubble” sign, where there are multi-focal pockets of air throughout the subarachnoid cisterns, putatively within the subarachnoid space. These radiographic findings are helpful clues; however, the diagnosis of tension pneumocephalus also requires a worsening neurological exam consistent with increased intracranial pressure and impending herniation. A burr hole may be used to relieve intracranial pressure per the discretion of the neurosurgical team.

When not associated with tension dynamics, the clinical significance and management of pneumocephalus depends on the underlying mechanism. There are thought to be two major mechanisms by which pneumocephalus develops. First, the “ball valve” mechanism involves the passage of air into the intracranial cavity during periods of positive pressure, whereby the defect in the skull acts as a one-way valve. In these cases, management involves avoiding positive pressure ventilation, and laying the head of the bed flat to minimize air traveling upwards into the cranial cavity. Second, the “inverted bottle” mechanism involves air entering the intracranial space due to a negative pressure gradient created by the drainage of CSF. In most cases, drainage occurs through a traumatic or iatrogenic CSF leak, but it may also occur through ventricular or lumbar drainage. In these cases, management should be focused on minimizing CSF drainage through the defect. If the CSF leak is at the skull base, as is the case following basilar skull fractures, or those involving the mastoid air cells, then the head of bed must be elevated so as to reduce hydrostatic pressure in the ventricular CSF fluid column, and controlled CSF diversion can be performed using an extraventricular or lumbar drain (see “Skull Fractures” for further discussion). Definitive repair of the skull-based defect can also be considered, but this is often done on an elective basis. In general, nontension pneumocephalus will resolve on its own with time as it is resorbed into the blood stream. Supplemental 100% oxygen may be used to increase the rate of resorption by increasing the diffusion gradient of nitrogen-predominant intracranial air pocket and the blood stream.

**Management of Traumatic Brain Injury**

**General Medical Management** Several medical steps may be taken to minimize secondary injury and the systemic consequences of head injury. Patients with a documented CHI and evidence of intracranial hemorrhage or a depressed skull fracture should receive a 1 g Keppra loading dose, followed by 1 week of therapeutic maintenance Keppra, typically 500 mg twice a day. Antiseizure prophylaxis has been shown to decrease the incidence of early posttraumatic seizures. There is no evidence to support long-term use of prophylactic antiepileptic agents. Even though the clinical studies supporting early antiseizure prophylaxis used phenytoin, Keppra is typically used in clinical practice due to a more favorable side effect profile. Blood glucose levels should be closely monitored by free blood sugar checks and controlled with sliding scale insulin. Fevers also should be evaluated and controlled with antipyretics, as well as source-directed therapy when possible. Hyperglycemia and hyperthermia are toxic to injured neurons and contribute to secondary injury. Head-injured patients have an increased prevalence of peptic ulceration and GI bleeding. Peptic ulcers occurring in patients with head injury or high ICP are referred to as Cushing’s ulcers. Ulcer prophylaxis should be used. Compression stockings or athrombic pumps should be used when the patient cannot be mobilized rapidly for prophylaxis of deep venous thrombosis.

**Steroids and Traumatic Brain Injury** Per a level 1 recommendation (high-quality evidence) from the Brain Trauma Foundation, steroids are not recommended for the management of TBI or reduction of elevated ICP. Also, high-dose methylprednisolone is contraindicated in severe TBI. A large randomized controlled trial (CRASH; n = 9673, 6-month follow-up, 24
demonstrated an increased risk of 6-month mortality in severe TBI (GCS 3–8) that received methylprednisolone (47%), as compared to placebo (42%, \( P = 0.0024 \)). This effect was also present when analyzing TBI of all severity levels (25.7% methylprednisolone vs. 22.3% placebo, \( P = 0.0001 \)).

**Blood Pressure Management** Blood pressure management in TBI is a complex issue. On one hand, hypotension results in hypoperfusion that may worsen brain injury that occurs following TBI. On another hand, hypertension may result in expansion of intracranial hematomas that are often seen in TBI.

There is clear evidence from retrospective studies, that frank hypotension (SBP <90 mmHg) is associated with increased mortality in TBI, particularly in the prehospital setting and during resuscitation.\(^2\) A large retrospective cohort study (\( n = 15,733 \)) identified hypotension thresholds that were associated with an increased risk of mortality in patients with TBI of varying age.\(^3\) Based on these data, the Brain Trauma Foundation guidelines\(^10\) provide a level III (low-quality) recommendation that maintaining systolic blood pressures >100 mmHg (ages 50–69 years), or >110 mmHg (ages 15–49 years or >70 years) may be considered to reduce mortality and improve outcomes. More recently, a large retrospective study\(^3\) demonstrated a dose-dependent relation between the duration of prehospital hypotension and increased mortality in patients with TBI, such that a 10-point increase in systolic blood pressure across a broad range (40–119 mmHg) was associated with an 18.8% decrease in adjusted odds of in-hospital mortality. These results suggest that having a single “hypotension threshold” may not be sufficient in management of TBI and may require more aggressive management than currently employed. Furthermore, an important and underappreciated consideration in blood pressure management is the baseline blood pressure of the patient. Future studies should assess blood pressure management goals that are tailored to each individual patient’s baseline blood pressure.

On the other hand, hypertension in TBI may have implications for intracranial hematoma expansion. It is common in clinical practice to recommend that systolic blood pressures are maintained <160 mmHg to mitigate the risk of hematoma expansion. Evidence supporting this practice is largely extrapolated from non-TBI patients. A small retrospective study (\( n = 69 \))\(^4\) demonstrated an increased risk of postcraniotomy intracranial hematoma in patients with intraoperative hypertension (62% vs. 34% controls, \( P < 0.001 \)), and postoperative hypertension in the first 12 hours after surgery (62% vs. 25% controls, \( P < 0.001 \)). A recent large retrospective study in patients with anticoagulant-associated intracranial hematoma demonstrated that lowering SBP to less than 160 mmHg within 4 hours of admission was associated with a reduced risk of hematoma expansion (\( n = 691 \), 33.1% <160 mmHg vs. 52.4% in ≥160 mmHg; \( P < 0.001 \)).\(^5\) However, there are no specific recommendations from the Brain Trauma Foundation on a hypertension threshold to avoid in patients with traumatic intracranial hematoma.

**Anticoagulation Reversal and Prophylaxis** Patients with intracranial hematoma who are on anticoagulation for cardiovascular indications (atrial fibrillation, cardiac stents, or mechanical valves) or stroke prevention present a challenging population. Anticoagulation reversal is important to reduce the risk of hematoma expansion; however, anticoagulant reversal is also associated with thrombotic cardiovascular complications. A recent retrospective study in patients with nontraumatic, oral-anticoagulant–associated intracranial hematoma showed that lowering the INR to <1.3 within 4 hours of admission was an independent predictor of hematoma expansion (\( n = 853 \); 19.8% vs. 41.5% in INR of ≥1.3; \( P < 0.001 \)). Furthermore, this study showed that the risk of ischemic complications was greater in patients that were not restarted on oral anticoagulation as compared to those that were subsequently restarted (\( n = 719 \); 5.2% vs. 15%, no restart, \( P < 0.001 \)); however, they did not observe a significant increase in the risk of hemorrhage with anticoagulation restart (\( n = 719 \); 8.1%, vs. 6.6%, \( P = 0.48 \)). The median time to anticoagulation restart was 30 days after discharge (inter-quartile range 18–65), as such, these data do not speak to risks and benefits of restarting anticoagulation in the acute postbleed interval. It is important to note that the risks and benefits of restarting anticoagulation will vary based on the individual patient and the patient’s indications for anticoagulation (e.g., mechanical heart valve vs. atrial fibrillation). As such, close collaboration between the neurosurgery and cardiology teams are important in optimizing a management strategy for these patients.

Anticoagulation prophylaxis for prevention of venous thrombosis also involves a risk-benefit analysis. Per a level III (low-quality) recommendation of Brain Trauma Foundation Guidelines,\(^10\) anticoagulation prophylaxis with low-molecular-weight heparin or low-dose unfractionated heparin may be used to reduce the risk of venous thrombosis, even though it is associated with an increased risk of intracranial hematoma expansion. It may be reasonable to initiate prophylactic anticoagulation 24 hours after an intracranial hematoma is deemed to be stable. A single-center retrospective study (\( n = 236 \)) found that such a strategy was associated with a decreased risk of DVT (0% vs. 5.6% (\( n = 6 \)), \( P < 0.001 \)), but did not observe significant differences in the rates of pulmonary embolism (0.78% (\( n = 1 \)) vs. 3.74% (\( n = 4 \), \( P = 0.18 \)) or intracranial hematoma expansion (0.7% [1] vs. 2.8% [3], \( P = 0.3 \)). However, because of the low rate of clinical events observed in this series, the study may have been underpowered to identify small differences in pulmonary embolism or hematoma expansion.

**Indications for Invasive Intracranial Monitoring** In patients with severe TBI (GCS <8), the Brain Trauma Foundation guidelines endorse a level IIB recommendation (low-quality of evidence) for ICP and cerebral perfusion pressure (CPP) to reduce short-term mortality (within 2 weeks of hospitalization). They also provide level IIB recommendations for treating ICP >22 mmHg and treating CPP level between 60 and 70 mmHg to optimize outcomes. These recommendations are supported by a recent retrospective cohort study (\( n = 459 \))\(^6\) that identified ICP and CPP thresholds that best discriminated between survivors and nonsurvivors in severe TBI, and also between survivors with “poor” and “favorable” outcomes (Glasgow Outcome Scale 1–3 vs. 4–6). A large, multicenter randomized controlled trial performed in 6 hospitals in Ecuador and Bolivia did not support the claim that intracranial monitoring in severe TBI results in improved clinical outcomes. Chestnut et al in 2012 did not observe a significant difference in mortality or favorable outcomes (as assessed by the Glasgow Outcome Scale) when severe TBI patients were managed with an intracranial monitor (\( n = 56 \), or with imaging and clinical exam (\( n = 53 \), \( P = 0.43 \)). Advanced multimodal monitoring such as brain tissue oxygen (PbrO2) monitoring, jugular bulb monitoring of arteriovenous oxygen content difference (AVDO2), cerebral autoregulation with TCD, and micro dialysis are under active investigation. Only jugular bulb monitoring of AVDO2 is associated with a level III (poor-quality evidence) recommendation to guide management in severe TBI.
The Brain Tissue Oxygen Monitoring in TBI (BOOST) trials are actively investigating the added benefit of brain tissue oxygenation beyond intracranial pressure monitoring in severe TBI.  

**Decompressive Cranietomy for Severe TBI** Decompressive craniectomy can be performed to relieve intracranial pressure associated with diffuse cerebral edema in cases of severe TBI without mass lesions (e.g., extra-axial hematoma). This is a controversial issue as there is a paucity of high-quality evidence providing clear support for or against this intervention. The DECRAs trial (a multicenter, randomized, controlled trial, n = 155) compared bifrontal decompressive craniectomy to medical management for the treatment of patients with severe TBI and elevated intracranial pressure refractory to first-tier therapies (ICP >20 mmHg for at least 15 minutes within an hour). They found no significant difference in mortality at six months, and found that functional outcomes (as measured by the Extended Glasgow Outcome Scale) were worse in patients who underwent surgery. They found a clear improvement in ICP and number of days in the ICU in patients that underwent surgery as compared to medical management. Of note, they used an intention-to-treat analysis, such that 18% of patients in the medical management group underwent a delayed craniotomy as a life-saving procedure. More recently, the RESCUE-ICP trial\(^{16}\) (a multicenter, randomized, controlled trial, n = 408) compared decompressive craniotomy and ongoing medical care in patients with severe TBI (without mass lesions) with elevated ICP (>25 mmHg) refractory to first- and second-tier interventions (medical management and ventriculostomy). Patients were randomized to either receive a barbiturate infusion (medical group) or undergo decompressive craniotomy (surgery group; unilateral hemicraniectomy vs. bifrontal craniotomy depending on degree of bilateral swelling and surgeon discretion). Again, they used an intention-to-treat analysis such that 37% of patients of the medical group underwent decompressive hemicraniectomy. At 6 months, decompressive craniectomy in patients with traumatic brain injury and refractory intracranial hypertension resulted in lower mortality and higher rates of vegetative state, lower severe disability, and upper severe disability than medical care. The rates of moderate disability and good recovery were similar in the two groups. The recent Brain Trauma Foundation Guidelines offer a level II (moderate-quality) recommendation against performing a bifrontal decompressive hemicraniectomy to improve functional outcomes at 6 months in patients with severe TBI with diffuse injury and no mass lesions, and with elevated ICP that is medically refractory. They note that this procedure has been demonstrated to reduce time in the ICU and ICP. However, they have not made an updated recommendation since the results of the RESCUE-ICP trial have been published. 

The results of the DECRAs and RESCUE-ICP trials suggest caution and careful consideration prior to performing decompressive craniotomy in treating severe TBI without mass lesions. There is now evidence that this procedure can be lifesaving and reduce mortality at 6 months; however, it is not clear that the survivors have a favorable functional outcome (as grossly measured by the Extended Glasgow Outcome Scale). By improving ICP and reducing time in the ICU, it may hasten the recovery process by allowing patients to begin rehabilitation earlier. Also, several unanswered questions remain. For example, might outcomes be improved if decompressive craniectomy was performed earlier, prior to the patient developing refractory ICP, and presumably secondary brain injury? As such, the decision of whether or not to perform decompressive craniotomy must be carefully considered within the context of each individual patient’s clinical scenario, the patient’s available social support system, and the family’s disposition and goals of care.  

**Vascular Injury.** Trauma to the head or neck may cause damage to the carotid or vertebrobasilar systems. Generally, dissection refers to violation of the vessel wall intima. Blood at arterial pressures can then open a plane between the intima and media, within the media, or between the media and adventitia. The newly created space within the vessel wall is referred to as the false lumen. Tissue or organs supplied by dissected vessels may subsequently be injured in several ways. Expansion of the hematoma within the vessel wall can lead to narrowing of the true vessel lumen and reduction or cessation of distal blood flow. Slow-flowing or stagnant blood within the false lumen exposed to thrombogenic vessel wall elements may thrombose. Pieces of thrombus may then detach and cause distal embolic arterial occlusion. Also, the remaining partial-thickness vessel wall may rupture, damaging adjacent structures.

Traumatic dissection may occur in the carotid artery (anterior circulation) or the vertebral or basilar arteries (posterior circulation). Dissections may be extradural or intradural. Intradural dissection can present with subarachnoid hemorrhage (SAH). Traditional angiography remains the basis of diagnosis and characterization of arterial dissection. Angiographic abnormalities include stenosis of the true lumen, or “string-sign,” visible intimal flaps, and the appearance of contrast in the false lumen. Four-vessel cerebral angiography should be performed when suspicion of dissection exists.

Historically, patients with documented arterial dissection have been anticoagulated with heparin and then warfarin to prevent thromboembolic stroke. Trauma patients often have concomitant absolute or relative contraindications to anticoagulation, complicating management. Antiplatelet therapy is often implemented in lieu of full anticoagulation, however, there is no randomized clinical trial comparing the two therapies.\(^{37}\) Consider surgical or interventional techniques for persisting embolic disease and for vertebral dissections presenting with SAH. Surgical options include vessel ligation and bypass grafting. Interventional radiology techniques include stenting and vessel occlusion. Occlusion techniques require sufficient collateral circulation to perfuse the vascular territory previously supplied by the occluded vessel.  

**Carotid Dissection** Carotid dissection may result from neck extension combined with lateral bending to the opposite side, or trauma from an incorrectly placed shoulder belt tightening across the neck in a motor vehicle accident. Extension or bending stretches the carotid over the bony transverse processes of the cervical vertebrae, while seat belt injuries cause direct trauma. Symptoms of cervical carotid dissection include contralateral neurologic deficit from brain ischemia, headache, and ipsilateral Horner’s syndrome from disruption of the sympathetic tracts ascending from the stellate ganglion on the surface of the carotid artery. The patient may complain of a bruise.

Traumatic vessel wall injury to the portion of the carotid artery running through the cavernous sinus may result in a carotid-cavernous fistula (CCF). This creates a high-pressure, high-flow pathophysiologic blood flow pattern. CCFs classically present with pulsatile proptosis (the globe pulses outward with arterial pulsation), retro-orbital pain, and decreased visual acuity or loss of normal eye movement (due to damage
to cranial nerves III, IV, and VI as they pass through the cavernous sinus. Symptomatic CCFs should be treated to preserve eye function. Fistulae may be closed by balloon occlusion using interventional neuroradiology techniques. Fistulae with wide necks are difficult to treat and may require total occlusion of the parent carotid artery.

Vertebrobasilar Dissection  Vertebrobasilar dissection may result from sudden rotation or flexion/extension of the neck, chiropractic manipulation, or a direct blow to the neck. Common symptoms are neck pain, headache, and brain stem stroke or SAH. The risks and benefits of aspirin therapy are unclear when a vertebral dissection extends intracranially. The theoretically increased friability of the vessel wall may increase the risk of SAH when coupled with an anticoagulant agent. Consultation of a stroke neurologist is recommended in this situation.

Brain Death.  Brain death occurs when there is an absence of signs of brain stem function or motor response to deep central pain in the absence of pharmacologic or systemic medical conditions that could impair brain function.

Clinical Examination  A neurologist, neurosurgeon, or intensivist generally performs the clinical brain death examination. Two examinations consistent with brain death 12 hours apart, or one examination consistent with brain death followed by a consistent confirmatory study generally is sufficient to declare brain death (see following paragraphs). Hospital regulations and local laws regarding documentation should be followed closely.

Establish the absence of complicating conditions before beginning the examination. The patient must be normotensive, euvolemic, and oxygenating well. The patient may not be under the effects of any sedating or paralytic drugs.

Documentation of no brain stem function requires the following: nonreactive pupils; lack of corneal blink, oculocephalic (doll’s eyes), oculovestibular (cold calorics) reflexes; and loss of drive to breathe (apnea test). The apnea test demonstrates no spontaneous breathing even when Paco2 is allowed to rise above 60 mmHg.

Deep central painful stimuli are provided by bilateral forceful twisting pinch of the supraclavicular skin and pressure to the medial canthal notch. Pathologic responses such as flexor or extensor posturing are not compatible with brain death. Spinal reflexes to peripheral pain, such as triple flexion of the lower extremities, are compatible with brain death.

Confirmatory Studies  Confirmatory studies are performed after a documented clinical examination consistent with brain death. A study consistent with brain death may obviate the need to wait 12 hours for a second examination. This is especially important when the patient is a potential organ donor, as brain-dead patients often have progressive hemodynamic instability. Lack of cerebral blood flow consistent with brain death may be documented by cerebral angiography or technetium radionuclide study. A “to-and-fro” pattern on transcranial Doppler ultrasonography indicates no net forward flow through the cerebral vasculature, consistent with brain death. An electroencephalogram (EEG) documenting electrical silence has been used but generally is not favored because there is often significant artifact which impairs interpretation.

Spine Trauma  The spine is a complex biomechanical structure. The spine provides structural support for the body as the principal component of the axial skeleton, while protecting the spinal cord and nerve roots. Trauma may fracture bones or cause ligamentous disruption. Often, bone and ligament damage occur together. Damage to these elements reduces the strength of the spine and may cause instability, which compromises both supportive and protective functions. Spine trauma may occur with or without neurologic injury.

Neurologic injury from spine trauma is classified as either incomplete or complete. If there is some residual motor or sensory neurologic function below the level of the lesion, as assessed by clinical examination, the injury is defined as incomplete. A patient with complete neurologic dysfunction persisting 24 hours after injury has a very low probability of return of function in the involved area.

Neurologic injury from spine trauma may occur immediately or in delayed fashion. Immediate neurologic injury may be due to direct damage to the spinal cord or nerve roots from penetrating injuries, especially from stab wounds or gunshot wounds. Blunt trauma may transfer sufficient force to the spine to cause acute disruption of bone and ligament, leading to subluxation, which is a shift of one vertebral element in relation to the adjacent level. Subluxation decreases the size of the spinal canal and neural foramina and causes compression of the cord or roots. Such neural impingement can also result from retropulsion of bone fragments into the canal during a fracture. Transection, crush injury, and cord compression impairing perfusion are mechanisms leading to SCI. Delayed neurologic injury may occur during transportation, examination of an improperly immobilized patient, or during a hypotensive episode.

The Mechanics of Spine Trauma  Trauma causes a wide variety of injury patterns in the spine due to its biomechanical complexity. A mechanistic approach facilitates an understanding of the patterns of injury, as there are only a few types of forces that can be applied to the spine. Although these forces are discussed individually, they often occur in combination. Several of the most common injury patterns are then presented to illustrate the clinical results of these forces applied at pathologically high levels.

Flexion/Extension  Bending the head and body forward into a fetal position flexes the spine. Flexion loads the spine anteriorly (the vertebral bodies) and distracts the spine posteriorly (the spinous process and interspinous ligaments). High flexion forces occur during front-end motor vehicle collisions, and backward falls when the head strikes first. Arching the neck and back extends the spine. Extension loads the spine posteriorly and distracts the spine anteriorly. High extension forces occur during rear-end motor vehicle collisions (especially if there is no headrest), forward falls when the head strikes first, or diving into shallow water.

Compression/Distraction  Force applied along the spinal axis (axial loading) compresses the spine. Compression loads the spine anteriorly and posteriorly. High compression forces occur when a falling object strikes the head or shoulders, or when landing on the feet, buttocks, or head after a fall from height. A pulling force in line with the spinal axis distracts the spine. Distraction unloads the spine anteriorly and posteriorly. Distraction forces occur during a hanging, when the chin or occiput strikes an object first during a fall, or when a passenger submarines under a loose seat belt during a front-end motor vehicle collision.

Rotation  Force applied tangential to the spinal axis rotates the spine. Rotation depends on the range of motion of intervertebral facet joints. High rotational forces occur during off-center...
impacts to the body or head or during glancing automobile accidents.

**Patterns of Injury.** Certain patterns of injury resulting from combinations of the previously mentioned forces occur commonly and should be recognized during plain film imaging of the spine. Always completely evaluate the spine. A patient with a spine injury at one level has a significant risk for additional injuries at other levels.

**Cervical** The cervical spine is more mobile than the thoracolumbar spine. Stability comes primarily from the multiple ligamentous connections of adjacent vertebral levels. Disruption of the cervical ligaments can lead to instability in the absence of fracture. The mass of the head transmits significant forces to the cervical spine during abrupt acceleration or deceleration, increasing risk for injury.

**Jefferson Fracture** A Jefferson fracture is a bursting fracture of the ring of C1 (the atlas) due to compression forces. There are usually two or more fractures through the ring of C1. The open-mouth odontoid view may show lateral dislocation of the lateral masses of C1. The rule of Spence states that 7 mm or greater combined dislocation indicates disruption of the transverse ligament. The transverse ligament stabilizes C1 with respect to C2. Jefferson fractures dislocated <7 mm usually are treated with a rigid collar, while those dislocated 7 mm or greater usually are treated with a halo vest. Surgical intervention is not indicated.

**Odontoid Fractures** The odontoid process, or dens, is the large ellipse of bone arising anteriorly from C2 (the axis) and projecting up through the ring of C1 (the atlas). Several strong ligaments connect the dens to C1 and to the base of the skull. Odontoid fractures usually result from flexion forces. Odontoid fractures are classified as type I, II, or III. A type I fracture involves the tip only. A type II fracture passes through the base of the odontoid process. A type III fracture passes through the body of C2. Types II and III are considered unstable and should be externally immobilized or fused surgically. Surgery often is undertaken for widely displaced fractures (poor chance of fusing) and for those that fail external immobilization. Type I fractures usually fuse with external immobilization only.

**Hangman’s Fracture** Traditionally considered a hyperextension/distraction injury from placement of the noose under the angle of the jaw, hangman’s fractures also may occur with hyperextension/compression, as with diving accidents, or hyperflexion. The injury is defined by bilateral C2 pars interarticularis fractures. The pars interarticularis is the bone between superior and inferior facet joints. Thus, the posterior bony connection between C1 and C3 is lost. Hangman’s fractures heal well with external immobilization. Surgery is indicated if there is spinal cord compression or after failure of external immobilization.

**Jumped Facets—Hyperflexion Injury** The facet joints of the cervical spine slope forward. In a hyperflexion injury, the superior facet can “jump” over the inferior facet of the level above if the joint capsule is torn. Hyperflexion/rotation can cause a unilateral jumped facet, whereas hyperflexion/distraction leads to bilateral jumped facets. Patients with unilateral injury usually are neurologically intact. Those with bilateral injury, however, typically suffer from spinal cord damage, since the anteroposterior diameter of the spinal canal is compromised by bilateral injury, leading to spinal cord compression (Fig. 42-11).

**Thoracolumbar** The thoracic spine is stabilized significantly by the rib cage. The lumbar spine has comparatively large vertebræ. Thus, the thoracolumbar spine has a higher threshold for injury than the cervical spine. A three-column model is useful for categorizing thoracolumbar injuries.⁵⁹ The anterior longitudinal ligament and the anterior half of the vertebral body constitute the anterior column. The posterior half of the vertebral body and the posterior longitudinal ligament constitute the middle column. The pedicles, facet joints, laminae, spinous processes, and interspinous ligaments constitute the posterior column.

**Compression Fracture** Compression fracture is a compression/flexion injury causing failure of the anterior column only. It is stable and not associated with neurologic deficit, although the patient may still have significant pain (Fig. 42-12).

**Burst Fracture** Burst fracture is a pure axial compression injury causing failure of the anterior and middle columns. It is unstable, and perhaps half of patients have neurologic deficit due to compression of the cord or cauda equina from bone fragments retropropulsed into the spinal canal.

**Chance Fracture** Chance fracture is a flexion-distraction injury causing failure of the middle and posterior columns, sometimes with anterior wedging. Typical injury is from a lap seat-belt hyperflexion with associated abdominal injury. It often is unstable and associated with neurologic deficit.

**Fracture-Dislocation** Fracture-dislocation is failure of the anterior, middle, and posterior columns caused by flexion/distraction, shear, or compression forces. Neurologic deficit can result from retropulsion of middle column bone fragments into the spinal canal, or from subluxation causing decreased canal diameter (Fig. 42-13).

**Initial Assessment and Management.** The possibility of a spine injury must be considered in all trauma patients. A patient with no symptoms referable to neurologic injury, a normal neurologic examination, no neck or back pain, and a known mechanism of injury unlikely to cause spine injury is at minimal risk for significant injury to the spine. Victims of moderate or severe trauma, especially those with injuries to other organ systems, usually fail to meet these criteria or cannot be assessed adequately. The latter often is due to impaired sensorium or significant pain. Because of the potentially catastrophic consequences of missing occult spine instability in a neurologically intact patient, a high level of clinical suspicion should govern patient care until completion of clinical and radiographic evaluation.

The trauma patient should be kept on a hard, flat board with straps and pads used for immobilization. A hard cervical collar is kept in place. These steps minimize forces transferred through the spine and therefore decrease the chance of causing dislocation, subluxation, or neural compression during transport to the trauma bay. The patient is then moved from the board to a flat stretcher. The primary survey and resuscitation are completed. Physical examination and initial X-rays follow.

For the examination, approach the patient as described in “Neurologic Examination” earlier in this chapter. Evaluation for spine or SCI is easier and more informative in awake patients. If the patient is awake, ask if he or she recalls details of the nature of the trauma, and if there was loss of consciousness, numbness, or inability to move any or all limbs. Assess motor function by response to commands or pain, as appropriate. Assess pinprick, light touch, and joint position, if possible. Determining the anatomically lowest level of intact sensation can pinpoint the level of the lesion along the spine. Testing sensation in an ascending fashion will allow the patient to better discern the true stimulus
as opposed to determine when it is extinguished. Document muscle stretch reflexes, lower sacral reflexes (i.e., anal wink and bulbocavernosus), and rectal tone.

American Spinal Injury Association Classification. The American Spinal Injury Association provides a method of classifying patients with spine injuries. The classification indicates completeness and level of the injury and the associated deficit.

Figure 42-11. A. Lateral cervical spine X-ray of an elderly woman who struck her head during a backward fall. Arrowhead points to jumped facets at C5–C6. Note the anterior displacement of the C5 body with respect to the C6 body. B. Sagittal T2-weighted magnetic resonance imaging of the same patient, revealing compromise of the spinal canal and compression of the cord. Note the bright signal within the cord at the level of compression, indicating spinal cord injury. C. Lateral cervical spine X-ray of same patient after application of cervical traction and manual reduction. Note restoration of normal alignment. D. Lateral cervical spine X-ray after posterior cervical fusion to restabilize the C5–C6 segment of the spine.

A form similar to that shown in Fig. 42-14 should be available in the trauma bay and completed for any spine injury patient. The association also has worked to develop recommendations and guidelines to standardize the care of SCI patients in an effort to improve the quality of care.

Neurologic Syndromes. Penetrating, compressive, or ischemic cord injury can lead to several characteristic presentations
Figure 42-12. A. Lateral lumbar spine X-ray showing a compression fracture of L2. Arrowhead points to anterior wedge deformity. Note the posterior wall of the vertebral body has retained normal height and alignment. B. Axial computed tomography scan through the same fracture. Arrowhead demonstrates a transverse discontinuity in the superior endplate of the L2 body.

Figure 42-13. Sagittal reconstruction of an axial fine-slice computed tomography scan through the lumbar spine demonstrating a severe fracture-dislocation through the body of L2.

Based on the anatomy of injury. The neurologic deficits may be deduced from the anatomy of the long sensory and motor tracts and understanding of their decussations (Fig. 42-15). Four patterns are discussed. First, injury to the entire cord at a given level results in anatomic or functional cord transection with total loss of motor and sensory function below the level of the lesion. The typical mechanism is severe traumatic vertebral subluxation reducing spinal canal diameter and crushing the cord. Second, injury to half the cord at a given level results in Brown-Séquard syndrome, with loss of motor control and proprioception ipsilaterally and loss of nociception and thermoception contralaterally. The typical mechanism is a stab or gunshot wound. Third, injury to the interior gray matter of the cord in the cervical spine results in a central cord syndrome, with upper extremity worse than lower extremity weakness and various degrees of numbness. The typical mechanism is transient compression of the cervical cord by the ligamentum flavum buckling during traumatic neck hyperextension. This syndrome occurs in patients with preexisting cervical stenosis. Fourth, injury to the ventral half of the cord results in the anterior cord syndrome, with paralysis and loss of nociception and thermoception bilaterally. The typical mechanism is an acute disc herniation or ischemia from anterior spinal artery occlusion.

Studies. Anteroposterior and lateral plain films provide a rapid survey of the bony spine. Plain films detect fractures and dislocations well. Adequate visualization of the lower cervical and upper thoracic spine often is impossible because of the shoulder girdle. Complete plain film imaging of the cervical spine includes an open-mouth view to assess the odontoid process and the lateral masses of C1. Fine-slice CT scan with sagittal and coronal reconstructions provides good detail of bony anatomy and is good for characterizing fractures seen on plain films, as well as visualizing C7–T1 when not well seen on plain films. MRI provides the best soft tissue imaging. Canal compromise from subluxation, acute disc herniations, or ligamentous disruption is clearly seen. MRI also may detect EDHs or damage to the spinal cord itself, including contusions or areas of ischemia.
Indications for Screening for Vascular Injury with Cervical Spine Trauma

It is important to consider the presence of blunt cerebrovascular injury in patients with cervical spine trauma; however, the specific indications for obtaining a screening CT angiography study are controversial. Many trauma centers rely on the Denver Criteria. These criteria indicate that screening should be employed for a cervical vertebral body or transverse foramen fracture, subluxation, or ligamentous injury at any level, or any fracture at the level of C1–C3, among other indications such as concerning mechanism, Lefort mid-face fractures, or basilar fractures through the carotid canal. A recent single-center retrospective study (n = 1717 cervical spine fractures) found a higher risk of vertebral artery injury only in the setting of fractures of C1 and C2 (combined), those that involve the transverse foramen, or had significant subluxation. They did not find that “high-risk” cervical spine fractures as defined by the Denver Criteria were associated with an increased risk of blunt cerebrovascular injury. Based on these data, we feel that it is appropriate to use a more defined set of screening criteria as outlined by Lockwood et al to reduce cost and contrast-exposure in patients with cervical spine fractures.
Definitive Management

**Spinal-Dose Steroids** Several studies have investigated the use of methylprednisolone in acute spinal cord injury. The National Acute Spinal Cord Injury studies (NASCIS I, II, and III) provided some support for the view that administration of high-dose methylprednisolone in acute spinal cord injury results in improved neurologic outcomes. A post-hoc analysis performed as part of NASCIS II demonstrated improved neurologic outcomes if methylprednisolone was administered within 8 hours of injury. A post-hoc analysis in NASCIS III showed improved outcomes at 6 weeks and 6 months, but not 1 year when methylprednisolone was administered within 3 and 8 hours of injury. However, these findings are tempered by the fact that these benefits were modest and only demonstrated in post-hoc analyses, and by the high rate of medical complications associated with corticosteroid administration. All three NASCIS trials showed that methylprednisolone was associated with a higher rate of complications such as pneumonia, severe sepsis, and poor wound healing. A recent Cochrane review did not observe a significant increase in complications or mortality associated with methylprednisolone administration in acute spinal cord injury, but did observe a trend towards this effect. Despite the lack of clear evidence on this issue, the most recent acute spinal cord injury guidelines provide a controversial level I recommendation against the use of corticosteroids in acute spinal cord injury. Some authors have argued for the use of methylprednisolone within 8 hours of acute spinal cord injury in carefully selected patients (e.g., young males that are less prone to medical complications associated with corticosteroids). Thus, clear consensus on the use of spinal-dose steroids does not exist. A decision to use or not use spinal-dose steroids may be dictated by local or regional practice patterns, especially given the legal liability issues surrounding SCI. Patients with gunshot or nerve root (cauda equina) injuries, or those who are pregnant, <14 years old, or on chronic steroids were excluded from the NASCIS studies and should not receive spinal-dose steroids. In addition to steroids, hypothermia for SCI has also received attention. There is even less evidence supporting the use of this treatment, and thus, it is not currently recommended.

**Orthotic Devices** Rigid external orthotic devices can stabilize the spine by decreasing range of motion and minimizing stress transmitted through the spine. Commonly used rigid cervical orthoses include Philadelphia and Miami-J collars. Cervical collars are inadequate for C1, C2, or cervicothoracic instability. Cervicothoracic orthoses brace the upper thorax and the neck, improving stabilization over the cervicothoracic region. Minerva braces improve high cervical stabilization by bracing from the upper thorax to the chin and occiput. Halo vest assemblies provide the most external cervical stabilization. Four pins are driven into the skull to lock the halo ring in position. Four posts arising from a tight-fitting rigid plastic vest immobilize the halo ring. Lumbar stabilization may be provided by thoracolumbosacral orthoses. A variety of companies manufacture lines of spinal orthotics. A physician familiar with the technique should fit a halo vest. Assistance from a trained orthotics technician improves fitting and adjustment of the other devices.

**Surgery** Neurosurgical intervention has two goals: decompression of the spinal cord and nerve roots, and stabilization of the spine. When spinal cord injury is caused by a hyperflexion-distraction injury that results in cord compression and an unstable spine, both surgical decompression and fusion are typically required. However, in cases of a hypextension injury causing central cord syndrome due to chronic cervical stenosis, surgical decompression may be needed without the need for internal fixation. In cases where there significant anterolisthesis (subaxial cervical jumped or perched facets), reduction of the fracture may be important for both decompression and stabilization. However, in some cases, reduction of the fracture may not be sufficient for decompression, and further decompressive surgery may still be needed.

Several controversial topics require consideration here. First, the timing of surgery has been a controversial topic. In the past, it has been suggested that patients with incomplete injury, or a deteriorating exam warrant emergent decompression, whereas patients with complete injuries can undergo surgery in a nonemergent manner. It is important to appreciate the risks of taking a medically unstable patient to surgery, such as a polytrauma patient with hemorrhagic shock or a complete spinal cord injury patient in neurogenic shock. However, there are clear benefits to early surgical decompression: it can allow early mobilization, aggressive nursing care, and physical therapy. Furthermore, a recent prospective cohort study found that the odds of observing a two-point increase in ASIA grade at 6 months was higher in patients that underwent surgery within 24 hours, as compared to those that underwent surgery after 24 hours. These data suggest that a subpopulation of patients may significantly benefit from early surgery; however, the characteristics of these patients were not described, suggesting a heterogeneous population in terms of preoperative ASIA grade and imaging.

In general, spine trauma patients with complete neurologic deficit, without any signs of recovery, or those without any neurologic deficits who have bony or ligamentous injury requiring open fixation, may be medically stabilized before undergoing surgery. Surgical stabilization may be indicated for some injuries that would eventually heal with conservative treatment. Solid surgical stabilization may also allow a patient to be managed with a rigid cervical collar who would otherwise require halo-vest immobilization.

**Continued Care.** Regional SCI centers with nurses, respiratory therapists, pulmonologists, physical therapists, physiatrists, and neurosurgeons specifically trained in caring for these patients may improve outcomes. Frequently encountered ICU issues include hypotension due to neurogenic shock (due to loss of sympathetic tone) and aspiration pneumonia. The recent guidelines recommend maintaining MAPs >85 for 7 days after injury. Chronically, prevention and treatment of deep venous thrombosis, autonomic hyperreflexia, and decubitus ulcer formation are important. Many patients with cervical or high thoracic cord injuries require prolonged ventilatory support until the chest wall becomes stiff enough to provide resistance for diaphragmatic breathing. Patients with high cervical cord injuries (C4 or above) will often require permanent ventilatory support. Patients should be transferred to SCI rehabilitation centers after stabilization of medical and surgical issues.

**Peripheral Nerve Trauma**

The peripheral nervous system extends throughout the body and is subject to injury from a wide variety of trauma. Peripheral nerves transmit motor and sensory information from the CNS to the body. An individual nerve may have pure motor, pure sensory, or mixed motor and sensory functions. The key
information-carrying structure of the nerve is the axon. The axon transmits information from the neuronal cell body and may measure from <1 mm to >1 m in length. Axons that travel a significant distance are often covered with myelin, which is a lipid-rich, electrically insulating sheath formed by Schwann cells. Myelinated axons transmit signals much more rapidly than unmyelinated axons because the voltage shifts and currents that define action potentials effectively jump from gap to gap over the insulated lengths of the axon.

Axons, whether myelinated or unmyelinated, travel through a collagenous connective tissue known as [endoneurium]. Groups of axons and their endoneurium form bundles known as [fascicles]. Fascicles run through a tubular collagenous tissue known as [perineurium]. Groups of fascicles are suspended in mesoneurium. Fascicles and their mesoneurium run through another tubular collagenous tissue known as [epineurium]. The epineurium and its contents form the nerve.

There are four major mechanisms of injury to peripheral nerves. Nerves may be lacerated, stretched, compressed, or contused. Knives, passing bullets, or jagged bone fractures may lacerate nerves. Adjacent expanding hematomas or dislocated fractures may stretch nerves. Expanding hematomas, external orthoses such as casts or braces, or blunt trauma over a superficial nerve may compress or crush nerves. Shock waves from high-velocity bullets may contuse nerves. These mechanisms of injury cause damage to the various anatomic components of the nerve. The patterns of damage are categorized in “Types of Injury.”

Certain nerve segments are particularly vulnerable to injury. The following four characteristics make a nerve segment more vulnerable: proximity to a joint, superficial course, passage through a confined space, and being fixed in position.

**Types of Injury.** The traditional classification system for peripheral nerve injury is the Seddon classification. Seddon described three injury patterns as defined in the “Neurapraxia,” “Axonotmesis,” and “Neurotmesis” sections. The Seddon classification provides a simple, anatomically based approach to peripheral nerve injury.50

**Neurapraxia** Neurapraxia is defined as the temporary failure of nerve function without physical axonal disruption. Axon degeneration does not occur. Return of normal axonal function occurs over hours to months, often in the 2- to 4-week range.

**Axonotmesis** Axonotmesis is the disruption of axons and myelin. The surrounding connective tissues, including endoneurium, are intact. The axons degenerate proximally and distally from the area of injury. Distal degeneration is known as Wallerian degeneration. Axon regeneration within the connective tissue pathways can occur, leading to restoration of function. Axons regenerate at a rate of 1 mm per day. Significant functional recovery may occur for up to 18 months. Scarring at the site of injury from connective tissue reaction can form a neuroma and interfere with regeneration.

**Neurotmesis** Neurotmesis is the disruption of axons and endoneurial tubes. Peripheral collagenous components, such as the epineurium, may or may not be intact. Proximal and distal axonal degeneration occurs. The likelihood of effective axonal regeneration across the site of injury depends on the extent of neuroma formation and on the degree of persisting anatomic alignment of the connective tissue structures. For instance, an injury may damage axons, myelin, and endoneurium, but leave perineurium intact. In this case, the fascicle sheath is intact, and appropriate axonal regeneration is more likely to occur than if the sheath is interrupted.

**Management of Peripheral Nerve Injury.** The sensory and motor deficits should be accurately documented. Deficits are usually immediate. Progressive deficit suggests a process such as an expanding hematoma and may warrant early surgical exploration. Clean, sharp injuries may also benefit from early exploration and reanastomosis. Most other peripheral nerve injuries should be observed. EMG/NCS studies should be done 3–6 to weeks postinjury if deficits persist. Axon segments distal to the site of injury will conduct action potentials normally until Wallerian degeneration occurs, rendering EMG/NCS before 3 weeks uninformative. Continued observation is indicated if function improves. Surgical exploration of the nerve may be undertaken if no functional improvement occurs over 3 months. If intraoperative electrical testing reveals conduction across the injury, continue observation. In the absence of conduction, the injured segment should be resected and end-to-end primary anastomosis attempted. However, anastomoses under tension will not heal. A nerve graft may be needed to bridge the gap between the proximal and distal nerve ends. The sural nerve is often harvested, as it carries only sensory fibers and leaves a minor deficit when resected. The connective tissue structures of the nerve graft may provide a pathway for effective axonal regrowth across the injury.

**Patterns of Injury**

**Brachial Plexus** The brachial plexus may be injured in a variety of ways. Parturition or a motorcycle accident can lead to plexus injury due to dislocation of the glenohumeral joint. Attempting to arrest a fall with one’s hands can lead to a stretch injury of the plexus due to abrupt movement of the shoulder girdle. An apical lung (Pancoast) tumor can cause compression injury to the plexus. There are many patterns of neurologic deficits possible with injury to the various components of the brachial plexus, and understanding them all would require extensive neuroanatomic discussion. Two well-known eponymous syndromes are Erb’s palsy and Klumpke’s palsy. Injury high in the plexus to the C5 and C6 roots resulting from glenohumeral dislocation causes Erb’s palsy with the characteristic “bellhop’s tip” position. The arm hangs at the side, internally rotated. Hand movements are not affected. Injury low in the plexus, to the C8 and T1 roots, resulting from stretch or compression injury, causes Klumpke’s palsy with the characteristic “claw hand” deformity. There is weakness of the intrinsic hand muscles, similar to that seen with ulnar nerve injury.

**Radial Nerve** The radial nerve courses through the axilla, then laterally and posteriorly in the spiral groove of the humerus. Improper crutch use can cause damage to the axillary portion. The section of the nerve traversing the spiral groove can be damaged by humerus fractures or pressure from improper positioning during sleep. This classically occurs when the patient is intoxicated and is called “Saturday night palsy.” The key finding is wrist drop (i.e., weakness of hand and finger extensors). Axillary (proximal) injury causes triceps weakness in addition to wrist drop.

**Common Peroneal Neuropathy** The common peroneal nerve forms the lateral half of the sciatic nerve (the medial half being the tibial nerve). It receives contributions from L4, L5, S1, and S2. It emerges as a separate nerve in the popliteal fossa and laterally wraps around the fibular neck, after which it splits to
form the deep and superficial peroneal nerves. The superficial, fixed location at the fibular neck makes the common peroneal nerve susceptible to compression. The classic cause of traumatic peroneal neuropathy is crush injury from a car bumper striking the lateral aspect of the leg at the knee. Symptoms of common peroneal neuropathy include foot drop (weakness of the tibialis anterior), evasion weakness, and numbness over the anterolateral surface of the lower leg and dorsum of the foot. In contrast, a foot drop due to L5 radiculopathy spares evasion because the S1 fibers are intact. Surgical exploration of a common peroneal crush lesion is typically a low yield endeavor. Rare cases may be due to compressive fibers or adhesions that may be lysed, with the possibility of return of function.

**CEREBROVASCULAR DISEASE**

Cerebrovascular disease is the most frequent cause of new, rapid-onset, nontraumatic neurologic deficit. It is far more common than seizures or tumors. Vascular structures are subject to a variety of chronic pathologic processes that compromise vessel wall integrity. Diabetes, high cholesterol, high blood pressure, and smoking – common comorbidities in the general population – are important risk factors for vascular disease. These conditions can lead to vascular damage by such mechanisms as atheroma deposition causing luminal stenosis, endothelial damage promoting thrombogenesis, and weakening of the vessel wall resulting in aneurysm formation or dissection. These processes may coexist. For instance, a vessel containing an atheromatous plaque will have a decreased luminal diameter. The plaque also may have compromised endothelium, providing the opportunity for thrombus formation, which can lead to acute total occlusion of the remaining lumen. Aneurysms and dissection often occur in atheromatous vessels. Specific patterns of disease relevant to the cerebrovascular system include atheromatous and thrombotic carotid occlusion, brain ischemia from proximal embolic disease, vessel wall rupture leading to hemorrhage, and rupture of abnormal, thin-walled structures, specifically aneurysms and AVMs.

**Ischemic Diseases**

Ischemic stroke accounts for approximately 85% of acute cerebrovascular events. Symptoms of acute ischemic stroke vary based on the functions of the neural tissues supplied by the occluded vessel, and the presence or absence of collateral circulation. The circle of Willis provides extensive collateral circulation, as it connects the right and left carotid arteries to each other and each to the vertebrobasilar system. Patients with complete occlusion of the carotid artery proximal to the circle of Willis may be asymptomatic if the blood flow patterns can shift and provide sufficient circulation to the ipsilateral cerebral hemisphere from the contralateral carotid and the basilar artery. However, the anatomy of the circle of Willis is highly variable. Patients may have a congenitally hypoplastic or missing communicating artery with resultant bilateral ACA supply by one carotid, or the PCA may be supplied by the carotid artery rather than the basilar. Similarly, one vertebral artery is often dominant and the other is hypoplastic. These variations may make disease in a particular vessel more neurologically devastating than in a patient with full collateral circulation. Occlusion distal to the circle of Willis generally results in a stroke in the territory supplied by that particular artery.

Neurologic deficit from occlusive disease may be temporary or permanent. A patient with sudden-onset focal neurologic deficit that resolves within 24 hours has had a transient ischemic attack (TIA). A patient with permanent deficit has had a completed stroke.

**Thrombotic Disease**

The most common area of neurologically significant vessel thrombosis is the carotid artery in the neck. Disease occurs at the carotid bifurcation. Thrombosis of a carotid artery chronically narrowed by atheroma can lead to acute carotid occlusion. As discussed previously, this can be asymptomatic due to sufficient collateralization. The more common concern is thromboembolus. Intracranial arterial occlusion by local thrombus formation may occur, but it is rare compared to embolic occlusion.

**Management.** Complete occlusion of the carotid artery without referable neurologic deficit requires no treatment. A patient with new neurologic deficit and an angiographically confirmed complete carotid occlusion contralateral to the symptoms should be considered for emergent carotid endarterectomy. Surgery should be performed within 2 hours of symptom onset and should not be performed on obtunded or comatose patients. These restrictions significantly reduce the number of operative candidates. In nonemergent cases, the results of the large-scale North American Symptomatic Carotid Stenosis Trial (NASCET) demonstrated a stroke reduction benefit to surgical revascularization in patients with severe stenosis, defined as occlusion of 70% to 99% of the carotid. Practice guidelines recommend revascularization at this level of stenosis even if asymptomatic. Surgical options for these patients include both carotid endarterectomy and carotid stenting, which have been shown to produce equal outcomes over long-term follow-up.

**Embolic Disease**

Emboli causing strokes may originate from a number of sources, including: the left atrium in atrial fibrillation, a hypokinetic left ventricular wall segment, valvular vegetations, an atheromatous aortic arch, stenotic/atheromatous carotid bifurcations, or from the systemic venous system in the presence of a right-to-left shunt, such as a patent foramen ovale. The majority of emboli enter the anterior (carotid) circulation rather than the posterior (vertebrobasilar) circulation. Characteristic clinical syndromes result from embolic occlusion of various vessels within these circulations.

**Common Types of Strokes**

**Anterior Cerebral Artery Stroke** The ACA supplies the medial frontal and parietal lobes as it courses into the interhemispheric fissure. Due to its vascular supply of the motor cortex, ACA stroke characteristically results in contralateral leg weakness.

**Middle Cerebral Artery Stroke** The MCA supplies the lateral frontal and parietal lobes as it courses into the interhemispheric fissure. Due to its vascular supply of the motor cortex, ACA stroke characteristically results in contralateral leg weakness. MCA stroke results in contralateral face and arm weakness. Dominant-hemisphere MCA stroke causes language deficits due to its supply of Broca’s area, Wernicke’s area, and the white matter tracts that connect the two. Proximal MCA occlusion with ischemia and swelling in the entire MCA territory can lead to significant intracranial mass effect and midline shift (see Fig. 42-6), termed malignant MCA stroke.

**Posterior Cerebral Artery Stroke** The PCA supplies the occipital lobe. PCA stroke results in a contralateral homonymous hemianopsia (see Fig. 42-6).

**Posterior Inferior Cerebellar Artery Stroke** The PICA supplies the lateral medulla and the inferior half of the cerebellar
hemispheres. PICA stroke results in nausea, vomiting, nystagmus, dysphagia, ipsilateral Horner’s syndrome, and ipsilateral limb ataxia. The constellation of symptoms resulting from PICA occlusion is referred to as the lateral medullary or Wallenberg’s syndrome.

Management. Ischemic stroke management has two goals: reopen the occluded vessel and maintain blood flow to ischemic “penumbra” tissues bordering the vascular territory. Reopening the vessel has historically been attempted with recombinant tPA. tPA administration within 3 hours of the onset of neurologic deficit improves outcome at 3 months. In the setting of suspected ischemic stroke, a head CT must be performed immediately to differentiate ischemic from hemorrhagic stroke. Intracranial hemorrhage, major surgery within the previous 2 weeks, GI or genitourinary hemorrhage in the previous 3 weeks, platelet count less than 100,000/µL, and systolic blood pressure >185 mmHg are among the contraindications to tPA therapy.

In recent years, a paradigm shift in ischemic stroke management has occurred with the advent of endovascular mechanical thrombectomy. Though tPA can be effective for strokes of smaller vessels, it produces recanalization in only 20% of large vessel ischemic strokes, and even less for internal carotid artery occlusion. In mechanical thrombectomy, the intracranial circulation is accessed endovascularly, and stent-retriever devices can be deployed to definitively remove clot and stent open involved vessels. Initial investigations into the technology began in 1999 with varied success. Technological advances proceeded, and in 2015, large-scale clinical trials investigating mechanical thrombectomy for large vessel occlusion were published. The MR-CLEAN trial was one of these. In this trial, 500 patients were randomized to tPA and medical therapy vs. mechanical thrombectomy. The latter group had a significantly better 90-day outcome. A domino effect ensued in which four other similarly large-scale trials (e.g., REVASCAT) were terminated early.

Benefits to mechanical thrombectomy include less stringent criteria than tPA administration, focused therapy, and time windows between 6 and 12 hours in trials. Currently, mechanical thrombectomy is being used for large vessel occlusion within 6 hours of symptom onset for those patients not eligible for tPA, which accounts for about 10% of ischemic stroke. Its indications and associated technology continue to evolve. It should be noted that current guidelines still support tPA as first-line therapy even for those eligible for mechanical thrombectomy.

Patients not eligible for tPA or mechanical thrombectomy require hemodynamic optimization and neurologic monitoring. Admit such patients to the ICU stroke service for blood pressure management and frequent neurologic checks. Permissive hypertension allows for maximal cerebral perfusion. Systolic blood pressure >180 mmHg may require treatment, but the optimal mean arterial pressure goal is between 100 and 140 mmHg. Give normal saline solution without glucose (which could injure neurons in the penumbra due to osmotic fluid shift), and aim for normovolemia. A stroke patient who worsens clinically should undergo repeat head CT to evaluate for hemorrhage or increasing mass effect from swelling, which typically peaks 3 to 5 days after the stroke. Significant swelling from an MCA or cerebellar stroke may cause herniation and brain stem injury. A decompressive hemicraniectomy or suboccipital craniectomy can be a life-saving intervention for these select stroke patients. In studies of malignant MCA syndrome, decompressive hemicraniectomy showed favorable mortality and functional outcomes.

This treatment option, however, should be considered with the understanding of potentially poor functional recovery regardless of therapy. One study showed that less than half of malignant MCA patients who underwent decompressive hemicraniectomy returned home following rehabilitation, which is even fewer for those undergoing medical therapy.

Hemorrhagic Diseases

Intracranial hemorrhage from abnormal or diseased vascular structures accounts for approximately 15% of acute cerebrovascular events. Hypertension and amyloid angiopathy account for most intraparenchymal hemorrhages, although AVMs, aneurysms, venous thrombosis, tumors, hemorrhagic conversion of ischemic infarct, and fungal infections also may be the cause. The term intracranial hemorrhage is frequently used to signify intraparenchymal hemorrhage and will be used here. Intracranial hemorrhage causes local neuronal injury and dysfunction and also may cause global dysfunction due to mass effect if sufficiently large. The Intracerebral Hemorrhage (ICH Score) is commonly used to risk-stratify these patients, and it takes into account GCS, age, hemorrhage volume, presence of intraventricular hemorrhage, and location to predict mortality. AVM or aneurysm rupture (along with trauma, discussed previously in this chapter) result in subarachnoid hemorrhage (SAH) because the major cerebral and cortical blood vessels travel in the subarachnoid space, between the pia and the arachnoid membranes. SAH can cause immediate concussive-like neuronal dysfunction by exposure of the brain to intra-arterial pressure pulsations during the hemorrhage. Moreover, it can cause delayed ischemia from cerebral arterial vasospasm, which can present as acute worsening of the patient’s neurologic status days to weeks after the injury. Patients presenting with intracranial hemorrhages that do not follow typical patterns should undergo cerebral angiography or MRI to evaluate for possible underlying lesions, such as AVM or tumor.

Hemorrhagic stroke most commonly occurs within the basal ganglia or cerebellum. The patient is usually hypertensive on admission and has a history of poorly controlled hypertension. Such patients are more likely to present with lethargy or obtundation compared to those who suffer an ischemic stroke. Depressed mental status results from mass effect from the hematoma in deep structures, which can produce midline shift or herniation. Ischemic stroke does not cause mass effect acutely. Therefore, patients are more likely to present with normal consciousness and a focal neurologic deficit. Hemorrhagic strokes tend to present with a relatively gradual decline in neurologic function as the hematoma expands, rather than the immediately maximal symptoms caused by ischemic stroke. Table 42-3 provides a listing of relative incidences of intracranial hemorrhage by anatomic distribution.

Hypertension. Hypertension increases the relative risk of intracranial hemorrhage by approximately fourfold, likely due to chronic degenerative vasculopathy. Hypertensive hemorrhages often present in the basal ganglia, thalamus, or pons, and result from breakage of small perforating arteries that branch off of much larger parent vessels (Fig. 42-16).

Most hypertensive hemorrhages should be medically managed. The hematoma often contains intact, salvageable axons because the blood dissects through and along neural tracts, and surgical clot evacuation destroys these axons. Factors potentially favoring surgery include: superficial clot location, young age, nondominant hemisphere, rapid deterioration, and significant
Table 42-3

Anatomic distribution of intracranial hemorrhages and correlated symptoms

<table>
<thead>
<tr>
<th>% OF INTRACRANIAL HEMORRHAGES</th>
<th>LOCATION</th>
<th>CLASSIC SYMPTOMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>Basal ganglia (putamen, globus pallidus), internal capsule</td>
<td>Contralateral hemiparesis</td>
</tr>
<tr>
<td>15</td>
<td>Thalamus</td>
<td>Contralateral hemisensory loss</td>
</tr>
<tr>
<td>10–20</td>
<td>Cerebral white matter (lobar)</td>
<td>Depends on location (weakness, numbness, partial loss of visual field)</td>
</tr>
<tr>
<td>10–15</td>
<td>Pons</td>
<td>Hemiparesis; may be devastating</td>
</tr>
<tr>
<td>10</td>
<td>Cerebellum</td>
<td>Lethargy or coma due to brain stem compression and/or hydrocephalus</td>
</tr>
<tr>
<td>1–6</td>
<td>Brain stem (excluding pons)</td>
<td>Often devastating</td>
</tr>
</tbody>
</table>

mass effect. However, the most comprehensive randomized clinical trials to date did not show an overall improved outcome in surgically evacuated intracranial hemorrhage, except for the subgroup of patients with clot <1 cm from the cortical surface. More recent studies have assessed the role of minimally-invasive catheter evacuation of clot; these investigations are ongoing. Medical management remains the gold standard, however, and includes moderate blood pressure control, normalizing platelet and clotting function, phenytoin or levetiracetam for seizure prophylaxis, and electrolyte management. Intubate patients who cannot clearly follow commands to prevent aspiration and hypercarbia. Follow and document the neurologic examination and communicate with the family regarding appropriateness for rehabilitation vs. withdrawal of care.

**Amyloid Angiopathy.** The presence of pathologic amyloid deposition in the media of small cortical vessels compromises vessel integrity and tends to cause more superficial (lobar) hemorrhages than hypertensive intracranial hemorrhage. Amyloid laden vessels may hemorrhage multiple times. The superficial location of amyloid hemorrhages may make surgical evacuation less morbid compared to typical deep hypertensive hemorrhages. Nonetheless, medical management and family counseling should be approached similarly to patients with hypertensive hemorrhages.

**Figure 42-16.** A. Head computed tomography scan of a patient with left-sided weakness and progressive lethargy reveals a right basal ganglia hemorrhage (arrowhead). The blood clot is bright white. Hypodensity around the clot represents cerebral edema. There is blood within the ventricular system. B. Another patient with intraventricular extension of a basal ganglia hemorrhage. The patient developed right-sided weakness and then lethargy. Head computed tomography indicated hydrocephalus. A ventriculostomy was placed for cerebrospinal fluid drainage (arrowhead indicates cross-sectional view of the catheter entering the anterior horn of the right lateral ventricle).
The Hunt-Hess clinical grading system for subarachnoid hemorrhage (SAH) categorizes patients clinically (Table 42-5).

### Table 42-5
The Hunt-Hess clinical grading system for subarachnoid hemorrhage

<table>
<thead>
<tr>
<th>HUNT-HESS GRADE</th>
<th>CLINICAL PRESENTATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Asymptomatic; unruptured aneurysm</td>
</tr>
<tr>
<td>1</td>
<td>Awake; asymptomatic or mild headache; mild nuchal rigidity</td>
</tr>
<tr>
<td>2</td>
<td>Awake; moderate to severe headache, cranial nerve palsy (e.g., cranial nerve III or IV), nuchal rigidity</td>
</tr>
<tr>
<td>3</td>
<td>Lethargic; mild focal neurologic deficit (e.g., pronator drift)</td>
</tr>
<tr>
<td>4</td>
<td>Stuporous; significant neurologic deficit (e.g., hemiplegia)</td>
</tr>
<tr>
<td>5</td>
<td>Comatose; posturing</td>
</tr>
</tbody>
</table>

Cerebral Aneurysm. An aneurysm is a focal dilatation of the vessel wall and is most often a balloon-like outpouching, but may also be fusiform. Aneurysms usually occur at branch points of major vessels (e.g., internal carotid artery bifurcation), or at the origin of smaller vessels (e.g., posterior communicating artery or ophthalmic artery). Approximately 85% of aneurysms arise from the anterior circulation (carotid) and 15% from the posterior circulation (vertebrobasilar). Table 42-4 shows the percentage distribution of cerebral aneurysms by location. Aneurysms are thin walled and at risk for rupture. The major cerebral vessels, and therefore aneurysms, lie in the subarachnoid space. Rupture results in SAH. The aneurysmal tear may be small and seal quickly, or it may not. SAH may consist of a thin layer of blood in the CSF spaces, or thick layers of blood around the brain and extending into brain parenchyma, resulting in a clot with mass effect. Because the meningeal linings of the brain are sensitive with free nerve endings, SAH usually results in a sudden, severe “thunderclap” headache. A patient will classically describe “the worst headache of my life.” Presenting neurologic symptoms may range from mild headache to coma to sudden death. The Hunt-Hess grading system categorizes patients clinically (Table 42-5).

The World Federation of Neurological Surgery (WFNS) SAH Grading Scale is also used for this purpose. The Fisher Scale and a more recent modified form use head CT characteristics, described in the following section, to stratify patients based on risk of vasospasm-induced delayed cerebral ischemia. Both scales are used in conjunction as a quick way to communicate severity of aneurysmal SAH.

Patients with symptoms suspicious for SAH should have a head CT immediately. Acute SAH appears as a bright signal in the fissures and CSF cisterns around the base of the brain, as shown in Fig. 42-17. CT is rapid, noninvasive, and approximately 95% sensitive. In patients with suspicious symptoms but negative head CT, a lumbar puncture (LP) should be performed. An LP with xanthochromia and high red blood cell counts (usually 100,000/mL), which do not decrease between tubes 1 and 4, is consistent with SAH. Negative CT and LP essentially rules out SAH. Patients diagnosed with SAH require four-vessel cerebral angiography within 24 hours to assess for aneurysm or other vascular malformation. Catheter angiography remains the gold standard for assessing the patient’s cerebral vasculature, relevant anomalies, and presence, location, and morphology of the cerebral aneurysms. Figure 42-18A demonstrates the typical anteroposterior digital subtraction angiographic view of a cerebral aneurysm. Figure 42-18B shows the anatomy of the circle of Willis in a simplified graphic representation to assist in visualizing the locations of various cerebral aneurysms.

SAH patients should be admitted to the neurologic ICU. Hunt-Hess grade 4 and 5 patients require intubation and hemodynamic monitoring and stabilization. The current standard of care for ruptured aneurysms requires early aneurysmal occlusion. There are two options for occlusion. The patient may undergo craniotomy with microsurgical dissection and placement of a titanium clip across the aneurysm neck to exclude the aneurysm from the circulation and reconstitute the lumen of the parent vessel. The second option is to utilize an endovascular approach for treatment, which has traditionally taken the form of “coiling.” The patient is taken to the interventional neuroradiology suite for placement of looped titanium coils inside the aneurysm dome. The coils support thrombosis and prevent
options, as described earlier. Current guidelines favor endovascular therapy as the preferred first-line approach. The International Subarachnoid Aneurysm Trial researchers suggested that endovascular occlusion resulted in better mortality outcomes for certain types of cerebral aneurysms, although this trial was marred by poor selection and randomization techniques, and the validity of its conclusions have been questioned.\textsuperscript{63} Long-term outcomes may be better in younger patients with clipped aneurysms, as demonstrated in the Barrow Ruptured Aneurysm Trial (BRAT).\textsuperscript{64} However, this trial has a number of similar criticisms as well. Debate also continues regarding optimal care for unruptured intracranial aneurysms, with a recent large-scale study showing no clear benefit to open surgical vs. endovascular approaches.\textsuperscript{65}

SAH patients often require 1 to 3 weeks of ICU care after aneurysm occlusion for medical complications that accompany neurologic injury. In addition to routine ICU concerns, SAH patients are also at risk for cerebral vasospasm. In vasospasm, cerebral arteries constrict pathologically and can cause ischemia or stroke from 4 to 21 days after SAH. Current vasospasm prophylaxis includes maintenance of optimal perfusion with hypertension and mild hypervolemia, as well as administration of nimodipine, a calcium channel blocker that may decrease incidence and degree of spasm, though its mechanism is debated. Neurointerventional options for treating symptomatic vasospasm include intra-arterial papaverine or nicardipine, and balloon angioplasty for larger caliber vessels.

Aneurysmal SAH has an approximate mortality rate of 50\% in the first month. Approximately one-third of survivors return to pre-SAH function, and the remaining two-thirds have mild to severe disability. Most require rehabilitation after hospitalization.

Arteriovenous Malformations. AVMs are abnormal, dilated arteries and veins without an intervening capillary bed. The nidus of the AVM contains a tangled mass of vessels but no neural tissue. AVMs may be asymptomatic or present with SAH, intraparenchymal hemorrhage, or seizures. Small AVMs present with hemorrhage more often than large AVMs, which tend to present with seizures. Headache, bruit, or focal neurologic deficits are less common symptoms. AVMs hemorrhage at an average rate of 2\% to 4\% a year. Figure 42-19 demonstrates the angiographic appearance of an AVM in arterial and venous phases.

For unruptured AVMs, recent evidence from over two hundred patients supports medical management alone rather than intervention due to risk of stroke.\textsuperscript{66} Because AVM rupture can present radiographically as SAH, it is important to consider several management differences as compared to aneurysmal SAH. Definitive therapy for the AVM usually is delayed 3 to 4 weeks to allow the brain to recover from acute injury. There is less risk of devastating early rebleeding from AVMs, and vasospasm is much less common. Three therapeutic modalities for AVMs are currently in common use: microsurgical excision, interventional radiology or endovascular embolization, and stereotactic radiosurgery (SRS). AVMs that are large, near eloquent cortex, or that drain to deep venous structures are considered high grade and more difficult to surgically resect without causing a significant neurologic deficit. Radiosurgery can treat these lesions, although it is limited to lesions <3 cm in diameter and has a 2-year lag time (i.e., the AVM may bleed in the interval). Embolization reduces flow through the AVM. It is usually considered adjunctive therapy, but it may serve as the sole treatment for deep, inaccessible lesions.
TUMORS OF THE CENTRAL NERVOUS SYSTEM

A wide variety of tumors affect the brain and spine. Primary benign and malignant tumors arise from the various elements of the CNS, including neurons, glia, and meninges. Tumors metastasize to the CNS from many primary sources. Presentation varies widely depending on relevant neuroanatomy. Prognosis depends on histology and anatomy. Modern brain tumor centers use team approaches to CNS tumors, as patients may require a combination of surgery (including newer, more minimally invasive approaches), radiation therapy, chemotherapy, SRS, and research protocol enrollment for studies assessing the efficacy of newer approaches such as immunotherapy. Tumors affecting the peripheral nervous system are discussed in the “Peripheral Nerve” section.

Intracranial Tumors

Intracranial tumors can cause brain injury from mass effect, dysfunction or destruction of adjacent neural structures, swelling, or abnormal electrical activity (seizures). Supratentorial tumors commonly present with focal neurologic deficits, such as contralateral limb weakness, visual field deficit, headache, or seizure. Infratentorial tumors often cause increased ICP due to hydrocephalus from compression of the fourth ventricle, leading to headache, nausea, vomiting, or diplopia. Cerebellar hemisphere or brain stem dysfunction can result in ataxia, nystagmus, or cranial nerve palsies. Infratentorial tumors rarely cause seizures.

All patients with symptoms concerning for brain tumor should undergo MRI with and without gadolinium. Gadolinium-based contrast can identify locations of blood-brain barrier breakdown of tumors and, when used in conjunction with other MRI sequences, is essential in narrowing the differential diagnosis. Initial management of a patient with a symptomatic brain tumor generally includes dexamethasone for reduction of vasogenic edema, and phenytoin or levetiracetam if the patient has seized. Patients with significant weakness, lethargy, or hydrocephalus should be admitted for observation until definitive care is administered.

Metastatic Tumors

Cerebral metastases are the most common type of intracranial tumor. Prolonged cancer patient survival and improved CNS imaging have increased the likelihood of diagnosing cerebral metastases. The sources of most cerebral metastases are (in decreasing frequency): lung, breast, kidney, GI tract, and melanoma. Lung and breast cancers account for more than half of cerebral metastases. Metastatic cells usually travel to the brain hematogenously and frequently seed the gray-white junction due to characteristic blood vessel caliber change. Other common locations are the cerebellum and the meninges. Meningeal involvement may result in carcinomatous meningitis, also known as leptomeningeal carcinomatosis. MRI pre- and post-contrast administration is the study of choice for evaluation. Figure 42-20 demonstrates bilateral cerebellar metastases. These lesions are typically well circumscribed, round, and multiple. Such findings should prompt a metastatic work-up, including CT scan of the chest, abdomen, and pelvis, and a bone scan.

Management largely depends upon the primary tumor, overall tumor burden, patient’s medical condition, and location and number of metastases. The beliefs of the patient and family regarding aggressive care must be considered, with the primary goal of optimizing survival time while maintaining or improving neurological function. Neurosurgical intervention can be indicated for a number of reasons. Biopsy can be obtained to provide a tissue diagnosis and further direct therapy. Hydrocephalus from increased intracranial pressure due to intracranial tumor burden can be temporized via placement of a ventriculoperitoneal shunt. Craniotomy can be used for debulking of intracranial tumor burden or resection. Data from randomized controlled trials have supported the use of craniotomy for tumor resection plus whole-brain radiation therapy (WBRT) or stereotactic radiosurgery (SRS) for patients with a single surgically accessible metastatic lesion, compared to radiation therapy alone. In one randomized trial assessing surgery and WBRT vs. WBRT alone, local recurrence decreased and median...
survival increased from 15 to 40 weeks. For multiple metastases, it should be noted that craniotomy primarily for resection is typically not indicated unless all detectable metastases can be resected. It may however still be useful for symptomatic relief from a primary lesion. Recent data suggest that SRS (e.g., Gamma Knife) may be applied to multiple metastases in one session with improved outcome.

Glial Tumors

Glial cells provide the anatomic and physiologic support for neurons and their processes in the brain. Tumors arising from glial cells are termed gliomas, and they represent the most common primary brain tumor. The several types of glial cells give rise to distinct primary CNS neoplasms.

Astrocytoma. Astrocytoma is the most common primary CNS neoplasm. The term glioma often is used to refer to astrocytomas specifically, excluding other glial tumors. Astrocytomas are graded from I to IV. Grades I and II are referred to as low-grade astrocytoma or low grade glioma, grade III as anaplastic astrocytoma, and grade IV as glioblastoma multiforme (GBM). Prognosis varies significantly between grades I/II, III, and IV, but not between I and II. Median survival is 8 years after diagnosis with a low-grade tumor, 2 to 3 years with anaplastic astrocytoma, and roughly 1 year with a GBM. GBMs account for almost two-thirds of all astrocytomas, anaplastic astrocytomas account for two-thirds of the rest, and low-grade astrocytomas the remainder. Fig. 42-21 demonstrates the typical appearance of a GBM.

The great majority of astrocytomas infiltrate adjacent brain. Juvenile pilocytic astrocytomas and pleomorphic xanthoastrocytomas are exceptions. These tumors are circumscribed, low grade, and associated with a good prognosis. Histologic features associated with higher grade include hypercellularity, nuclear atypia, and endovascular hyperplasia. Necrosis is present only with GBMs; it is required for the diagnosis.

Gross total resection should be attempted for suspected astrocytomas. Motor cortex, language centers, deep or midline structures, or brainstem location may make this impossible without unacceptable, devastating neurologic deficit. Advanced imaging, such as diffusion tensor imaging (DTI) and functional MRI (fMRI), are seeing increased use as means of assessing peritumoral structure and function to guide surgical decision-making. However, some lesions may be in such precarious regions as to be limited to stereotactic needle biopsy specimen. Gross total resection followed by fractioned radiotherapy (FRT) improves survival for all grades, although radiation therapy may be delayed until recurrence in low-grade tumors. Alongside FRT, adjuvant chemotherapy with temozolomide was demonstrated in a randomized controlled trial to increase short-term survival rate. Bevacizumab, an anti-VEGF antibody, is another treatment option under investigation. There are various ongoing research studies for GBM adjuvant therapy; these should be discussed with the patient and family. Other options include Iotrex-containing balloons for conformal radiation brachytherapy (Glia-Site), placed in the resection cavity at the time of surgery.
for recurrence. Adjuvant therapy remains marginally effective; survival has changed little over the last several decades.

**Oligodendroglioma.** Oligodendroglioma accounts for approximately 10% of gliomas, arising from the oligodendrocytes that create myelin in the CNS. They often present with seizures. Calcifications and hemorrhage on CT or MRI suggest the diagnosis. Oligodendrogliomas are also graded from I to IV; grade portends prognosis. Prognosis is better overall than for astrocytomas. Median survival ranges from 2 to 7 years for highest and lowest grade tumors, respectively. Aggressive resection improves survival. Many oligodendrogliomas will respond to procarbazine, lomustine (CCNU), and vincristine (PCV) chemotherapy. A particular chromosomal deletion, 1p19q, has been associated with robust response to the chemotherapeutic agent temozolomide. Radiation has not been clearly shown to prolong survival.

Recent updates to brain tumor classification by the WHO (discussed in the following section) note that high-grade (at least II or III) oligodendrogliomas and astrocytomas are classified together as diffuse gliomas. In fact, more similarity is seen between high-grade astrocytomas and oligodendrogliomas than between high-grade astrocytomas and low-grade astrocytomas. Further discussion of this nosology is beyond the scope of this chapter, but implications for the future of neuro-oncology are discussed here.

**Ependymoma.** The lining of the ventricular system consists of cuboidal/columnar ependymal cells from which ependymomas may arise. Although most pediatric ependymomas are supratentorial, two-thirds of adult ependymomas are infratentorial. Supratentorial ependymomas arise from the lateral or third ventricles. The infratentorial tumors arise from the floor of the fourth ventricle (i.e., off the posterior brainstem). The most common symptoms are headache, nausea, vomiting, or vertigo, secondary to increased ICP from obstruction of CSF flow through the fourth ventricle. The tumors may grow out the foramina of Luschka to form a cerebellopontine angle mass. They may also spread through the CSF to form “drop mets” in the spinal canal. The two main histologic subtypes are papillary and anaplastic, the latter characterized by increased mitotic activity and areas of necrosis. Gross total resection often is impossible because the tumor arises from the brain stem. The goal of surgery is to achieve maximal resection without injuring the very delicate brainstem. Suboccipital craniotomy and midline separation of the cerebellar hemispheres allows access to tumors in the fourth ventricle. Postoperative radiation therapy significantly improves survival. Patients with CSF spread documented by LP or contrast MRI should also have whole-spine radiation plus focused doses to visualized metastases.

**Choroid Plexus Papilloma.** The choroid plexus is composed of many small vascular tufts covered with cuboidal epithelium. It represents part of the interface between blood and brain. The choroid cells create CSF from blood via ultrafiltration and release it into the ventricular system. Choroid plexus papillomas and choroid plexus carcinomas (rare, mostly pediatric) may arise from these cells. Papillomas usually occur in infants (typically supratentorial in the lateral ventricle) but also occur in adults (usually infratentorial in the fourth ventricle). Papillomas
are well circumscribed and vividly enhance due to extensive vasculature. Like ependymomas, adult choroid plexus papillomas usually present with symptoms of increased ICP. Treatment is surgical excision. Total surgical excision is curative; recurrent papillomas should be re-resected. Radiation and chemotherapy are not indicated for papillomas. Radiation is adjunctive to aggressive surgery for carcinomas, but the results are generally poor.

**Neural Tumors and Mixed Tumors**

Neural and mixed tumors are a diverse group that includes tumors variously containing normal or abnormal neurons and/or normal or abnormal glial cells. Primitive neuroectodermal tumors arise from bipotential cells, capable of differentiating into neurons or glial cells.

**Medulloblastoma.** Medulloblastoma is classically described as the most common type of primitive neuroectodermal tumor (PNET), although this term has been removed in the latest WHO classification of central nervous system tumors. Most occur in the first decade of life, but there is a second peak around age 30. Medulloblastoma is the most common malignant pediatric brain tumor. They are usually midline. Most occur in the cerebellum and present with symptoms of increased ICP. Histologic characteristics include densely packed small round cells with large nuclei and scant cytoplasm. They are generally not encapsulated, frequently disseminate within the CNS, and should undergo surgical resection followed by radiation therapy and chemotherapy.

**Ganglioglioma.** Ganglioglioma is a mixed tumor in which both neurons and glial cells are neoplastic. They occur in the first three decades of life, often in the medial temporal lobe, as circumscribed masses that may contain cysts or calcium and may enhance. The presenting symptom is usually a seizure, due to the medial temporal location. Patients have a good prognosis after complete surgical resection.

**Neural Crest Tumors**

Multipotent neural crest cells develop into a variety of disparate cell types, including smooth muscle cells, sympathetic and parasympathetic neurons, melanocytes, Schwann cells, and arachnoid cap cells. They migrate in early development from the primitive neural tube throughout the body.

**Miscellaneous Tumors**

**Meningioma.** Meningiomas are derived from arachnoid cap cells of the arachnoid mater. They appear to arise from the dura mater grossly and on MRI and are commonly referred to as dural-based tumors. The most common intracranial locations are along the falx (Fig. 42-22), the convexities (i.e., over the cerebral hemispheres), and the sphenoid wing. Less common locations include the foramen magnum, olfactory groove, and inside the lateral ventricle. Most are slow growing, encapsulated, benign tumors. Aggressive atypical or malignant meningiomas may invade adjacent bone or cerebral cortex. Previous cranial irradiation increases the incidence of meningiomas. Approximately 10% of patients with a meningioma have multiple meningiomas. Total resection is curative, although involvement with small perforating arteries or cranial nerves may make total resection of skull base tumors impossible without significant neurologic deficit. The Simpson grading scale is used to characterize the extent of resection. Small, asymptomatic meningiomas can be followed until symptomatic or until significant growth is documented on serial imaging studies.

Atypical and malignant meningiomas may require postoperative radiation. Patients may develop recurrences from the surgical bed or distant de novo tumors.

**Vestibular Schwannoma (Acoustic Neuroma).** Vestibular schwannomas predominantly arise from the superior half of the vestibular portion of the vestibulocochlear nerve (cranial nerve VIII) (Fig. 42-23). Commonly, patients present with progressive hearing loss, tinnitus, or balance difficulty. Large tumors may cause brain stem compression and obstructive hydrocephalus. Bilateral acoustic neuromas are pathognomonic for neurofibromatosis type 2 (NF2), a syndrome resulting from mutation of chromosome 22. NF2 patients have an increased incidence of spinal and cranial meningiomas and gliomas.

Vestibular schwannomas may be treated with microsurgical resection or SRS (Gamma Knife radiosurgery or linear accelerator technology). The main complication with treatment is damage to the facial nerve (cranial nerve VII), which runs through the internal auditory canal with the vestibulocochlear nerve. Risk of facial nerve dysfunction increases with increasing tumor diameter. SRS is preferred for tumors <3 cm, and microsurgical resection for those >3 cm (SRS can be supplemented for any residual tumor following resection).

**Pituitary Adenoma.** Pituitary adenomas arise from the anterior pituitary gland (adenohypophysis). Tumors <1 cm diameter are considered microadenomas; larger tumors are macroadenomas.
Pituitary tumors may be functional (i.e., secrete endocrinologically active compounds at pathologic levels) or nonfunctional (i.e., secrete nothing or inactive compounds). Functional tumors are often diagnosed when quite small, due to endocrine dysfunction. The most common endocrine syndromes are Cushing’s disease, due to adrenocorticotropic hormone secretion, Forbes-Albright syndrome, due to prolactin secretion, and acromegaly, due to growth hormone secretion. Nonfunctional tumors are typically diagnosed as larger lesions causing mass effects such as visual field deficits due to compression of the optic chiasm or panhypopituitarism due to compression of the gland. Figure 42-24 demonstrates a large pituitary adenoma. Hemorrhage into a pituitary tumor causes abrupt symptoms of headache, visual disturbance, decreased mental status, and endocrine dysfunction. This is known as pituitary apoplexy.

Symptomatic pituitary tumors should be decompressed surgically to eliminate mass effect and/or to attempt an endocrine cure. However, prolactin-secreting tumors (prolactinomas) usually shrink with dopaminergic therapy alone. First-line pharmacotherapy for small prolactinomas is cabergoline, a dopamine agonist that inhibits production and secretion of prolactin, and is preferred over bromocriptine for its superior side effect profile. Consider surgery for prolactinomas with persistent mass effect or endocrinologic dysfunction in spite of adequate dopamine agonist therapy. Most pituitary tumors are approached transsphenally via the transsphenoidal approach, and minimally invasive, endoscopic surgical techniques are being used increasingly.

**Hemangioblastoma.** Hemangioblastomas occur almost exclusively in the posterior fossa, with about 20% occurring in patients with von Hippel-Lindau (VHL) disease, a multisystem neoplastic disorder. Other tumors associated with VHL are renal cell carcinoma, pheochromocytoma, and retinal angiomas. Many appear as cystic tumors with an enhancing tumor on the cyst wall known as the mural nodule. Surgical resection is curative for sporadic (non-VHL associated) tumors. Pathology reveals abundant thin-walled vascular channels; internal debulking may be bloody. En bloc resection of the mural nodule alone, leaving the cyst wall, is sufficient.

**Lymphoma.** CNS lymphoma may arise either primarily in the CNS or secondarily from systemic disease. Recent rising incidence may be due to growing transplant and AIDS populations. Presenting symptoms include mental status changes, headache due to increased ICP, and cranial nerve palsy due to lymphomatous meningitis (analogous to carcinomatous meningitis). Often, lymphoma appears hyperdense on CT due to dense cellularity. Most lesions typically enhance with contrast on MRI, and can be differentiated by diffusion restriction on diffusion-weighted sequences. Surgical excision is not indicated. A stereotactic needle biopsy specimen usually confirms the diagnosis. Subsequent treatment includes steroids, whole-brain radiation, and chemotherapy. Intrathecal methotrexate is an additional treatment option.

**Embryologic Tumors**

Embryologic tumors result from embryonal remnants that fail to involute completely or differentiate properly during development.

**Craniopharyngioma.** Craniopharyngiomas are benign cystic lesions that occur in the sellar region that occur most frequently in children. A second peak of incidence also exists around 50 years of age. Craniopharyngiomas arise from remnant embryonic tissue in the pituitary stalk. Calcification occurs in all pediatric and roughly half of adult craniopharyngiomas. Symptoms result from compression of adjacent structures, especially the optic chiasm. Pituitary or hypothalamic dysfunction or hydrocephalus may develop. Treatment is primarily surgical. Excision
is somewhat easier in children, as the tumor is often soft and easily suctioned. Adult tumors are often firm and adherent to adjacent vital structures. Visual loss, pituitary endocrine hypo-function, diabetes insipidus, and cognitive impairment from basal frontal injury are common complications.

**Epidermoid.** Epidermoid tumors are cystic lesions with stratified squamous epithelial walls from trapped ectodermal cell rests that grow slowly and linearly by desquamation into the cyst cavity. The cysts contain keratin, cholesterol, and cellular debris (Fig. 42-25). They occur most frequently in the cerebellopontine angle and may cause symptoms due to brainstem compression. Recurrent bouts of aseptic meningitis may occur due to release of irritative cyst contents into the subarachnoid space (Mollaret’s meningitis). Treatment is surgical drainage and removal of the cyst wall. Intraoperative spillage of cyst contents may lead to severe chemical meningitis and must be avoided by containment and aspiration.

**Dermoid.** Dermoids are less common than epidermoid tumors. They contain hair follicles and sebaceous glands in addition to a squamous epithelium. Dermoids may be found anywhere along the craniospinal axis. They are more commonly midline structures and are associated with more anomalies than epidermoids. They may be associated with trauma, as from a lumbar puncture that drags skin structures into the spine. Bacterial meningitis may occur when dermoids are associated with a dermal sinus tract. Treatment of symptomatic lesions is surgical resection, again with care to control cyst contents.

**Teratoma.** Teratomas are germ cell tumors that arise in the midline, often in the pineal region (the area behind the third ventricle, above the midbrain and cerebellum). They contain elements from all three embryonal layers: ectoderm, mesoderm, and endoderm. Teratomas may contain skin, cartilage, GI glands, and teeth. Teratomas with more primitive features are more malignant, while those with more differentiated tissues are more benign. Surgical excision may be attempted. However, the prognosis for malignant teratomas is very poor.

**Spinal Tumors**

Approximately 20% of CNS tumors occur in the spine, and a wide variety of spinal tumors exist. Unlike cranial tumors, the majority of spinal tumors are histologically benign. Understanding two major spinal concepts—stability and neural compression—facilitates an understanding of the effects of spinal tumors. Destruction of bones or ligaments can cause spinal instability, leading to deformities such as compression, kyphosis, subluxation, all of which harbor the potential for subsequent neural compression. Tumor growth in the spinal canal or neural foramina can cause direct compression of the spinal cord or nerve roots and cause pain and loss of function. Classically, the pain is worse at night. Anatomic categorization provides the most logical approach to these tumors. Certain tumors present in characteristic locations. An understanding of the anatomy leads to an understanding of the clinical presentation and possible therapeutic options.

**Extradural Tumors.** Extradural tumors account for approximately 55% of spinal tumors. This category includes tumors...
arising within the bony vertebral structures and from within the epidural space. Destruction of the bone can lead to instability and fractures, causing pain and/or deformity. Epidural expansion can lead to spinal cord or nerve root compression with myelopathy, radiculopathy, or a combination thereof.

**Metastatic Tumors** Metastatic tumors are the most common extradural tumors. Spinal metastases most commonly occur in the thoracic and lumbar vertebral bodies because the greatest volume of red bone marrow is found in these regions. The most common primary sources of spine metastases are lymphoma, lung, breast, and prostate. Other sources include renal, colon, thyroid, sarcoma, and melanoma. Most spinal metastases create osteolytic lesions. Osteoblastic, sclerotic lesions suggest prostatic cancer in men and breast cancer in women.

Patients with progressive neurologic dysfunction due to a metastatic lesion should undergo urgent surgery followed by radiation therapy. Patients with debilitating pain may undergo radiation therapy with close observation for neurologic deterioration. Preoperative neurologic function correlates with postoperative function. Patients may lose function over hours. These patients should be given high-dose IV dexamethasone, taken immediately to MRI, and then to the OR or radiation therapy suite. Indications for surgery include failure of radiation therapy, spinal instability, recurrence after radiation therapy, and the need for diagnosis in cases of unknown primary tumors. Most cases with significant bone involvement require both decompression and fusion. Bony fusion usually takes 2 to 3 months. Prognosis governs operative decisions. Surgery is unlikely to improve quality of life for patients with a life expectancy of 3 months or less, but it is likely to improve quality of life for patients with life expectancy of 6 months or more. Benefit for patients with 3- to 6-months’ life expectancy is unclear and requires frank discussion with the patient and family. Patients who are unlikely to tolerate general anesthesia, are already completely paralyzed, or who have very radiosensitive tumors such as multiple myeloma and lymphoma, should not generally undergo surgery.

**Management Principles of Spinal Cord Compression in Metastatic Cancer of the Spine** Spinal cord compression due to tumor burden is important to distinguish because it can, as with any other form of cord compression, cause paralysis and loss of bowel and bladder function. A randomized controlled trial demonstrated that patients with spinal cord compression from metastatic tumor have better outcomes with decompressive surgery and radiotherapy compared to radiotherapy alone. Patients with radiosensitive tumors, such as multiple myeloma and lymphoma, are excluded from this group. Another important tool in assessing these patients is the SINS (Spinal Instability Neoplastic Score). This scale grades the utility of operative intervention in metastatic cancer to the spine on the basis of pain, deformity, location, type of bone lesion, integrity of vertebral body, and posterior element involvement.

**Primary Tumors** Hemangiomas are benign tumors found in 10% of people at autopsy. They occur in the vertebral bodies of the thoracolumbar spine and are frequently asymptomatic. They are often vascular and may hemorrhage, causing pain or neurologic deficit. Large hemangiomas can destabilize the spine and predispose to fracture. Osteoblastic lesions include osteoid osteoma and osteoblastoma. The latter tends to be larger and more destructive. Aneurysmal bone cysts are nonneoplastic, expansile, lytic lesions containing thin-walled blood cavities that usually occur in the lamina or spinous processes of the cervicothoracic spine. They may cause pain or sufficiently weaken the bone to cause a fracture. Cancers arising primarily in the bony spine include Ewing’s sarcoma, osteosarcoma, chondrosarcoma, and plasmacytoma.

**Intradural Extramedullary Tumors** Intradural extramedullary tumors constitute approximately 40% of spinal tumors and arise from the meninges or nerve root elements. They may compress the spinal cord, causing myelopathy, or the nerve roots, causing radiculopathy. The most common intradural extramedullary tumors are typically benign, slow growing, and well circumscribed. Rare benign epidural masses include arachnoid cysts, dermoids, and epidermoids. Rare malignant epidural tumors include metastases and high-grade gliomas, or “drop” metastases from posterior fossa gliomas.

**Meningioma** Meningiomas arise from the arachnoidea mater. They appear to be dural based and enhance on MRI. An enhancing “dural tail” may be seen. They occur most commonly in the thoracic spine (Fig. 42-26) but also arise in the cervical and lumbar regions. Some spinal meningiomas grow into the

---

**Figure 42-26.** T2-weighted sagittal magnetic resonance imaging of the midthoracic spine demonstrating a well-encapsulated tumor arising from the dura posteriorly and compressing the spinal cord. Arrowhead points to dorsal location of the mass. The patient presented with worsening gait and lower extremity spasticity. Pathology demonstrated meningioma.
epidural space. Growth causes cord compression and progressive myelopathy with hyperreflexia, spasticity, and gait difficulties. Surgical excision is the treatment of choice. The surgeon often finds a clean margin between the tumor, dura, and spinal cord, allowing en bloc resection without damage to the cord.

**Schwannoma** Schwannomas are derived from peripheral nerve sheath Schwann cells. They are benign, encapsulated tumors that rarely undergo malignant degeneration. While two-thirds are entirely intradural, one-sixth are entirely extradural, and one-sixth have a classic “dumbbell” shape from intradural and extradural components. Symptoms result from radiculopathy, often presenting as pain or myelopathy. Symptomatic lesions should be surgically resected. The parent nerve root usually can be preserved. Patients with multiple schwannomas likely have NF2. In these patients, a careful neurologic examination is needed to determine which lesions are symptomatic and require resection.

**Neurofibroma** In contrast to schwannomas, neurofibromas tend to appear more fusiform and to grow within the parent nerve, rather than forming an encapsulated mass branching off the nerve. Neurofibromas are benign but not encapsulated. They present similarly to schwannomas, and the two may be difficult to differentiate on imaging. Salvage of the parent nerve is more challenging with neurofibromas. To improve the likelihood of total resection, thoracic and high cervical nerve roots may be sacrificed with minimal deficit. Patients with multiple neurofibromas likely have NFI, also known as von Recklinghausen’s neurofibromatosis. Resection for symptomatic lesions should be offered.

**Intradural Tumors.** Intradural tumors constitute approximately 5% of spinal tumors. They arise from within the parenchyma of the spinal cord. Common presenting symptoms are local dysesthesia, burning pain, radicular pain, sensory loss, weakness, or sphincter dysfunction. Patients with such symptoms should undergo MRI of the entire spine with and without enhancement.

**Ependymoma** Ependymomas are the most common intramedullary tumors in adults. There are several histologic variants. The myxopapillary type occurs in the conus medullaris or the filum terminale in the lumbar region and has the best prognosis after resection. The cellular type occurs more frequently in the cervical cord. Many spinal ependymomas have cystic areas and may contain hemorrhage. Surgical removal can improve function. A distinct tumor margin often exists, allowing safer excision. Postoperative radiation therapy after subtotal resection may prolong disease control.

**Astrocytoma** Astrocytomas are the most common intramedullary tumors in children, although they also occur in adults. They may occur at all levels, although more often in the cervical cord. The tumor may interfere with the CSF-containing central canal of the spinal cord, leading to a dilated central canal, referred to as syringomyelia (syrinx). Spinal astrocytomas are usually low grade, but complete excision is rarely possible due to the nonencapsulated, infiltrative nature of the tumor. As a result, patients with astrocytomas fare worse overall than patients with ependymomas.

**Other Tumors.** Other types of rare tumors include high-grade astrocytomas, dermoids, epidermoids, teratomas, hemangiomas, hemangioblastomas, and metastases. Patients usually present with pain. Prognosis generally depends on preoperative function and the histologic characteristics of the lesion.

### Future Directions

Future directions in the neurosurgical management of brain tumors are related to improved genetic characterization of brain tumors and technological advances. Regarding the former, the recent update of the World Health Organization classification of central nervous system tumors, alluded to earlier in this chapter, emphasizes an integrated classification approach utilizing both histologic and molecular characteristics. The addition of the latter feature underscores significant advances in the genetics of CNS tumors, which are being translated to treatment. Immunotherapy, for example, is an active area of research in the treatment of GBM. In this approach, immune cells such as T cells and dendritic cells are leveraged to target tumor-specific tissue. These approaches will play an important role as adjuvant therapy to neurosurgical approaches.

Recent advances in neurosurgical technology are being used to address previous operative limitations. Intraoperative fluorescein is being studied as a means of marking abnormal tissue that appears grossly normal in order to maximize resection and minimize recurrence. Indications for SRS are expanding and provide a noninvasive option that can be used as monotherapy or as supplemental therapy to operative intervention. Moreover, new minimally invasive techniques, such as MRI-guided laser interstitial thermal therapy (MRgLITT), are actively being studied for less morbid access to otherwise difficult-to-reach tumors.

### SPINE: BASIC CONCEPTS

The spine is a complex structure and is subject to an extensive array of pathologic processes, including degeneration, inflammation, infection, neoplasia, and trauma. Discussions of spine trauma, tumor, and infection are addressed separately in this chapter in the “Infection—Spine,” “Spinal Tumors,” and “Spine Trauma” sections. General concepts, common patterns of disease, and basic operative interventions are presented here.

The spine consists of a series of stacked vertebrae, intervening discs, and longitudinal ligaments. The vertebrae consist of the vertebral body anteriorly and the pedicles, articular facets, laminae, and spinous processes posteriorly. The intervertebral discs have two components. The tough, fibrous ring that runs around the outer diameter of the two adjacent vertebral bodies is known as the annulus fibrosus. The spongy material inside the ring of the annulus is known as the nucleus pulposus. The annulus and the nucleus provide a cushion between adjacent vertebral bodies, absorb forces transmitted to the spine, and allow some movement between the vertebral bodies. The ligaments stabilize the spine by limiting the motion of adjacent vertebrae.

**Stability**

Stability and neural compression are the two concepts critical to understanding the mechanics and pathologic processes affecting the spine.

**Future Directions**

Future directions in the neurosurgical management of brain tumors are related to improved genetic characterization of brain tumors and technological advances. Regarding the former, the recent update of the World Health Organization classification of central nervous system tumors, alluded to earlier in this chapter, emphasizes an integrated classification approach utilizing both histologic and molecular characteristics. The addition of the latter feature underscores significant advances in the genetics of CNS tumors, which are being translated to treatment. Immunotherapy, for example, is an active area of research in the treatment of GBM. In this approach, immune cells such as T cells and dendritic cells are leveraged to target tumor-specific tissue. These approaches will play an important role as adjuvant therapy to neurosurgical approaches.
intermediate mobility. The sacral spine is fused together and has no intrinsic mobility. The load borne by the lumbar spine is transmitted to the sacrum, and then the pelvis through the sacroiliac joints. The coccyx is the most inferior segment of the spine and has no significant contribution to load bearing or mobility.

A stable spine is one that can bear normally experienced forces resulting from body mass, movement, and muscle contraction, while maintaining normal structure and alignment. An unstable spine will shift or sublux under these forces. The determinants of spinal stability vary throughout the cervical, thoracic, and lumbar portions. In elementary form, stability depends on the structural integrity of the hard, bony elements of the vertebral column, as well as the tensile integrity and security of the supporting ligamentous attachments. Plain X-rays and CT scans are sensitive for detecting bony defects such as fractures or subluxation, while MRI better detects disruptions of the soft tissues, including ligaments and intervertebral discs. Specific patterns of abnormalities seen on imaging studies may suggest or diagnose spinal instability.

A common form of nontraumatic spinal instability is lumbar spondylolisthesis, which is typically a forward slippage of a lumbar vertebra relative to the lower vertebra on which it rests. This results from congenital or degenerative disruption of the pars interarticularis, the critical bridge of bone that spans adjacent facet joints. In the setting of a pars defect, there is no solid bony connection between the adjacent vertebrae. The spine is unstable and anterior listhesis (slippage) may result. Patients typically present with severe low back pain that is exacerbated with movement and load bearing (mechanical low back pain). Radiculopathy in this setting indicates neuroforaminal compression. Figure 42-27 demonstrates an L4 and L5 spondylolisthesis.

Neural Compression

Besides providing a stable, central element of the body’s support structure, the spine also protects the spinal cord and nerve roots as they pass through the neural foramina to form the peripheral nervous system. In a healthy spine, the spinal cord and nerve roots are suspended in CSF, free of mechanical compression. Pathologic processes that can lead to CSF space impingement and neural compression include: hypertrophic degenerative changes in the intervertebral discs and facet joints, expansion of epidural masses such as tumors or abscesses, and subluxation (i.e., slippage) of adjacent vertebral bodies. Subluxation may be due to trauma that exceeds the spine’s load-bearing capabilities and leads to structural failure, or chronic structural degradation by degenerative disease, infection, or tumor. Subluxation reduces the cross-sectional area of the central canal and the neural foramina (see Fig. 42-10B). Reduced central canal area can lead to myelopathy. Reduced neural foraminal area can lead to radiculopathy.

Myelopathy. Compression of the spinal cord can cause disturbance of function known as myelopathy. This dysfunction may be secondary to the direct effects of compression, cord ischemia due to reduced perfusion, or pathologic changes due to repeated cord trauma. These mechanisms lead to demyelination of the corticospinal tracts, which are long descending motor tracts. Corticospinal tract damage leads to upper motor neuron signs and symptoms, including hyperreflexia, spasticity, and weakness. These mechanisms also cause damage to the dorsal columns, which carry ascending proprioception, vibration, and two-point discrimination information. Loss of proprioception makes fine motor tasks and ambulation difficult.

Radiculopathy. Compression of the nerve roots causes disturbance of root function, known as radiculopathy. Characteristic features of radiculopathy include lower motor neuron signs and symptoms (hyporeflexia, atrophy, and weakness) and sensory disturbances such as numbness or tingling sensations (paresthesias), burning sensations (dysesthesias), and shooting (radicular) pain. Myelopathy and radiculopathy often present together in diseases that involve the central canal and the neural foramina. This combination can lead to lower motor neuron dysfunction at the level of disease, and upper motor neuron dysfunction below that level.

Patterns of Disease

Cervical Radiculopathy. The cervical nerve roots exit the central canal above the pedicle of the same-numbered vertebra and at the level of the higher adjacent intervertebral disc. For example, the C6 nerve root passes above the C6 pedicle at the level of the C5–C6 discs. The cervical nerve roots may be compressed acutely by disc herniation, or chronically by hypertrophic degenerative changes of the discs, facets, and ligaments. Table 42-6 summarizes the effects of various disc herniations. Most patients with acute disc herniations will improve without surgery. NSAIDs or cervical traction may help alleviate symptoms. Patients whose symptoms do not resolve or who have significant weakness should undergo decompressive surgery. The two main options for nerve root decompression are anterior cervical discectomy and fusion (ACDF) and posterior cervical foraminotomy (keyhole foraminotomy). ACDF allows more direct access to and removal of the pathology (anterior to the nerve root). However, the procedure requires fusion because discectomy causes a collapse of the interbody space and instability will likely occur. Figure 42-28 demonstrates a C6–C7 ACDF with the typical interposed graft and plating system. Keyhole foraminotomy allows for decompression without requiring
fusion, but it is less effective for removing centrally located canal pathology.

**Cervical Spondylotic Myelopathy.** The term *spondylosis* refers to diffuse degenerative and hypertrophic changes of the discs, intervertebral joints, and ligaments, which collectively result in spinal stenosis. Spinal cord dysfunction (myelopathy) due to cord compression from cervical spinal degenerative disease is therefore referred to as *cervical spondylotic myelopathy* (CSM). Classically CSM presents with spasticity and hyperreflexia due to corticospinal tract dysfunction, upper extremity weakness and atrophy from degeneration of the motor neurons in the anterior horns of the spinal gray matter, and loss of lower extremity proprioception due to dorsal column injury.

Table 42-6

<table>
<thead>
<tr>
<th>LEVEL</th>
<th>FREQUENCY (%)</th>
<th>ROOT INJURED</th>
<th>REFLEX</th>
<th>WEAKNESS</th>
<th>NUMBNESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>C4–C5</td>
<td>2</td>
<td>C5</td>
<td>—</td>
<td>Deltoid</td>
<td>Shoulder</td>
</tr>
<tr>
<td>C5–C6</td>
<td>19</td>
<td>C6</td>
<td>Biceps</td>
<td>Biceps brachii</td>
<td>Thumb</td>
</tr>
<tr>
<td>C6–C7</td>
<td>69</td>
<td>C7</td>
<td>Triceps</td>
<td>Wrist extensors (wrist drop)</td>
<td>Second and third digits</td>
</tr>
<tr>
<td>C7–T1</td>
<td>10</td>
<td>C8</td>
<td>—</td>
<td>Hand intrinsics</td>
<td>Fourth and fifth digits</td>
</tr>
</tbody>
</table>


Figure 42-28. **A.** Anteroposterior cervical spine X-ray showing the position of an anterior cervical plate used for stabilization after C6–C7 discectomy. Patient presented with right triceps weakness and dysesthesias in the right fifth digit. Magnetic resonance imaging revealed a right paracentral C6–C7 herniated disc compressing the exiting C7 nerve root. **B.** Lateral cervical spine X-ray of the same patient clearly demonstrates the position of the plate and screws. The allograft bone spacer placed in the drilled-out disc space is also apparent.

Figure 42-29 demonstrates typical findings. Some patients complain of difficulty buttoning shirts, using utensils, and ambulating. Spondylosis is usually diffuse, so the usual treatment for CSM is multilevel (usually C3–C7) cervical laminectomy, although patients with disease localized over one to three levels may be candidates for anterior decompression and fusion. Fig. 42-30 demonstrates the postoperative appearance of a vertebral corpectomy and fusion for CSM. Thorough cervical laminectomy decompresses the cord posteriorly. Patients often have slow recovery due to the extensive chronic changes in the cervical cord and may benefit from rehabilitation programs. The other disease that classically presents with combined upper and lower motor neuron symptoms is amyotrophic lateral sclerosis (ALS). Care must be taken...
to avoid offering cervical laminectomy to a patient with undiagnosed ALS. Two findings help differentiate CSM from ALS: cranial nerve dysfunction such as dysphagia (not typically caused by cervical spine disease) and sensory disturbance (not found in ALS).

**Thoracic Disc Herniation.** Thoracic disc herniation accounts for <1% of herniated discs. A patient may present with radicular pain or sensorimotor changes in the lower extremities due to cord compression. A posterior approach via midline incision and laminectomy should be avoided because of the high incidence of cord injury from manipulation and retraction. Anterior approaches via thoracotomy minimize risk to the cord and allow excellent access to the disc. The radicular arteries running from the aorta to the thoracic cord should be spared, when possible, to avoid ischemia. Alternatively, a posterolateral approach is possible via resection of the rib head and facet joint. Finally, a transpedicular approach may be attempted for lateral disc herniations.73

**Lumbar Radiculopathy.** Lumbar nerve roots exit the thecal sac, pass over the higher adjacent disc space, and exit the canal under the pedicle of the same-numbered vertebra. Therefore, the L5 nerve root passes over the L4–L5 disc space and exits under the L5 pedicle (Fig. 42-31). Lumbar discs may herniate with or without a history of trauma or straining. Normally they cause lancinating (radicular) pain down the leg (Table 42-7). Most acute herniated lumbar discs improve symptomatically without surgery. Surgery is indicated for symptoms persisting more than 6 to 8 weeks, progressive motor deficit (e.g., foot drop), or for patients with incapacitating pain not manageable with analgesics. A recent randomized control trial (Spine Patients Outcome Research Trial [SPORT])74 did not observe significant differences between patients randomized to undergo surgery vs. conservative management when using an intent-to-treat analysis. Because of the high amount of bi-directional cross-over between the surgery and conservative group, a subsequent reanalysis of the data as observational cohort analysis, demonstrated improved functional outcomes in terms of pain and physical function, more so at 3 months than at 2 years.74 Discectomy is performed using a midline incision, partial removal of the overlying laminae (hemilaminectomy or laminotomy), identification of the thecal sac and nerve root, and extraction of disc fragments. Free-floating disc fragments may be found. Often, however, the herniated disc material is still contained within the annulus, requiring incision of the posterior longitudinal ligament and curettage of the disc space. After lumbar discectomy, approximately two-thirds of patients will have complete relief of pain, and up to 85% will have significant improvement.

**Neurogenic Claudication.** Neurogenic claudication is characterized by low back and leg pain that occurs while walking and is relieved by stopping, leaning forward, or sitting. It is normally caused by degenerative lumbar stenosis causing compression of the cauda equina. Neurogenic claudication must be distinguished from vascular claudication, which tends to resolve quickly with cessation of walking. There is typically no need to change position, and the pain follows a stocking distribution rather than a dermatomal distribution. Pallor and coldness of the feet, and normal neurologic examination are also typical, though diabetic patients may present a challenge with microvascular
neuropathy. Patients with neurogenic claudication have a slowly progressive course and may be surgical candidates when their pain interferes with their lifestyle. The usual surgery is an L3 to L5 lumbar laminectomy to decompress the nerve roots. A recent randomized control trial did not observe a significant difference in outcomes between surgical and conservative management of lumbar stenosis at 1 year. Thus, surgical management of lumbar stenosis should be reserved for patients that do not improve with physical therapy.

Cauda Equina Syndrome. Cauda equina syndrome is due to compression of the cauda equina and may result from massive disc herniation, EDH, epidural abscess, tumor, or subluxation from trauma. Patients with cauda equina compression often present with urinary retention, saddle anesthesia, or progressing leg weakness. Saddle anesthesia is numbness in the perineum, genitals, buttocks, and upper inner thighs. Patients with suspected cauda equina syndrome should undergo immediate MRI of the lumbar spine to evaluate for a surgical lesion. Mass lesions should be removed urgently via laminectomy to preserve sphincter function and ambulation.

Spine Fusion Surgery
Fusion surgery is often required for patients with spinal instability resulting from disease, surgical intervention, or both. Fusion procedures lock adjacent vertebrae together. Fusion occurs when the body forms a solid mass of bone incorporating the adjacent vertebrae, eliminating normal intervertebral movement. Stabilization and immobilization promote bony fusion. Internal instrumentation and external orthoses are often used to stabilize and immobilize the fused spinal segments.

Spinal Instrumentation
Internal fixation devices for spinal segmental immobilization have been developed for all levels of the spine. Most spinal instrumentation constructs have two elements. The first element is a device that solidly attaches to the vertebral bodies. Options include wires wrapped around laminae or spinous processes, hooks placed under the lamina or around the pedicles, or screws placed in the pedicles or the vertebral bodies. The second element is a device that traverses vertebral segments. Options include rods and plates that lock directly to the wires, hooks, or screws at each vertebral level. Spinal instrumentation devices are available for anterior and posterior fusion in the cervical, thoracic, and lumbar regions. Most modern spinal instrumentation devices are made of titanium to minimize problems with future MRI scanning (Fig. 42-32). All spinal instrumentation constructs will eventually fail by loosening or breaking if bony fusion does not occur.

<table>
<thead>
<tr>
<th>LEVEL</th>
<th>FREQUENCY (%)</th>
<th>ROOT INJURED</th>
<th>REFLEX</th>
<th>WEAKNESS</th>
<th>NUMBNESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>L3–L4</td>
<td>5</td>
<td>L4</td>
<td>Patellar</td>
<td>Quadriceps</td>
<td>Anterior thigh</td>
</tr>
<tr>
<td>L4–L5</td>
<td>45</td>
<td>L5</td>
<td>—</td>
<td>Tibialis anterior (foot drop)</td>
<td>Great toe</td>
</tr>
<tr>
<td>L5–S1</td>
<td>50</td>
<td>S1</td>
<td>Achilles</td>
<td>Gastrocnemius</td>
<td>Lateral foot</td>
</tr>
</tbody>
</table>

Arthrodesis

Arthrodesis refers to the obliteration of motion or instability by incorporating the relevant components into a solid mass of bone. Arthrodesis must occur in any fused segment to have long-term stability. Failure of arthrodesis results in failed fusion, often in the form of a fibrous nonunion. The rates of successful fusion are higher in the cervical spine than the lumbar spine. Arthrodesis requires ingrowth of new bone formed by the patient’s osteoblasts across the unstable defect. Inserting graft material, such as autograft or allograft, into the defect provides a bridge for osteoblasts and promotes fusion. The term autograft refers to the patient’s own bone, often harvested from the iliac crest. Iliac crest bone graft is a source of both cortical and cancellous bone. Cortical bone provides structural support, while cancellous bone provides a matrix for bony ingrowth. The term allograft refers to sterilized bone from human tissue banks. Allografts also may be cortical, cancellous, or both. Allograft lacks the array of osteoinductive endogenous compounds intrinsic to autograft, although supplemental products such as demineralized bone matrix paste can be added to encourage new bone formation. Other techniques for increasing the rates of successful fusion are being developed, including the integration of osteoinductive bone morphogenetic proteins, known as BMPs, into the fusion constructs.

Dynamic stabilization refers to the creation of spinal stability without achieving a bony fusion. The concept applies to both cervical and lumbar motion segments. Artificial lumbar and cervical disc replacement therapies are recent developments in degenerative spine disease that address this concept. However, their use is limited to very select cases. Another motion preservation technique that may hold promise is segmental “soft” stabilization. In cases of degenerative spondylolisthesis, such systems in the lumbar spine allow for decompressive laminectomy without increasing slippage. In theory, adjacent level facets and discs are spared the stresses of a neighboring bony fusion moment arm.

Peripheral Nerve

Common pathologic processes that compromise function of the peripheral nervous system include mechanical compression, ischemia, inflammation, and neoplasia.

Peripheral Nerve Tumors

Most peripheral nerve tumors are benign and grow slowly. Significant pain increases the likelihood that the patient has a malignant tumor. Treatment for peripheral nerve tumors is surgical resection to establish diagnosis and evaluate for signs of malignancy. These tumors have various degrees of involvement with the parent nerve. Some can be resected with minimal or no damage to the nerve. Tumors that grow within the nerve often contain functioning fascicles. Total excision of these tumors requires sacrifice of the parent nerve. The choice of subtotal resection, nerve preservation, and observation, vs. total resection with nerve sacrifice depends on tumor histology and the function of the parent nerve.

Schwannoma. Schwannomas are the most common peripheral nerve tumors, also referred to as neurilemomas or neurinomas. Most occur in the third decade of life. These benign tumors arise from Schwann cells, which form myelin in peripheral nerves. The most characteristic presentation is a mass lesion with point tenderness and shooting pains on direct palpation. Spontaneous
or continuous pain suggests malignancy. Schwannomas tend to grow slowly and eccentrically on parent nerves. The eccentric location and discrete encapsulated nature of these tumors often allow total resection without significant damage to the parent nerve. Subtotal resection and observation is reasonable for schwannomas entwined in important nerves, as the incidence of malignant transformation is extremely low.

**Neurofibroma.** Neurofibromas arise within the nerve and tend to be fusiform masses, unlike schwannomas, which tend to grow out of the nerve. Neurofibromas often present as a mass that is tender to palpation. They usually lack the shooting pains characteristic of schwannomas. Neurofibromas are often difficult to resect completely without sacrifice of the parent nerve. Neurofibromas have a higher incidence of malignant transformation; therefore, patients with known residual tumors require close observation. Patients with NF1 often have multiple neurofibromas. These patients should be offered resection for symptomatic tumors. Risk of malignant degeneration is up to 10%. Malignant neurofibromas have the histologic characteristics of sarcoma.

**Malignant Nerve Sheath Tumors.** Malignant nerve sheath tumors include solitary sarcomas, degenerated neurofibromas, and neuroepitheliomias. Patients with malignant peripheral nerve tumors typically complain of constant pain, rather than pain only on palpation, and are more likely to have motor and sensory deficits in the distribution of the parent nerve. Treatment for these tumors is radical excision. This often requires sacrifice of the parent nerve. Invasion of nearby soft tissues may occur and necessitate wide resection or amputation in an attempt to prevent systemic metastasis.

**Entrapment Neuropathies**

Entrapment neuropathy presents as neurologic dysfunction in nerves passing through a pathologically small, fixed space. Nerve dysfunction may result directly from chronic, repetitive pressure on the nerve, or from ischemic damage due to impaired perfusion.75 Entrapment causing dysfunction of nerve signaling may be associated with numbness, paresthesias, weakness, or muscle atrophy. The two most common sites of entrapment neuropathy are the ulnar nerve at the medial aspect of the elbow and the median nerve at the wrist. Usually EMG/NCS demonstrate slowing across the entrapped segment of nerve. Mechanical peripheral nerve disorders resulting from trauma (brachial plexus disruption, radial nerve damage from humerus fractures, and common peroneal nerve crush injuries) are discussed in the section “Trauma.”

**Ulnar Neuropathy.** The ulnar nerve has contributions from the C7, C8, and T1 nerve roots, arises from the medial cord of the brachial plexus, and supplies most of the intrinsic hand muscles (interossei and third and fourth lumbricals) and sensation to the fourth and fifth digits. It passes posteriorly to the medial epicondyle at the elbow in the condylar groove. This segment is superficial and subject to external compression and repetitive minor impacts. Patients with ulnar entrapment at the elbow present with numbness and tingling in the medial palm, as well as the fourth and fifth digits. Motor deficits include weakness and wasting of the intrinsic hand muscles. Treatment for symptomatic ulnar entrapment neuropathy is surgical exploration and incision of the fibrous aponeurotic arch that overlies the nerve. A 6-cm curvilinear incision centered between the medial epicondyle and the olecranon allows exploration of up to 10 cm of nerve and lysis of compressive tissues.

**Carpal Tunnel Syndrome.** The median nerve has contributions from the C5 to T1 nerve roots, arises from the medial and lateral cords of the brachial plexus, and supplies the muscles of wrist and finger flexion and sensation to the palmar aspect of the first, second, and third digits. The median nerve passes through the carpal tunnel in the wrist, lying superficial to the four deep and four superficial flexor tendons. The transverse carpal ligament is a tough, fibrous band that forms the roof of the carpal tunnel. The ligament attaches to the pisiform and hamate medially and the trapezium and scaphoid laterally. Patients complain of numbness and tingling in the supplied digits, clumsiness, and worsening with sleep or repetitive wrist movement. Patients may notice wasting of the thenar eminence. Treatment for symptomatic carpal tunnel syndrome unresponsive to splinting, analgesics, and rest is surgical division of the flexor retinaculum. This often provides prompt relief of pain symptoms and slow recovery of numbness and strength.

**Autoimmune and Inflammatory Disorders**

These are not surgical diseases, but they merit brief mention as they are included in the differential diagnosis for new-onset weakness. Their characteristic presentations help distinguish them from weakness due to structural lesions.

**Guillain-Barré Syndrome.** Guillain-Barré syndrome is an acute inflammatory demyelinating polyradiculopathy often occurring after viral infection, surgery, inoculations, or mycoplasmal infections. Patients classically present with weakness ascending from the legs to the body, arms, and even cranial nerves. Symptoms usually progress over 2 to 4 weeks and then resolve. Care is supportive. Respiratory weakness may require ventilatory support.

**Myasthenia Gravis.** Myasthenia gravis is an autoimmune process in which antibodies form to the acetylcholine receptors of muscles, leading to fluctuating weakness. Most patients have either thymic hyperplasia or thymoma. The most common symptoms are diplopia, ptosis, dysarthria, and dysphagia. More severe cases have limb or respiratory involvement. Weakness worsens with repetitive movement. Treatment is with acetylcholinesterase inhibitors and possible thymectomy.

**Eaton-Lambert Syndrome.** Eaton-Lambert syndrome is an autoimmune process with antibodies to the presynaptic calcium channels. This is a paraneoplastic syndrome most commonly associated with oat cell carcinoma. Patients have weakness of proximal limb muscles that improves with repetitive movement. This diagnosis must prompt oncologic evaluation.

**INFECTION**

CNS infections of interest to neurosurgeons include those that cause focal neurologic deficit due to mass effect, require surgical aspiration or drainage because antibiotic therapy alone is insufficient, cause mechanical instability of the spine, or occur after neurosurgical procedures.

**Cranial**

**Osteomyelitis.** The skull is highly vascular and resistant to infections. Osteomyelitis of the skull may develop by contiguous spread from pyogenic sinus disease or from contamination by penetrating trauma. *Staphylococcus aureus* and *S epidermidis* are the most frequent causative organisms. Patients usually present with redness, swelling, and pain. Contrast head CT aids
diagnosis and shows the extent of involved bone, along with associated abscesses or empyema. Osteomyelitis treatment entails surgical debridement of involved bone followed by 2 to 4 months of antibiotics. Craniotomy wound infections are a special concern because performing a craniotomy creates a devascularized free bone flap susceptible to infection and not penetrated by antibiotics. These wounds must be debrided and the bone flaps removed and discarded. Subsequent care involves appropriate antibiotic therapy, observation for signs of recurrent infection off antibiotics, and return to the OR for titanium or methylmethacrylate cranioplasty 6 to 12 months later.

**Subdural Empyema.** Subdural empyema is a rapidly progressive pyogenic infection. The subdural space lacks significant barriers to the spread of the infection, such as compartmentalization or septations. Subdural empyemas usually occur over the cerebral convexities. Potential infectious sources include sinus disease, penetrating trauma, and otitis. Streptococci and staphylococci are the most frequent sources. Presenting symptoms include fever, headache, neck stiffness, seizures, or focal neurologic deficit. Neurologic deficit results from inflammation of cortical blood vessels, leading to thrombosis and stroke. The most common deficit is contralateral hemiparesis. Patients with suggestive symptoms should undergo rapid contrast CT scan. LP frequently fails to yield the offending organism and risks herniation due to mass effect. Typical treatment is wide hemicraniectomy, dural opening, and lavage. The pus may be thick or septated, making burr hole drainage or small craniotomy insufficient. Patients then require 1 to 2 months of antibiotics. Subdural empyema has 10% to 20% mortality risk and common chronic sequelae, including development of a seizure disorder and residual hemiparesis. However, many patients do make a good recovery.

**Brain Abscess.** Brain abscess is encapsulated infection within the brain parenchyma. It may spread hematogenously in patients with endocarditis or intracardiac or intrapulmonary right-to-left shunts, by migration from the sinuses or ear, or via direct seeding by penetrating trauma. Disorganized cerebritis often precedes formation of the organized, walled-off abscess. Patients may present with nonspecific symptoms such as headache, nausea, or lethargy, or with focal neurologic deficit such as hemiparesis. Alternatively, patients may present in extremis if the abscess ruptures into the ventricular system. Abscesses appear as well-demarcated, ring-enhancing, thin-walled lesions on CT scan and MRI, and often have associated edema and mass effect. Patients require antibiotic therapy after needle aspiration or surgical evacuation. Antibiotic therapy without surgical evacuation may be considered for patients with small, multiple, or critically located abscesses. Abscesses that are large, cause mass effect, decreased mental status, or that fail to decrease in size after 1 week of antibiotics, should be evacuated. Nonsurgical management still requires aspiration or biopsy specimen for organism culture and sensitivities. Blood and CSF cultures rarely give definitive diagnosis. Removal of an encapsulated abscess significantly shortens the length of antibiotic therapy required to eliminate all organisms. Common chronic sequelae after successful treatment include seizures or focal neurologic deficit.

**Spine**

**Pyogenic Vertebral Osteomyelitis.** Pyogenic vertebral osteomyelitis is a destructive bacterial infection of the vertebrae, usually of the vertebral body. Vertebral osteomyelitis frequently results from hematogenous spread of distant disease, but may occur as an extension of adjacent disease, such as psoas abscess or perinephric abscess. *S aureus* and *Enterobacter* spp. are the most frequent etiologic organisms. Patients usually present with fever and back pain. Diabetics, IV drug abusers, and dialysis patients have increased incidence of vertebral osteomyelitis. Epidural extension may lead to compression of the spinal cord or nerve roots with resultant neurologic deficit. Osteomyelitis presents a lytic picture on imaging and must be distinguished from neoplastic disease. Adjacent intervertebral disc involvement occurs frequently with pyogenic osteomyelitis, but rarely with neoplasia. Plain films and CT help assess the extent of bony destruction or deformity such as kyphosis. MRI shows adjacent soft tissue or epidural disease. Most cases can be treated successfully with antibiotics alone, although the organism must be isolated to steer antibiotic choice. Blood cultures may be positive. Surgical intervention may be required for debridement when antibiotics alone fail, or for stabilization and fusion in the setting of instability and deformity.

**Tuberculous Vertebral Osteomyelitis.** Tuberculous vertebral osteomyelitis, also known as *Pott’s disease*, occurs most commonly in underdeveloped countries and in the immunocompromised. Several features differentiate tuberculous osteomyelitis from bacterial osteomyelitis. The infection is indolent and symptoms often progress slowly over months. Tuberculosis rarely involves the intervertebral disc. The involved bodies may have sclerotic rather than lytic changes. Multiple nonadjacent vertebrae may be involved. The upper lumbar and lower thoracic vertebrae are most commonly affected. Diagnosis requires documentation of acid-fast bacilli. Treatment involves long-term antimycobacterial drugs. Patients with spinal instability or neural compression from epidural inflammatory tissue should undergo debridement and fusion as needed.

**Discitis.** Primary infection of the intervertebral disc space, or discitis, is most commonly secondary to postoperative infections. Spontaneous discitis occurs more commonly in children. *S aureus* and *S epidermidis* account for most cases. The primary symptom is back pain. Other signs and symptoms include radicular pain, fevers, paraspinal muscle spasm, and localized tenderness to palpation. Many cases will resolve without antibiotics, which generally are given for positive blood or biopsy specimen cultures or persistent constitutional symptoms. Most patients will have spontaneous fusion across the involved disc and do not need debridement or fusion.

**Epidural Abscess.** Epidural abscesses may arise from or spread to the adjacent bone or disc, so distinguishing between vertebral osteomyelitis or discitis and a spinal epidural abscess may be difficult. The most common presenting signs and symptoms are back pain, fever, and tenderness to palpation of the spine. The most significant risk of epidural abscess is weakness progressing to paralysis due to spinal cord or nerve root damage. Cord and root damage may be due to direct compression or to inflammatory thrombosis resulting in venous infarction. *S aureus* and *Streptococcus* spp. are the most common organisms. Meticillin-resistant *S aureus* now constitutes a significant proportion of these infections, as high as 40%. The source may be hematogenous spread, local extension, or operative contamination. MRI best demonstrates the epidural space and degree of neural compromise. Patients with spinal epidural abscess and neurologic compromise should undergo surgical debridement for decompression and diagnosis, followed by culture-directed therapy.
antibiotic therapy. Relative contraindications to surgery include prohibitive comorbidities or total lack of neurologic function below the involved level. Patients with no neurologic deficits and an identified organism may be treated with antibiotics alone and very close observation. However, this management strategy remains somewhat controversial because these patients can undergo rapid and irreversible neurologic decline. Most epidural abscesses can be accessed via laminectomy without fusion. Collections predominantly anterior to the cervical or thoracic cord may require anterior approach and fusion.

**FUNCTIONAL NEUROSURGERY**

**Epilepsy Surgery**

Seizures result from uncontrolled neuronal electrical activity. Seizures may result from irritative lesions in the brain, such as tumors or hematomas, or from physiologic or structural abnormalities. Seizures may involve a part of the brain (focal) or the entire brain (generalized). Focal seizures may be associated with normal consciousness (simple) or decreased consciousness (complex). All generalized seizures cause loss of consciousness. Focal seizures may secondarily generalize. Patients with multiple unprovoked seizures over time are considered to have epilepsy. The type of epilepsy depends on such factors as type of seizures, electroencephalographic (EEG) findings, associated syndromes, and identifiable etiologies. All patients with unexplained seizures (i.e., no obvious cause such as head trauma or alcohol withdrawal) require thorough neurologic evaluation, including imaging to evaluate for a mass lesion. Antiepileptic drugs (AEDs) form the first line of therapy for epilepsy, initially as monotherapy, then as combination therapy. Epilepsy patients who have failed satisfactory trials of several AED combination regimens may be candidates for surgical intervention. Lack of seizure control or patient intolerance of the medications may constitute failure. Epilepsy surgery can decrease the frequency of seizures by resection of the electrical source of the seizures, or decrease the severity of seizures by disconnecting white matter tracts through which the abnormal electrical activity spreads. Four types of epilepsy surgery are discussed in sections that follow. Epilepsy surgery appears to be extremely underused, given the relatively low risk of the procedures, and the crippling social and economic effects of uncontrolled or partially controlled epilepsy. Patients with symptoms, imaging abnormalities, and EEG analysis compatible with a specific seizure focus are most likely to have good results from epilepsy surgery.

**Anterior Temporal Lobectomy.** Medial temporal lobe structural abnormalities can lead to complex partial seizures (CPS). Many patients with CPS have poor seizure control on medications. Patients with CPS may have significant reduction in seizure frequency or cessation of seizures after resection of the anterior temporal lobe. The amygdala and the head of the hippocampus are removed as part of the lobectomy. Resection may be taken back approximately 4.5 cm from the temporal tip in the language-dominant hemisphere, and 6 cm from the temporal tip in the language nondominant hemisphere, with low risk of significant neurologic deficits. The two main risks of anterior temporal lobectomy are memory impairment and visual loss. Removal of the hippocampus in a patient with an atrophied or nonfunctional contralateral hippocampus causes a global memory deficit. Interruption of the optic radiations, which carry visual signals from the contralateral superior visual quadrants of both eyes, causes a contralateral superior quadrant anopia, known as a *pie in the sky* field cut.

**Corpus Callosotomy.** Patients with generalized seizures, atonic seizures associated with drop attacks, or absence seizures, who are found to have bilaterally coordinated pathologic cortical discharges on EEG and who fail AED therapy, may be candidates for corpus callosotomy. The corpus callosum is a large white matter tract that connects the cerebral hemispheres. Loss of consciousness requires simultaneous seizure activity in both hemispheres. Focal or partial seizures may spread via the corpus callosum to the contralateral hemisphere, causing generalization and loss of consciousness. Division of the corpus callosum can interrupt this spread. Patients may have decreased numbers of seizures and/or fewer episodes of lost consciousness. Usually only the anterior half or two-thirds of the corpus callosum is divided, as more extensive division increases the risk of disconnection syndrome. Patients with disconnection syndrome are unable to match objects in the opposite visual hemifields, to identify objects held in one hand with the other hemifield, and to write with the left hand or name objects held in the left hand (in left hemisphere–dominant patients).

**Hemispherectomy.** Children with intractable epilepsy, structural anomalies in one hemisphere, and contralateral hemiplegia, may have improved seizure control after resection of the hemisphere (anatomic hemispherectomy) or disruption of all connections to the hemisphere (functional hemispherectomy). Functional hemispherectomy often is preferred over anatomic hemispherectomy because of the high incidence of complications such as hematoma formation and ventriculoperitoneal shunt dependence associated with the latter.

**Vagus Nerve Stimulation.** Neuromodulatory treatments like vagus nerve stimulation (VNS), approved by the U.S. Food and Drug Administration (FDA) in 1997, are less invasive and offer some titratability in addition to reversibility unlike the resective surgical options previously described. Since first reported in 1985, VNS has proven to be efficacious in certain patient populations for several disorders such as treatment-resistant major depressive disorder, bipolar disorder, and epilepsy. In VNS, a pulse generator is placed under the skin in the chest and is connected to the vagus nerve by an electrical lead. Chronic, intermittent VNS has been proven to be an effective option for patients suffering from medically refractory seizures who are not candidates for surgical resection. Although only a small minority of patients will be entirely seizure-free, three blinded, randomized-controlled trials have examined VNS and demonstrated significant clinical improvement compared to sham.

Generally VNS is well-tolerated and safe, as device implantation is associated with a low rate of perioperative complications. Additionally, the majority of side effects are stimulation-dependent and thus, reversible. For the most part, VNS is limited in its application because it can only exert its effects by altering neural activity via the vagus nerve. Procedures with brain region-specificity are being investigated.

**Deep Brain Stimulation**

The following summary of deep brain stimulation (DBS) will include a review of the current FDA-approved indications, as well the expanding applications of this therapy, currently being investigated preclinically and in clinical trials. While the
mechanism of action of DBS continues to elude our understanding, it is well established that administering electrical stimulation to a nucleus in the brain known to be involved in a given disease can disrupt the pathologic signals emanating to or from this brain region. A fine electrical lead is placed in a deep brain nucleus and connected to pulse generators placed in the chest in a manner similar to cardiac pacemakers. Connector wires travel from the generators in the subcutaneous space up the neck and in the subgaleal space in the head, to connect the pulse generators to the electrical leads. Proper lead placement is accomplished with stereotactic guidance. A frame is rigidly fixed to the patient’s head, and an MRI is obtained with the frame in place. Calculation of the coordinates of the millimeter-sized deep brain nuclei is performed in relation to the three-dimensional space defined by the fixed frame, allowing for accurate targeting of the nucleus (Fig. 42-33). Postoperatively, the pulse generators can be interrogated and adjusted with hand-held, transcutaneous, noninvasive devices as needed for symptom control.

**Essential Tremor.** Essential tremor is the most common movement disorder in the western world and is characterized by action tremor (4–8 Hz rhythmic oscillations) of the hands, forearms, head, and voice. Essential tremor often starts in the third or fourth decade of life and increases in frequency and amplitude with age. β-Blockers can decrease symptoms, but patients with poor medical control and significant functional impairment significantly benefit from placement of a deep brain stimulator in the contralateral ventralis intermediate nucleus of the thalamus. In properly selected patients, DBS of this region of the thalamus appears to provide robust and durable symptom control.84,85

**Parkinson’s Disease.** Parkinson’s disease is a progressive disorder characterized by rigidity, bradykinesia, and resting tremor, due to loss of dopamine-secreting neurons in the substantia nigra. Dopaminergic agents such as levodopa/carbidopa and anticholinergic agents such as amantadine and selegiline form the basis of medical therapy. Patients with poor medical control or significant drug side effects may benefit significantly from placement of bilateral deep brain stimulators in the subthalamic nuclei. Although the globus pallidus interna has also been a widely targeted area, the subthalamic nucleus is now the most accepted target in deep brain stimulation for Parkinson’s disease.86 Deep brain stimulation provides durable symptom relief with good postoperative neuropsychologic function in properly selected patients.87

Recently, a large randomized controlled trial compared bilateral DBS (n = 121) to best medical therapy in advanced Parkinson’s disease (n = 134).88 The DBS group did significantly better in both motor function and quality of life. While adverse events were 3.8 times more likely in the DBS group, 99% of these events had resolved by 6 months. There was a 0.8% risk of death due to the procedure, and there was no difference in risk of adverse events when comparing older (≥70 years) to younger patients (<70 years). Thus, the benefits of DBS over medical therapy are clear, especially when considering quality of life measures.

Another recent randomized controlled trial focused on defining the optimal targets for DBS in Parkinson’s disease.89 While the subthalamic nucleus (STN) and the globus pallidus interna (GPI) have been successfully targeted in the past, a direct comparison of the two was lacking. In this study, 299 subjects were randomized to receive either bilateral STN or GPI stimulators and were evaluated for 2 years. The primary outcome was motor function, as assessed by part III of the Unified Parkinson’s Disease Rating Scale (UPDRS). The study found no significant difference in motor improvement between target sites. However, a significant difference was found in a secondary outcome measuring depression. On the Beck Depression Inventory, the pallidal stimulation group improved slightly compared with the STN group, which actually worsened slightly. Nevertheless, the actual incidence of depressive episodes requiring prolonged or new hospitalization was 2.6% and 0.7% in GPI and STN, respectively, which was not significantly different. On the other hand, the STN group was found to require less adjunctive dopaminergic pharmacotherapy than the GPI group. In terms of overall severe adverse events, there was no difference between groups. The investigators concluded that both target sites are effective and that nonmotor factors such as psychiatric symptoms may be a consideration in DBS target selection.

**Dystonia.** The FDA humanitarian device exemption has been made for DBS for dystonia but is limited to patients ≥7 years of age with primary dystonia, including generalized and/or segmental dystonia, hemidystonia, or cervical dystonia (torticollis). Dystonia is characterized by sustained muscle contractions that cause repetitive movements and involuntary postures. Cognitive function is typically spared, and pharmacological therapy is frequently inadequate. The positive impact of DBS on Parkinson’s disease is currently considered the most efficacious target for dystonia, and controlled trials indicate approximately a 50% improvement in motor function and disability.90 Since many patients undergoing

![Figure 42-33.](image-url)
surgery for dystonia are children and young adults, DBS is an attractive surgical option because it can be titrated, revised, and reversed according to individual needs and growth patterns.

**Obsessive-Compulsive Disorder.** The safety and efficacy of DBS, as well as its titratability and reversibility, have been demonstrated for the treatment of movement disorders in the 1990s and 2000s has spawned an increasing interest and awareness of the capabilities of nonlesional surgical treatments for diseases of the brain. An obvious outgrowth of DBS for movement disorders has been the treatment of medically refractory psychiatric disorders. Despite the dark history of frontal leucotomy procedures that dominated the early 20th century, nonlesional DBS for psychiatric disorders are now considered potential treatment strategies.

Functional neuroimaging has implicated certain brain regions in the pathogenesis of a variety of psychiatric disorders. The FDA has approved a humanitarian device exemption for DBS targeting the ventral capsule/ventral striatum for severe obsessive-compulsive disorder (OCD). Recent case reports and pilot studies have reported remission in patients suffering from refractory OCD following DBS. A pilot study using a blinded, staggered-onset design found that four (66.7%) of six patients met a stringent criterion as “responders” (≥35% improvement), according to the Yale-Brown Obsessive Compulsive Scale after 12 months of stimulation.9 In this study, patients did not improve during the sham phase. Adverse events were generally mild and modifiable with setting changes, and stimulation interruption led to rapid yet reversible development of depressive symptoms in two cases. Thus, DBS has promise as a therapy of last resort for carefully selected cases of severe OCD.

**Expanding Indications of Deep Brain Stimulation.** There are multiple disorders, both psychiatric and neurologic, that have exhibited significant promise as potential indications for DBS in large-scale trials. Recently, there have been reports of significant improvements in refractory depression with DBS. Lozano and colleagues performed an open label study with extended follow-up on 20 patients targeting an area within the subcallosal cingulate gyrus (SCG) with bilateral DBS.9 At the last follow-up visit in this study (range: 3–6 years), the average response rate was 64%, according to the Hamilton Rating Scale for Depression. Of note, impairment in social functioning was improved, and no significant adverse events were reported. Because two patients died by suicide during depressive relapses, it remains unclear if DBS can only improve quality of life or significantly suppress relapses and extend life-span in this extremely delicate patient population. Of note, as seen in OCD, the ventral capsule/ventral striatum has also been targeted for depression, as well as the nucleus accumbens directly, which lies within the ventral striatum. Studies of DBS in this region report an approximate 40% to 60% response rate, and results from a recent, multicenter randomized controlled trial are pending.91

DBS as a potential therapy for epilepsy targeting the anterior nucleus of the thalamus has been investigated in a multicenter, double-blind, randomized trial (SANTE).94 In this trial, the group receiving DBS showed a 29% greater reduction in seizure frequency in relation to the sham group in the last month of the blinded phase. Complex partial and the “most severe” seizures were significantly reduced in the cohort who had the stimulator on DBS-on group. After the blinded phase of the trial was complete, 54% of patients had a seizure reduction of at least 50%. Fourteen patients were seizure-free for at least 6 months; eight were seizure-free for at least one year, four for at least two years, and one patient for more than four years. Because of the modest benefit during the blinded phase of this trial, FDA-approval has yet to be granted to DBS for epilepsy targeting the thalamus in the United States, though approval has been given in Europe and Canada.

The region-specific, neuromodulatory capabilities of DBS have inspired the open label use of this technique in many other neurologic and psychiatric disorders, including but not limited to Tourette syndrome, Huntington’s disease, and Alzheimer disease. Preclinical studies of both substance abuse and obesity have also shown promise.95,96 The opportunity to model reward-seeking behaviors associated with these disorders in animals provides the ability to not only test safety but also study mechanisms and inform the design of future clinical trials.

**Trigeminal Neuralgia**

Trigeminal neuralgia, also known as tic douloureux, is characterized by repetitive, unilateral, sharp, and lancinating pains in the distribution of, typically, the second, but sometimes third, branch of cranial nerve V, the trigeminal nerve. The patient may describe a “trigger point,” an area on the face that elicits the pain when touched. A current leading etiologic hypothesis for trigeminal neuralgia is irritation and pulsatile compression of the root entry zone of the nerve by an artery in the posterior fossa, usually a loop of the superior cerebellar artery. The pain is excruciating and can be debilitating. Medical therapy, including carbamazepine and amitriptyline, may reduce the frequency of events. Options for medically refractory cases include percutaneous injection of glycerol into the path of the nerve, peripher- al transection of the nerve branches, SRS, and microvascular decompression (MVD).

MVD involves performing a small posterior fossa craniotomy on the side of the symptoms, retraction of the cerebellar hemisphere, and exploration of cranial nerve V. If an artery is found near the nerve, the vessel is freed of any adhesions and nonabsorbable material is placed between the nerve root and the artery. MVD remains the first definitive management option because SRS is associated with a substantial incidence of facial numbness.97,98

**STEREOTACTIC RADIOSURGERY**

The term stereotactic radiosurgery (SRS) refers to techniques that allow delivery of high-dose radiation that conforms to the shape of the target and has rapid isodose fall-off, minimizing damage to adjacent neural structures. The two most common devices used for conformal SRS for intracranial lesions are the LINAC (linear accelerator) and the gamma knife. LINAC delivers a focused beam of x-ray radiation from a port that arcs part way around the patient’s head. Linear accelerators are commonly used to provide fractionated radiation for lesions outside the CNS. They are found in most radiation oncology departments. After upgrades to the software and collimators, SRS can be performed with these existing units. The gamma knife delivers 201 focused beams of gamma radiation from cobalt sources through a specially designed colander-like helmet. Gamma knife units are used only for intracranial disease and cost up to $5 million; thus, they are most appropriate in high patient–volume centers. There is ongoing debate in the literature...
regarding the two technologies. Both continue to evolve, allowing more precise and complex isodose conformation to complex lesions. Most lesions can be treated equally well with either technology. Lesions abutting the medulla or the spinal cord should not be treated with SRS because these structures do not tolerate the radiation dose delivered to structures within millimeters of the target. Also, medullary or spinal cord compression can result from swelling of the lesion after the radiosurgery dose, resulting in devastating neurologic deficit.

Proton beam is an evolving SRS technology that may play a specialized role in treatment of lesions where posttarget exiting radiation limits photon-based therapies. For example, the physical properties of photons cause destruction upon entry and exit from tissue, which can be particularly harmful to skull-base or clival lesions such as chordoma, in which the exiting pathway travels through the brain stem. Proton beam therapy uses accelerated protons, which dissipate energy upon uptake and do not cause additional exiting damage. Currently, there are very few centers using this technology.

CyberKnife is another radiosurgery system that has neurosurgical application. It is a frameless, robotic, LINAC-based system that allows for targeting of spinal neoplasms with higher resolution than conventional external beam radiotherapy. Using imaging tracking in real time, the CyberKnife is able to adjust to breathing artifact and patient movement. The application of this technology is rapidly growing.

Arteriovenous Malformations
SRS has been found to be an effective stand-alone therapy for AVMs up to 3 cm in diameter. SRS is best for lesions that are difficult to access surgically due to high likelihood of postoperative neurologic deficit. However, SRS is not effective for lesions >3 cm. Effective obliteration and elimination of the risk of hemorrhage takes 2 to 3 years. Overall, there is an approximately 2% annual incidence of AVM hemorrhage, although one study found a 50% decrease in hemorrhage rate during the latency period before angiographic obliteration. Nonetheless, surgical excision remains the preferred therapeutic modality, while SRS is reserved for cases deemed very high risk for surgery due to location or patient factors. Some patients with large AVMs who undergo surgery will have unresectable residual lesions. In these patients, SRS may be used as an effective adjunctive therapy.

Vestibular Schwannomas
SRS has been introduced as a therapeutic alternative to microsurgical resection for vestibular schwannomas up to 2.5 cm in maximum diameter. SRS provides high rates of tumor growth arrest and possible reduction in size with low rates of facial nerve palsy. Patients with functional ipsilateral preprocedure hearing may be more likely to retain functional hearing postprocedure than with microsurgery. The limitations of SRS include inability to treat tumors >2.5 cm, the possibility of radiation-induced malignant transformation of these benign tumors, and lack of long-term follow-up. SRS centers are accumulating experience with these tumors and accumulating data on long-term results. The indications for microsurgery and SRS will continue to evolve. Either approach should be undertaken at a high-volume center, as studies show the patient outcomes improve with increased surgeon experience.

Intracranial Metastases
Patients with solitary or multiple intracranial metastases may be treated primarily with SRS. Patients have improved survival after SRS compared to no treatment or WBRT, and similar survival to patients undergoing total surgical resection. Patients with lesions >3 cm in diameter or evidence of ICH should undergo surgical decompression rather than SRS. Some studies show improved survival with up to seven intracranial masses. Patients with multiple intracranial masses have almost zero long-term survival, and most will die of their intracranial disease. Patients with intracranial metastases live 3 to 6 months on average with medical care and WBRT. This can be extended to 9 to 16 months with SRS or surgery, depending on tumor type, age, and patient condition.

CONGENITAL AND DEVELOPMENTAL ANOMALIES

Dysraphism
Dysraphism describes defects of fusion of the neural tube involving the neural tube itself, or overlying bone or skin. Dysraphism may occur in the spine or head. Neural tube defects are among the most common congenital abnormalities. Prenatal vitamins, especially folic acid, reduce the incidence of neural tube defects.

Spina Bifida Occulta
Spina bifida occulta is congenital absence of posterior vertebral elements. The spinous process is always missing, the laminae may be missing to various degrees, but the underlying neural tissues are not involved. Spina bifida occulta is found in 25% of the general population, and it is asymptomatic unless associated with other developmental abnormalities.

Spina Bifida With Myelomeningocele
Spina bifida with myelomeningocele describes the congenital absence of posterior vertebral elements with protrusion of the meninges through the defect, with underlying neural structural abnormalities. Common findings include weakness and atrophy of the lower extremities, gait disturbance, urinary incontinence, constipation, and deformities of the foot. Myelomeningoceles have unique clinical issues. Patients with abnormal protrusion of meninges through the bony defect without abnormalities of the underlying neural tissue have a meningocele. Most of these patients are neurologically normal.

Encephalocele
Herniation of brain encased in meninges through the skull that forms an intracranial mass is referred to as encephalocele. Herniation of meninges without brain tissue is referred to as a meningocele. Most occur over the convexity of the skull. More rarely, the tissue protrudes through the skull base into the sinuses. Treatment involves excision of the herniated tissue and closure of the defect. Most patients with encephaloceles and meningoceles have impaired cognitive development. Patients with greater amounts of herniated neural tissue tend to have more severe cognitive deficits.

Craniosynostosis
Craniosynostosis is the abnormal early fusion of a cranial suture line with resultant restriction of skull growth in the affected area.
and compensatory bulging at the other sutures. Skull growth occurs at the cranial sutures for the first 2 years of life, at the end of which the skull has achieved >90% of its eventual adult size. Fusion of the sagittal suture, or sagittal synostosis, results in a boat-shaped head, known as scaphocephaly. Unilateral coronal synostosis results in ipsilateral forehead flattening and outward deviation of the orbit, known as plagiocephaly. The contralateral normal forehead appears to bulge by comparison. Bilateral coronal synostosis results in a broad, flattened forehead, known as brachycephaly, and is often associated with maxillary hypoplasia and proptosis. Unilateral or bilateral lambdoid synostosis results in flattening of the occiput. Occipital flattening can result from abnormal suture fusion (synostosis), or from physical remodeling of the skull caused by always placing the baby in the supine position for sleep (known as positional plagiocephaly). Placing the baby in the prone position or tilted onto the contralateral side may restore near-normal skull shape in most cases of lambdoid synostosis, avoiding surgery. Treatment for synostoses in general is surgical, involving resection of the fused suture, or more complex reconstructive techniques for severe or refractory cases.

**Hydrocephalus**

Excess CSF in the brain that results in enlarged ventricles is known as hydrocephalus. CSF flows from the ventricles to the subarachnoid space and is then absorbed into the venous blood through the arachnoid granulations. Hydrocephalus may be classified as communicating or obstructive (outlined in the next two sections), and congenital or acquired. Congenital lesions associated with or causing hydrocephalus include stenosis of the cerebral aqueduct, Chiari malformation, myelomeningocele, and intrauterine infection. Acquired hydrocephalus may result from occlusion of arachnoid granulations by meningitis, germinal matrix hemorrhage, or SAH. CSF pathways may be occluded by adjacent tumors (Fig. 42-34).

**Communicating Hydrocephalus.** Obstruction at the level of the arachnoid granulations constitutes communicating hydrocephalus. This usually causes dilation of the lateral, third, and fourth ventricles equally. The most common causes in adults are meningitis and SAH. Hydrocephalus may be transient after SAH, with reestablishment of normal CSF absorption after the protein content of the CSF returns to normal and the granulations reopen.

**Obstructive Hydrocephalus.** Obstruction of CSF pathways is known as obstructive hydrocephalus. Ventricles proximal to the obstruction dilate, while those distal to the obstruction remain normal in size. Typical patterns include dilation of the lateral ventricles due to a colloid cyst occluding the foramen of Monro, dilation of the lateral and third ventricles due to a tectal (midbrain) glioma or pineal region tumor occluding the cerebral aqueduct, or dilation of the lateral and third ventricles with obliteration of the fourth ventricle by an intraventricular tumor of the fourth ventricle. Obstructive hydrocephalus may present precipitously and require urgent shunting to prevent herniation.

**Chiari I Malformation**

Chiari I malformation is the caudal displacement of the cerebellar tonsils below the foramen magnum. It may be seen as an incidental finding on MRI scans in asymptomatic patients. Symptomatic patients usually present with headache, neck pain, or symptoms of myelopathy, including numbness or weakness in the extremities. A syrinx may be associated, but the brain stem and lower cranial nerves are normal in Chiari I malformations. Chiari II malformations are more severe and involve caudal displacement of the lower brain stem and stretching of the lower cranial nerves. Symptomatic patients may be treated with suboccipital craniectomy to remove the posterior arch of the foramen magnum, along with removal of the posterior ring...
of C1. Removal of these bony structures relieves the compression of the cerebellar tonsils and cervicomedullary junction, and may allow reestablishment of normal CSF flow patterns. Figure 42-35 demonstrates typical MRI appearance of a Chiari I malformation.

REFERENCES

Entries highlighted in bright blue are key references.


Orthopedic Surgery
Nabil A. Ebraheim, Bert J. Thomas, Freddie H. Fu, Bart Muller, Dharmesh Vyas, Matt Niesen, Jonathan Pribaz, and Klaus Draenert

Introduction 1880

Orthopedic Trauma 1881
Introduction / 1881
Open Fractures / 1881
Compartment Syndrome / 1883

Treatment of Fractures and Dislocations 1883
Clavicle Fractures / 1883
Scapula Fractures / 1883
Shoulder Dislocations / 1884
Proximal Humerus Fractures / 1884
Humeral Shaft Fractures / 1884
Distal Humerus Fractures / 1884
Radial Head Fractures / 1886
Olecranon Fractures / 1886
Forearm Fractures / 1886
Distal Radius Fractures / 1886
Scaphoid Fractures / 1887
Pelvic Fractures / 1887
Acetabular Fractures / 1888
Hip Dislocations / 1888
Hip Fractures / 1888
Femoral Shaft Fractures / 1890
Distal Femur Fractures / 1890
Knee Dislocations / 1890
Patella/Extensor Mechanism Injuries / 1891
Tibial Plateau Fractures / 1891
Tibial Shaft Fractures / 1891
Tibial Plafond (Pilon) Fractures / 1892
Ankle and Subtalar Dislocations / 1892
Ankle Fractures / 1893
Malleonuneau Fractures / 1893
Calcaneal Fractures / 1893
Talus Fractures / 1894
Foot Fractures / 1894

Sports Medicine 1895
Introduction / 1895
Shoulder 1895
Rotator Cuff / 1895
Shoulder Instability / 1896
Posterior Dislocation of the Shoulder / 1896
Superior Labrum and Biceps Tendon / 1896

Impingement Syndromes / 1896
The Acromioclavicular Joint / 1897
Knee 1897
Menisci / 1897
Collateral Ligaments / 1898
Cruciate Ligaments / 1898
Posterolateral Corner / 1899

Spine 1900
Spinal Trauma / 1900
Occipital Cervical Dislocation / 1900
Fractures of C1 (Jefferson Fracture) / 1900
Fractures of C2 (Odontoid Fracture) / 1900
Hangman's Fractures of C2 / 1901
Compression Fracture of the Cervical Spine / 1901
Burst Fractures of the Cervical Spine / 1901
Unilateral and Bilateral Facet Dislocation / 1901
Clay-Shoveler's Injury / 1902

Fractures of the Thoracic and Lumbar Spine 1902
Thoracic Lumbar Spine Injury / 1902
Compression Fracture / 1902
Burst Fracture / 1902
Seatbelt Injuries (Flexion Distraction Injuries) / 1902
Fracture Dislocations of the Spine / 1903
Disc Herniation / 1903
Cauda Equina Syndrome / 1903
Spinal Stenosis / 1903
Back Pain and Degenerative Disc Disease / 1904
Scoliosis / 1904
Idiopathic Scoliosis / 1904
Neuromuscular Scoliosis / 1904

Joint Reconstruction 1904
Introduction to Arthritis / 1904
Examination of the Patient / 1904
Nonoperative Management and Prevention of Arthritis / 1905
Injections / 1905

Surgical Management of Arthritis / 1906
Computer Navigation, Robotics, and Joint Arthroplasty / 1908
Fixation Options in Joint Arthroplasty / 1909
Complications in Joint Arthroplasty / 1910

Orthopedic Pathology and Oncology 1910
Diagnosis of Malignant Bone Tumors / 1910

Osteosarcoma 1911
Intramedullary Osteosarcoma / 1911
Parosteal Osteosarcoma / 1912
Periosteal Osteosarcoma / 1912
Paget’s Sarcoma / 1912
Radiation-Induced Sarcoma / 1912

Ewing’s Sarcoma 1912

Cartilage-Forming Tumors 1912
Chondrosarcomas / 1912

Fibrous Lesions of Bone 1912
Desmoplastic Fibroma / 1912
Malignant Fibrous Histiocytoma of Bone / 1913
Malignant Vascular Tumors / 1913

Miscellaneous Tumors 1913
Giant Cell Tumor of Bone / 1913
Adamantinoma and Osteofibrous Dysplasia / 1913
Primary Lymphoma of Bone / 1914
Chordoma / 1914
Multiple Myeloma / 1914

Metastatic Bone Tumors 1914

Pediatric Orthopedics 1915
Birth Injuries / 1915
Skeletal Growth / 1915
Pediatric Fractures / 1916
Classification of Growth Plate Injuries / 1916
Diaphyseal Injuries in a Pediatric Patient / 1916
Fractures of the Pediatric Hip / 1916
Fractures of the Femoral Shaft / 1916
Pediatric Ankle Fractures / 1917
Pediatric Elbow Fractures / 1917
INTRODUCTION

Every physician should be familiar with orthopedics and orthopaedic surgery. Anyone who cares for patients in an outpatient or emergency room setting will find that the majority of presenting complaints involve the musculoskeletal system. A basic understanding of the principles of care for musculoskeletal conditions is essential for the healthcare profession.

For physicians, the field of orthopedics offers an array of subspecialties with such diversity that it seems that “there is something for everyone.” Trauma specialists have the satisfaction of physically putting complex fractures back together. Sports medicine offers remarkably rapid recovery in athletes who have suffered fibrocartilage and ligament tears with ever-improving arthroscopic techniques and instrumentation. Spine surgeons see remarkable results from their minimally invasive microscopic techniques, while also managing massive deformities with new instrumentation and open surgery. Joint reconstruction is one of our most exciting subspecialties, working with orthopedic bioengineers to develop improved designs, biomaterials, and minimally invasive surgical approaches for faster return to function for patients debilitated by arthritis and injury. Musculoskeletal oncology offers the intellectual challenge of arriving at appropriate differential diagnoses as well as the technical challenge of limb salvage and major reconstructive surgery. Pediatric orthopedics is an especially challenging and rewarding subspeciality because of the remarkable ability of children to heal their fractures quickly and remodel their bones. The incredible array of congenital and developmental disorders makes pediatrics a uniquely intellectually challenging field as well. The authors hope that our readers will share our enthusiasm for orthopedic surgery and all of its subspecialties: trauma, sports, spine, joint replacement, musculoskeletal oncology, foot and ankle, hand, and pediatric orthopedics.
ORTHOPEDIC TRAUMA

Introduction

Musculoskeletal injuries resulting from trauma include fractures of bones, damage to joints, and injuries to soft tissues. Long bone fractures can be described as transverse, oblique, spiral, segmental, or comminuted (Fig. 43-1). The goals of treating musculoskeletal injuries are to restore the normal anatomy, stabilize fractures to allow early mobility and minimize complications related to multiple system trauma, and to repair or reconstruct these injuries to restore function.

Fractures frequently result from high-energy trauma as well as from falls onto an extremity (Fig. 43-2). The majority of fractures can heal well with immobilization, which stabilizes the fracture while new bone forms at the fracture site. Methods of immobilization can vary and depend on the fracture being treated. The most common tool used in orthopedics to treat fractures is immobilization with a splint, cast, or braces, and their proper application is important to successfully treat the injury without causing additional problems. A successful splint contains adequate padding on the underlying skin, particularly over bony prominences, to prevent pressure or burns that can be caused by plaster. Splints, which are not circumferential, are preferred for acute injuries because they allow room for swelling that inevitably occurs after a fracture. The splint may later be changed to a cast as the swelling subsides.

Fractures that are displaced or angulated require closed reduction to properly realign the bone. This is done using analgesia, local or general anesthesia, and often muscle relaxation. Reduction is performed with axial traction and reversal of the mechanism of injury in order to restore length, rotation, and angulation. A splint is then applied and can be gently molded to help hold the reduction in place. It is important to obtain X-rays after a closed reduction to verify acceptable alignment of the fracture, and to perform a neurovascular exam to ensure the splint is not too tight or that manipulation did not change the neurovascular status. Careful monitoring with timely clinical and radiological examination is necessary in the outpatient setting.

For certain fractures, splint or cast immobilization is inadequate, and in these instances internal fixation or external fixation is used. A variety of implants can be used to stabilize the fracture such as screws, plates, rods, and external fixators. The main principle of orthopedic implants for fracture care is to create a stable construct that will allow the fracture to heal in proper length, alignment, and rotation. Screws can be placed across a fracture to create compression at the fracture site, which promotes healing. Plates can be placed on the cortex of bones and held with screws, which creates a large area of fixation to stabilize the fracture. Intramedullary rods are commonly used for long bone fractures, such as the femur and tibia (Fig. 43-3A,B). Usually, the fracture is reamed prior to the insertion of the rod into the intramedullary canal. Screws can then be placed across the cortices of the bone through holes in the rod proximal and distal to the fracture to create a locked construct that further stabilizes the rod. When the fracture is locked proximally and distally, this is called static locking. In situations where patients are severely injured and cannot safely undergo more invasive open surgery, damage control orthopedics is done by utilizing an external fixator. External fixators are also used when the soft tissues are too swollen or injured to allow for surgical incisions to be safely made. The wrinkle test is helpful in guiding the most suitable time for definitive surgery. An external fixation device can be used to temporarily immobilize the fracture, especially if the fracture is open and contaminated. External fixators involve pins placed into bone proximal and distal to the fracture (through healthy tissues), which are then connected by strong rods on the outside of the extremity, creating a stable construct.

Open Fractures

An open fracture occurs when the bone breaks through the skin. These typically result from high-energy injuries and are often associated with significant damage to the surrounding soft tissues and contamination of the wound. These injuries are classified into three types according to the Gustillo-Anderson Classification.

- Type I injuries are low energy and wounds are usually less than 1 cm.
- Type II injuries have a wound length of 2 to 10 cm with moderate soft tissue damage and wound contamination.
• Type III injuries are high-energy wounds usually greater than 10 cm in length with extensive muscle devitalization. The wound is highly contaminated with extensive soft tissue damage. These injuries require immediate administration of antibiotics and irrigation and debridement of the wound. The goal of the treatment is to achieve fracture healing and to prevent wound infections and osteomyelitis. They are frequently associated with injuries to surrounding vessels and nerves, which must be addressed as well. When the wound is contaminated, an external fixator is initially used (Fig. 43-4A,B). Often, definitive treatment of the fracture is delayed until the wound is

![Figure 43-3](image1.png)  
**Figure 43-3.** A. Transverse femur fracture. B. Intramedullary rod stabilizes femur fracture.

![Figure 43-4](image2.png)  
**Figure 43-4.** A. Gustillo-Anderson fracture type III open fracture. B. Image of external fixator of the tibia.
sufficiently cleaned and healthy soft tissue is available to cover the fracture. Early coverage of the wound is important to avoid infection. Usually a large wound in the proximal or middle third of the tibia can be covered using local muscle flaps, while the distal third of the tibia will require a free flap. In general, an increase in Gustillo grade is associated with an increase in infection risk.

**Compartment Syndrome**

Compartment syndrome is an orthopedic emergency caused by significant swelling within a compartment of an injured extremity that jeopardizes blood flow and microcirculation to the limb. Increased pressure within the compartment compromises perfusion to muscles and nerves and can cause ischemia or necrosis. Patients complain of pain that is greater than expected for the injury or surgery. There may be an increase in analgesic requirements. Early high index of suspicion is necessary for timely diagnosis and treatment of compartment syndrome. The usual clinical findings are pain, swelling, and pain with passive stretch of the compartment muscles. Numbness, paralysis, and the absence of a pulse are late findings. While the diagnosis is usually based on clinical exam, compartment pressures can be measured with needles placed into the compartment, which is necessary in unconscious patients and those who will not cooperate with the exam. Compartment pressure within 30 mmHg of the diastolic pressure is diagnostic of compartment syndrome. When compartment syndrome is suspected, emergent fasciotomy (Fig. 43-5A,B) must be performed in which the overlying tight fascia is released through long incisions. Fasciotomy must be done as soon as possible to prevent damage to muscles and nerves that will result in irreversible necrosis and Volkmann’s ischemic contractures with severe loss of function.

**TREATMENT OF FRACTURES AND DISLOCATIONS**

**Clavicle Fractures**

Fractures of the clavicle are one of the most common fractures in orthopedics. They typically occur following a fall onto the shoulder. The majority of clavicle fractures occur in the middle third of the clavicle. Since the clavicle is subcutaneous, the fracture is often evident on inspection. Most clavicle fractures can be treated nonoperatively with a sling, early range of motion exercises, and gradual return to normal activities. Fractures that are significantly displaced and shortened or that penetrate or tent the skin are treated with open reduction and internal fixation, typically with plate and screw fixation.

Distal clavicle fractures are less common and may occur with coracoclavicular ligament ruptures. These injuries can be more troublesome and are at risk for nonunion if the bone ends are not in contact. If there is displacement of the fracture and the fracture is proximal to the coracoclavicular ligament, surgical management is often recommended.

Acromioclavicular (AC) joint injuries occur from either a fall directly onto the shoulder or onto an outstretched hand and can result in tears of the acromioclavicular and coracoclavicular ligaments. A step-off, or separation, of the AC joint may be apparent on radiographs. The majority of these injuries can be treated with a sling and gentle range of motion. Although controversial, injuries resulting in severe displacement of the clavicle at the AC joint usually require open reduction and surgical repair, especially in athletes and manual workers.

The sternoclavicular (SC) joint is the only articulation between the upper extremity and the axial skeleton. Injuries to this joint are rare. Anterior dislocations occur more frequently, and although closed reduction can be attempted, recurrence of the dislocation is typical. Patients are given a sling and the outcome is usually good, despite the visible bump and swelling. Posterior SC joint dislocations are rare and not grossly visible and can be easily missed. They can be dangerous injuries, resulting in pulmonary or neurovascular compromise. Therefore, closed or open reduction under general anesthesia is recommended with a cardiac surgeon back-up.

**Scapula Fractures**

Fractures of the scapula often result from significant high-energy trauma (Fig. 43-6) with about 80% associated injuries,
most commonly involving the head, ribs, and lungs. Pulmonary injuries occur in over one-third of patients. Most scapula fractures are treated nonoperatively with a sling and early range of motion. Surgery is performed when there is involvement of the glenoid with a major articular step-off or if there is a glenoid rim fracture with subluxation of the joint.

**Shoulder Dislocations**

The shoulder is the most commonly dislocated large joint. Most dislocations are anterior. They are often associated with injuries to the anterior inferior glenoid labrum (Bankart lesion), impaction fractures of the humeral head (Hill-Sachs lesion) (Fig. 43-7), and rotator cuff tears in the elderly. The axillary nerve is at risk of being injured in shoulder dislocation. If the patient is unable to raise the arm after reduction of shoulder dislocation, then it is most likely due to a rotator cuff tear in the elderly and axillary nerve injury in the young.

There is a high recurrence rate that correlates with the age of the patient at the time of dislocation. There is a 90% redislocation rate if the patient is younger than 20 years of age. Posterior dislocations are associated with seizures or electric shock. Adequate radiographs are required to diagnose a shoulder dislocation, with the axillary view being the most important. The patient’s shoulder is usually locked in internal rotation with limitation of external rotation and axillary view will show the posterior dislocation (Fig. 43-8A,B). If proper X-rays are not performed, then dislocations can be missed and can result in significant disability to the patient. A computed tomography (CT) scan should be performed if an axillary view is unable to be obtained. In general, dislocation of the shoulders can be managed with closed reduction followed by a short period of sling immobilization.

**Proximal Humerus Fractures**

Proximal humerus fractures occur most frequently in elderly female patients following a fall onto the shoulder, though they can also occur following high-energy trauma in young patients. They have historically been classified by the number of fracture fragments using Neer’s classification (Fig. 43-9), which divides the proximal humerus into four parts: the humeral head, greater tuberosities, lesser tuberosities, and the humeral shaft. Treatment is determined by the displacement of the fracture fragments, the amount of angulation of the fracture, and the amount of comminution (which means multiple fracture fragments). If there is suspicion of an intra-articular fracture, CT scan is often indicated. The majority of proximal humerus fractures are minimally displaced and can be treated with sling immobilization, followed by early shoulder motion and pendulum exercises. Physiotherapy should be started within 2 weeks of the injury to prevent stiffness, especially in the elderly. Displaced fractures and fractures involving the humeral head are at increased risk for osteonecrosis, and therefore surgery is often recommended. If there is adequate bone stock and the fracture can be successfully reduced, open reduction internal fixation with plate and screw fixation is the treatment of choice. Older patients with osteoporosis, comminuted fractures, head-splitting fractures, and four-part fractures or fracture dislocations are typically treated with a prosthetic replacement of the humeral head or a hemiarthroplasty. Reverse shoulder arthroplasty is gaining popularity in the elderly as well.

**Humeral Shaft Fractures**

The majority of humeral shaft fractures can heal with nonsurgical management if they are within an acceptable degree of angulation. The radial nerve spirals around the humeral shaft and is at risk for injury; therefore, a careful neurovascular exam is important. If you have a patient with a humeral shaft fracture, check the patient for wrist drop (Fig. 43-10). Most radial nerve injuries are neurapraxias, or stretching of the nerve, and function typically returns within 3 to 4 months. A spiral fracture of the distal one-third of the humeral shaft is commonly associated with neurapraxia of the radial nerve, and this fracture is called a Holstein-Lewis fracture. Humeral shaft fractures are typically treated with a coaptation splint or functional bracing, which consists of a plastic clamshell brace with Velcro straps. Criteria for acceptable alignment are less than 20° anterior angulation, less than 30° varus-valgus angulation, and less than 3 cm shortening. Radial nerve palsy is not a contraindication to conservative treatment. Close follow-up with serial radiographs
is important to verify healing of the fracture, and gentle motion exercises are begun within 1 to 2 weeks. Fractures with significant angulation are most commonly treated with open reduction and plate fixation, with care to protect the radial nerve as it often lies close to the fracture site. Intramedullary nailing can also be performed, though it carries the risk of shoulder pain from the nail insertion. A plate is usually more stable than a nail and allows early weight-bearing through the humerus. Spontaneous recovery of radial nerve palsy can occur up to 6 months after injury. The patient should have an EMG to monitor recovery of the nerve. In an open fracture of the humeral shaft with radial nerve palsy, the nerve should be explored for the possibility of a significant nerve injury or laceration.

**Distal Humerus Fractures**

Fractures of the distal humerus result from falls onto the elbow or onto an outstretched arm. Supracondylar fractures occurring above the elbow joint are most common and do not involve the articular surface. Minimally displaced fractures can occasionally be treated with a posterior long arm splint, with the elbow typically flexed to 90°. However, fixation is often recommended to allow early range of motion and prevent stiffness. Fractures involving the articular surface are treated with plate fixation, and depending on the fracture pattern they may require more than one (usually anatomically contoured) plate. As with other intra-articular fractures, the goals of treatment are anatomic reduction of the joint surface with stable fixation, restoration of the anatomic alignment of the joint, and early range of motion. Severely comminuted fractures, especially in the elderly, may be treated with a total elbow replacement. Fractures about the elbow are notorious for developing stiffness and therefore early motion of the elbow is paramount to a successful outcome. Range of motion should be started as soon as the patient can tolerate therapy.

**Elbow Dislocations**

Dislocations of the elbow are common and typically occur posteriorly after a fall on an outstretched hand. A dislocation results in injury to the joint capsule and rupture of the lateral collateral ligament, with possible involvement of the medial collateral ligament, as well as possible fractures of the radial head and coronoid. This combination of injuries is called the “terrible triad,” which is a challenging injury and carries the
worst prognosis. Simple elbow dislocations should be urgently reduced with the patient under sedation and treated with a short period of immobilization, utilizing a posterior splint. Stiffness of the elbow is a common complication following elbow dislocations and therefore only short-term immobilization (about 7–10 days) followed by early range of motion is recommended.

Dislocations associated with fractures may be treated surgically if there is any instability of the elbow joint. The “terrible triad” is an unstable injury comprising of an elbow dislocation as well as fractures to the radial head and coronoid, which requires surgery. Surgery includes repair of the torn lateral collateral ligament, fixation or replacement of the radial head, and possible fixation of the coronoid, depending on the size of this fracture fragment.

Radial Head Fractures
Most fractures of the radial head can be treated nonoperatively, simply with a sling for 1 to 2 days followed by motion exercises. Surgery is recommended if there is a displaced fracture, if the fracture blocks pronation or supination of the forearm, if there is an associated dislocation of the elbow, or if the patient has associated wrist pain (Essex-Lopresti fracture). Surgery can be fixation or replacement. If the fracture can be well reduced, it is fixed with 1 or 2 screws. If the radial head is fractured into multiple pieces, the treatment of choice is a radial head replacement with a metallic implant. Simple excision of the radial head can also be performed in low demand patients with an isolated radial head fracture; otherwise, it may lead to instability of the elbow and the wrist over time.

Olecranon Fractures
Olecranon fractures usually occur following a fall directly onto a flexed elbow (Fig. 43-11). Nondisplaced fractures are treated with a splint in 45° to 90° of flexion for a short time followed by range of motion exercises to prevent stiffness. Because the triceps inserts on the olecranon, the pull of the muscle often causes active extension of the elbow and displacement of the fracture, and therefore the olecranon fracture should be fixed surgically. Simple transverse fractures can be fixed with a tension band construct, which consists of wire passing through the ulna, distal to the fracture, and wrapped in a figure-of-8 fashion around two or more pins placed proximally into the olecranon, crossing and stabilizing the fracture. This tension band construct creates a compressive force across the articular aspect of the fracture that will promote healing. Fractures that are comminuted or have large fragments are usually treated with plate and screw fixation. Excision of the olecranon with advancement of the triceps can be done in elderly patients when the fracture involves less than 50% of the joint surface. Because of the subcutaneous location of the olecranon, symptomatic hardware is a frequent complication, causing irritation to the patient; it may need to be removed after the fracture has healed. Stiffness of the elbow is another complication seen in a large number of patients.

Forearm Fractures
Forearm fractures are common injuries that result from high-energy trauma or from falls onto an outstretched arm. Both bone forearm fractures generally require surgery with plate and screw fixation. The radius has a bow and rotates around the straight ulna for proper pronation and supination of the forearm, and therefore this anatomic relationship needs to be restored to maintain function. An isolated fracture of the ulna shaft, or a “nightstick fracture,” occurs from a direct blow to the side of the forearm. These can usually be treated in a cast, splint, or brace. Fractures that are angulated or displaced can be treated with open reduction and plate fixation. A Monteggia fracture is a fracture of the proximal third of the ulna associated with a radial head dislocation. The radial head dislocation may be missed. Careful evaluation of the radiograph, especially the relationship with the radial head to the capitellum is necessary for the diagnosis of this injury. These fractures are common in children and rare in adults. These injuries require surgery to fix the ulna fracture with plate and screw fixation and to reduce the radial head dislocation. A Galeazzi fracture is a fracture of the distal third radial shaft associated with distal radioulnar joint (DRUJ) injury at the wrist. If the fracture of the radius is less than 7.5 cm from the joint, the distal radioulnar joint is injured in a large number of cases. After the radius is fixed with plate and screw fixation, the DRUJ is assessed for stability and may need wires placed across the joint temporarily.

Distal Radius Fractures
Distal radius fractures commonly occur in older patients due to a fall or osteoporosis. In younger patients, these fractures usually occur due to high-energy trauma. A Colles fracture is a low energy fracture that is extra-articular and usually dorsally displaced. It has a characteristic appearance of a fork, naming the fracture the “dinner-fork” deformity. A Smith’s fracture is a reverse Colles fracture, usually extra-articular and volarly displaced. A Chauffer’s fracture involves the radial styloid process and may cause occult carpal disruption. A Barton’s fracture can be either volar or dorsal. It is a fracture dislocation of the radiocarpal joint, with an intra-articular volar or dorsal fracture.

Every attempt should be made to rule out fractures that extend intra-articularly into the wrist joint or involve the DRUJ. Patients should be evaluated for a median nerve injury and osteoporosis if suspected. Loss of thumb extension from extensor pollicus longus tendon rupture can occur especially in nondisplaced distal radius fractures. Treatment is often a closed

Figure 43-11. Displaced olecranon fracture.
reduction and immobilization. Surgery utilizing a variety of surgical techniques is done for unstable fractures as well as those with significant intra-articular involvement.

**Scaphoid Fractures**
Scaphoid fracture is the most common fracture of the carpal bone. Its diagnosis can be easily missed, and the fracture can lead to nonunion and avascular necrosis. It usually occurs in the waist of the scaphoid but can occur in the proximal or distal pole. Proximal scaphoid fracture will have a higher incidence of avascular necrosis due to interruption of the retrograde blood supply. Tenderness in the anatomic snuffbox after trauma should be considered a scaphoid fracture until proven otherwise. Magnetic resonance imaging (MRI) will be helpful in early diagnosis if no fracture is visible on an X-ray. A thumb spica cast is used for stable nondisplaced fracture, while reduction and screw fixation of the fracture is usually done for displaced fractures. The dorsal approach is used for proximal fractures, and the volar approach is used for the majority of other fractures.

**Pelvic Fractures**
Pelvic fractures are indicative of high-energy trauma and are associated with head, chest, abdominal, and urogenital injuries. Hemorrhage from pelvic trauma can be life-threatening and patients can present with hemodynamic instability, requiring significant fluid resuscitation and blood transfusions. The bleeding that occurs is often due to injury to the venous plexus in the posterior pelvis or from the fracture itself. It can also be due to a large vessel injury such as the superior gluteal artery at the greater sciatic notch. Immediate resuscitation with fluids and blood is critical. In hemodynamically unstable patients, blood, fresh frozen plasma, and platelets are given in a 1:1:1 ratio. These patients may require surgical exploration or interventional radiology embolization to stop the bleeding. An important first-line treatment in the emergency department is the application of a pelvic binder or sheet that is wrapped tightly around the pelvis to help control bleeding. This is important when there is an increase in the volume of the pelvis by the anteroposterior compression mechanism (an open book mechanism). The pelvic binder is clearly the initial management of an unstable open book fracture of the pelvis with bleeding. Traction pins may be applied in the operating room, but it is less frequently used. Other associated injuries are bladder and urethral injuries that manifest with bleeding from the urethral meatus or blood in the urinary catheter, and these need to be assessed with a retrograde urethrogram.

The pelvis is a ring structure made up of the sacrum and the two innominate bones that are held together by strong ligaments. Because it is a ring, displacement can only occur if the ring is disrupted in two places. This may occur either from fractures of the bones or tears of the ligaments that can cause dislocation. When you see an anterior fracture of the ring, check for a posterior injury (Fig. 43-12). There are three main fracture patterns that occur from trauma to the pelvis. An anteroposterior force to the pelvis causes an “open book” injury pattern in which the pelvis springs open, hinged on the intact posterior ligaments with widening of the pubic symphysis. A lateral compression pattern results from a crush injury that causes fractures to the ilium, sacrum, and pubic rami. Vertical shear injuries are very unstable since they result from disruption of the strong posterior pelvic ligaments and are associated with significant blood loss and visceral injuries. Fractures of the sacrum may be difficult to see on X-ray, and therefore CT scans are often needed to completely visualize the fracture pattern. The sacral nerves pass through foramen in the sacrum, and therefore fractures that are close to the foramen can result in nerve injuries. Fractures that involve the sacral canal have a high incidence of nerve injuries and cauda equina syndrome. Fractures that involve the ala of the sacrum may involve the L5 nerve root. Vertical fractures of the

![Figure 43-12.](image-url) **A.** Pelvic fracture showing anterior and posterior disruption of the pelvis. **B.** Image depicting a vertical shear fracture with cephalad migration of the hemi-pelvis.
sacrum can be highly unstable even after fixation and may be associated with sacral nerve root injuries.

Treatment of pelvic fractures depends on the fracture pattern. Stable, minimally displaced fractures such as many lower energy lateral compression fractures can be treated nonoperatively with protected weight-bearing. Open book injuries in which the pubic symphysis is widened more than 2.5 cm may require an anterior plate, and if the posterior pelvic ligaments are also injured, the patient will need posterior fixation. Posterior stabilization is typically performed with screws placed percutaneously through the ilium into the sacrum to stabilize the pelvis posteriorly, and a plate is applied over the pubic symphysis for anterior stabilization. Displaced sacral fractures and iliac wing fractures are treated with screws or plates, while pubic rami fractures can usually be managed nonoperatively. While most pelvic fractures are caused by high-energy trauma, elderly patients with osteoporotic bone can also suffer pelvic fractures after a fall, usually fracturing the pubic rami. Since these are stable injuries, they can be managed nonoperatively with protected weight-bearing.

**Acetabular Fractures**

The acetabulum forms the socket of the hip joint, and fractures occur when the femoral head is driven into the acetabulum in the setting of high-energy trauma. Sciatic nerve function should be examined carefully after an acetabulum fracture. It is important to rule out dislocation of the hip, which should be reduced immediately to prevent avascular necrosis of the femoral head. Usually 45° oblique views, called Judet views, are utilized. CT scans are very important to visualize the fracture pattern. According to Judet and Letournel, there are ten acetabular fracture patterns: five simple and five complex fracture types (Fig. 43-13). These fractures often require surgery in order to obtain anatomic reduction and to minimize the development of degenerative arthritis.

**Hip Dislocations**

Hip dislocations almost always result from high-energy trauma; they most commonly occur posteriorly and less commonly anteriorly (Fig. 43-14). They can cause injury to the sciatic nerve, which runs directly posterior to the hip joint. Examine the patient for foot drop and numbness at the top of the foot. Hip dislocation can be simple, or it may be associated with a fracture of the acetabulum or femoral head. Hip dislocations need to be emergently reduced because of the risk of osteonecrosis of the femoral head if the reduction is delayed. Closed reduction is usually successful with adequate sedation or under general anesthesia. Once reduction is done, a CT scan is ordered to define the extent of the injury. A CT scan will show associated fracture patterns, trapped intraarticular fracture fragments, and the congruity of the reduction. If the reduction is unsuccessful, or if there is a fracture fragment inside the joint, then an open reduction is indicated. Hip dislocations that are associated with a femoral head fracture are at increased risk for osteonecrosis of the femoral head and posttraumatic osteoarthritis. The femoral head fracture associated with hip dislocation is called a Pipkin fracture. If the dislocation is associated with posterior wall fractures, the stability of the hip joint should be assessed carefully, even if the fragment is small. This is usually done by an examination of the patient under anesthesia.

**Hip Fractures**

Hip fractures are an extremely common injury seen in orthopedics and are associated with significant morbidity and mortality. They most often occur in elderly patients after ground level falls, are much more common in women than men, and occur more commonly in patients with osteoporosis. The three most common fractures in the elderly are those of the wrist, spine, and hip. Patients who suffer hip fractures are at increased risk for many complications, including deep vein thrombosis, pulmonary embolism, pneumonia, deconditioning, pressure sores, and even death. The mortality rate in the first year following a hip fracture is around 25%. One of the most important reasons for performing surgery is to prevent these complications because getting patients out of bed and walking as soon as possible diminishes their risk for many of these adverse events. Performing early surgery also decreases the complications in these patients. Therefore, surgery is almost always the treatment of choice for hip fractures. The type of surgery performed is determined by the anatomic location of the fracture and the fracture pattern. Surgery should be performed as soon as possible, typically within 24 to 48 hours; however, since many of these patients suffer other comorbidities, they must be properly medically optimized before surgery. The goals of surgery are to minimize pain, restore hip function, and allow early mobilization, the importance of which cannot be overemphasized. The functional outcome for patients following a hip fracture is largely based on their level of mobility and independence.

![Figure 43-13. Types of acetabular fractures.](image1)

![Figure 43-14. Posterior and anterior dislocation.](image2)
before their injury. Many patients become less independent, may require assistive devices to help them walk, and some may require a long-term nursing or rehabilitation facility. Hip fractures can be femoral neck fractures, intertrochanteric fractures, or subtrochanteric fractures (Fig. 43-15).

**Femoral Neck Fractures.** Femoral neck fractures occur within the capsule of the hip joint. The main blood supply to the femoral neck and head comes from the deep branches of the medial femoral circumflex arteries, which run along the femoral neck, and when the fracture is displaced, there is an interruption in the blood supply of the femoral head, which can lead to osteonecrosis. Femoral neck fractures that are nondisplaced have a low risk of disruption of blood flow and therefore can be treated with in situ internal fixation. Three partially threaded cancellous screws are placed through a small incision over the lateral proximal femur, directed through the femoral neck and into the femoral head. Patients can usually begin protected weight-bearing immediately after surgery. Displaced femoral neck fractures will likely disrupt the blood supply and therefore need to be treated with a prosthetic replacement in older adults. Most commonly a hemi-arthroplasty is performed in which the femoral head and neck are replaced with a metal head and neck into the femoral canal. Higher demand patients and those who have osteoarthritis of the hip joint and hip pain before their fracture may receive a total hip replacement, in which the acetabulum is also replaced with a prosthesis, typically a plastic cup inside a metal shell. Patients can begin weight-bearing immediately after surgery. Displaced femoral neck fractures in young patients are the result of a high-energy trauma and are usually treated by reduction with screw fixation. The reduction may be closed or open.

**Intertrochanteric Hip Fractures.** Intertrochanteric hip fractures occur between the greater and lesser trochanters of the proximal femur. Because the blood supply to this area is abundant, osteonecrosis is uncommon, and therefore these fractures can be treated with reduction and internal fixation. Displaced fractures need to be realigned, and this often involves placing the patient on a fracture table where traction and rotation can be applied to the affected leg to reduce the fracture. There are two devices that can be used. In stable fractures, a sliding hip screw includes a large screw placed from the lateral cortex of the proximal femur across the fracture and into the femoral neck and head, followed by a side plate along the lateral cortex of the femur, which is then fixed to the shaft with screws. A cephalomedullary nail includes a nail placed down the medullary canal of the femur and a large screw that engages the nail as it is passed from the lateral cortex up into the neck and head. Nails are usually used in unstable fractures and allow protected weight-bearing postoperatively. The reverse oblique intertrochanteric fracture is a specific type of fracture that exits on the lateral cortex (Fig. 43-16). This is best treated with a cephalomedullary

![Figure 43-15. Types of hip fractures.](image1)

![Figure 43-16. Classic intertrochanteric fracture and reverse oblique fracture. Notice that the fracture line of the reverse oblique fracture exits on the lateral cortex.](image2)
nail; a dynamic hip screw is the wrong device to be used in reverse oblique fractures because it will lead to sliding, shortening, and medial displacement of the fracture.

**Subtrochanteric Hip Fractures.** Subtrochanteric hip fractures occur in the proximal femoral shaft just distal to the lesser trochanter in an area of high biomechanical stresses. While they can occur in older adult patients after a fall, they are also seen in high-energy trauma. Because of the forces of muscles attached to the fractured segments, they tend to be significantly displaced (Fig. 43-17) and may be difficult to reduce. They are most often treated with a long cephalomedullary nail that includes a screw distally to lock the nail in place and prevent rotation of the femur. Fixed angle plates or blade plates are sometimes used in the treatment of subtrochanteric fractures. In most cases, protected weight-bearing can begin soon after surgery. Complications usually include malunion and nonunion of the fracture.

Bisphosphonate-related subtrochanteric fractures are an example of insufficiency fractures that may be related to the long-term use of bisphosphonates. These fractures have been recently identified. An intramedullary nail is the treatment of choice for this fracture.

**Femoral Shaft Fractures**

Fractures of the femoral shaft are caused by high-energy trauma and may be associated with other severe injuries. Long bone fractures, such as femoral shaft fractures, put these patients at risk for complications such as thromboembolic events and acute respiratory distress syndrome (ARDS), and therefore it is important to fix these quickly, typically within 24 hours. They are most commonly fixed with an intramedullary nail that can be placed antegrade (from the piriformis fossa or greater trochanter down the canal) or retrograde (through an incision into the knee joint and up the canal), with screws placed through proximal and distal holes to lock the nail in place, creating a stable construct to allow weight-bearing. Trauma patients who are hemodynamically unstable or who have other life-threatening injuries are treated temporarily with an external fixator until they can safely undergo surgery. This is called “damage control orthopedics.” The base deficit and lactic acid levels are monitored and used as guides to indicate if the patient is adequately resuscitated. When their levels are normal, it means the tissue is adequately oxygenated and the patient can undergo definitive fixation of the femur.

**Distal Femur Fractures**

Distal femur fractures are the result of a fall from a height or from high-energy trauma. They can also occur in elderly patients with osteoporotic bone after a fall onto the knee. While nondisplaced fractures in the elderly may be treated nonoperatively with a hinged knee brace and early motion exercises, most require surgery. These fractures can involve the articular surface of the knee joint, so anatomic reduction of the joint surface is crucial. They are fixed with plates and screws, often utilizing a locking construct. The plate is placed over the lateral, or rarely the medial cortex depending on the fracture pattern. A retrograde intramedullary rod inserted through the knee can also be used, especially in extraarticular fracture patterns. The goal of surgery is to achieve anatomic reduction, stable fixation, and allow early knee range of motion. Intra-articular fractures require the patient to be non-weight-bearing until the fracture shows signs of healing. Complications of these fractures include nonunion, malunion, and stiffness of the knee. Be aware of Hoffa fractures, a coronal fractures that usually involve the lateral femoral condyle. They can be missed on X-rays, but they are easily diagnosed by CT scan. It may need a different fixation than that required for the associated supracondylar fracture component.

**Knee Dislocations**

Dislocation of the knee is a rare but devastating injury that can be limb-threatening. Some dislocations spontaneously reduce and can be underdiagnosed. When the knee dislocates, the anterior cruciate ligament (ACL) and posterior cruciate ligament (PCL) are torn, and various degrees of injury occur to the lateral collateral ligament (LCL), medial collateral ligament (MCL), posterolateral corner, joint capsule, and menisci. However, the danger is due to the close proximity of the popliteal artery, which runs directly behind the knee and may kink or sustain a tear of the intimal wall when the knee dislocates. A neurovascular exam is extremely important, focusing on the common peroneal nerve and the vascular status of the extremity, followed by immediate reduction of the knee and repeat neurovascular exam. If the pulses are normal, the ankle brachial index (ABI) should be measured. If the ABI is more than 0.9, then the patient should be monitored with serial examination. If the ABI is less than 0.9, then a CTA or an arterial duplex ultrasound should be performed. If there is evidence of diminished pulses after reduction, an angiogram must be performed. If the pulses are absent after reduction, immediate surgical exploration and/or repair should be done by a vascular surgeon. Prophylactic fasciotomy of the leg is usually done. Time is critical to reestablish the circulation of the limb. If ischemia time is more than 8 hours, then there is a very high rate of amputation. With regard to the ligamentous injuries, an external fixator may be initially used to stabilize the unstable knee and protect the reduction. Subsequently, an MRI will identify what structures have been torn. Because a dislocation causes so much damage to the knee, a delayed multiligamentous reconstruction is recommended on an elective basis in order to stabilize the knee joint.
Stiffness and instability of the knee are common complications after this injury.

**Patella/Extensor Mechanism Injuries**
The extensor mechanism is comprised of the quadriceps tendon, the patella, and the patella ligament. This mechanism functions to extend the knee. Injuries can result after a fall directly onto the knee or from forcible contraction of the quadriceps. It is important to examine the knee for the ability to actively extend the knee. Quadriceps tendon ruptures, patella fractures, or patella ligament ruptures can result in a loss of active knee extension requiring surgery. Nondisplaced patella fractures with intact active knee extension can be treated nonoperatively with a cast or knee immobilizer, holding the knee in full extension, and weight-bearing is permitted. Displaced or comminuted fractures require surgery with tension band wiring and/or screws. Symptomatic hardware is a common complication. Acute osteochondral fractures can be managed with internal fixation. Quadriceps tendon and patella tendon ruptures with loss of active knee extension are treated with suture repair. After surgery, the knee is held in extension, and knee flexion is slowly increased over several weeks using a hinged knee brace.

Patella dislocations are common injuries that occur when the femur is forcibly internally rotated on an externally rotated tibia while the foot is planted on the ground. They typically dislocate laterally and often relocate spontaneously. The medial patellofemoral ligament is the primary stabilizer of the patella. Patients present with a significant knee effusion and medial-sided tenderness. During the physical exam, these patients may elicit a positive apprehension test, in which a lateral force to the patella elicits pain and the sensation of an impending dislocation. Dislocated patellas can be reduced by extending the knee and manual reduction and are treated with temporary knee immobilization. Make sure that there is no fracture or loose bodies, which would be an indication for surgery. MRIs will show the classic bone bruise and edema involving the medial facet of the patella and the lateral condyle of the femur. There is a high risk of recurrent dislocation with nonoperative treatment, which may require surgical intervention.

**Tibial Plateau Fractures**
The tibial plateau is comprised of the articular surfaces and underlying cancellous bone of the medial and lateral plateaus of the proximal tibia. Fractures of the plateau result from axial loads sustained in falls from a height or high-energy trauma, and they are often associated with injuries to the menisci and cartilage of the knee. Fractures can involve the medial, lateral, or both plateaus with significant comminution, angulation, and depression, creating a challenging injury to fix. The Schatzker classification is commonly used in tibial plateau fractures (Fig. 43-18).

- Type I: Lateral split fracture
- Type II: Lateral split-depressed fracture
- Type III: Lateral pure depression fracture
- Type IV: Medial plateau fracture
- Type V: Bicondylar fracture
- Type VI: Metaphyseal-diaphyseal disassociation

Meniscal tears occur more on the lateral side and tend to be peripheral tears, especially if there is more than 6 mm depression or separation of the joint. Type IV, which is the medial tibial plateau fracture, could be a variant of a knee dislocation. The ankle brachial index (ABI) should be used in this situation and in more complex types of tibial plateau fractures. Clinically, laxity of more than 10° may indicate instability of the fracture; however, the test may be painful and hard to perform. A CT scan is important to visualize the intra-articular involvement of the fracture. Minimally displaced fractures may be treated nonoperatively with strict non–weight-bearing until the fracture heals. Fractures associated with displaced articular fragments require surgery in order to restore the smooth contour of the articular surface. They are treated with plates and screws placed medially, laterally, or both. Stabilization of a posteromedial fragment may require a separate posteromedial approach. Since there is often a depression of the cancellous bone, bone graft or bone substitutes, particularly calcium phosphate which resists compression, may be needed to buttress the articular surface and restore the anatomic alignment of the tibia. Patients are kept strictly non–weight-bearing for several months until the fracture begins to heal, though early range of motion is encouraged to prevent stiffness. Repair of ligament or meniscus injuries may also be indicated at the time of surgery. Knee stiffness and osteoarthritis are common complications of these injuries. The goal of the surgery is to restore joint stability and alignment.

**Tibial Shaft Fractures**
Tibial shaft fractures are the most common long bone fractures and occur following high-energy trauma, direct blows, and severe twisting injuries. Trauma and direct blows to the
tibia result in transverse or comminuted fracture patterns, while torsional injuries cause spiral fractures. Fractures with minimal angulation can be treated with reduction and casting, followed by transition to a functional brace and slow return to weight-bearing. Such fractures may need to be immobilized for several months since these fractures can be slow to heal. Most tibial shaft fractures, especially comminuted and angulated fractures, are treated with an intramedullary nail placed down the tibial canal, with interlocking screws placed proximally and distally. Weight-bearing can begin soon after surgery. Proximal third tibial fractures are challenging and can result in malalignment, usually valgus and apex anterior angulation. Knee pain is common after intramedullary rod placement. Plate and screw fixation can also be used; however, since the tibia is subcutaneous, hardware placed along the shaft can increase the risk of wound complications, making intramedullary nailing the preferred treatment. Fibula shaft fractures often occur along with tibial shaft fractures, though they usually heal well without surgery. Tibial fractures, both closed and open, can be associated with compartment syndrome. Patients usually have pain out of proportion with swelling of the leg and pain with passive stretch. Compartment pressure within 30 mmHg of the diastolic pressure is diagnostic of compartment syndrome.

Tibial Plafond (Pilon) Fractures
The tibial plafond is the distal tibial articular surface of the ankle joint. Pilon fractures are typically high-energy injuries that usually result from axial compression. These injuries can cause significant soft tissue injury, severely comminuted intra-articular and metaphyseal fragments (Fig. 43-19A,B), and wound healing problems, making these fractures very difficult to treat. Due to the soft tissue injury, these fractures are initially treated with external fixation until the swelling subsides, which may take several days to weeks. The wrinkle test is helpful in this situation to assess when the soft tissues are amenable to definitive fixation. A CT scan is usually obtained after the fracture is stabilized by an external fixator. The CT scan will clearly define the fracture fragments and helps in planning the surgical approach and fixation. Minimal incision techniques and minimal fixation are becoming popular in some situations. The main goal of surgery is to restore the articular surface. Fixation of the fibula in order to maintain and establish anatomic length is done in some cases. Bone grafts or bone substitutes may be used to fill the void in the metaphyseal region. A variety of fixation techniques may be used including plates to stabilize the metaphysis to the diaphysis. Patients are kept non-weight-bearing for many months until the fracture heals. Despite best efforts, patients may suffer from ankle pain and stiffness, arthritis, wound healing problems, infection, nonunion, and some patients may eventually require ankle fusion. Early fixation of pilon fractures with plates can increase the incidence of wound complications significantly.

Ankle and Subtalar Dislocations
Ankle Dislocations. The ankle joint is a complex hinge joint comprised of the distal tibial plafond, medial malleolus, and lateral malleolus and their articulation with the talus. Several ligaments also contribute to the stability of the ankle joint, including the deltoid ligament medially, the syndesmotic ligaments between the tibia and fibula, and the anterior talofibular, posterior talofibular, and calcaneofibular ligaments laterally. Dislocations of the ankle joint result from a severe twisting injury and often occur with fractures. At times, dislocations

Figure 43-19. A. Tibial pilon fracture with comminution. B. Pilon fracture and its main fracture fragments.
can place significant pressure on the overlying skin and can cause neurovascular compromise; therefore, prompt reduction is extremely important followed by splinting.

**Subtalar Dislocations.** Subtalar dislocations can be medial or lateral, depending on the position of the foot. The medial dislocation is more common. Lateral dislocations are less common, can be open, and are more likely to be associated with fractures. Irreducible lateral subtalar dislocations may occur from a trapped tibialis posterior tendon, which will block the reduction. The main complication of subtalar dislocations is subtalar arthritis.

**Ankle Fractures**

Ankle fractures are very common and result from a twisting injury to the ankle. The patterns of ankle fractures depend on the direction of force and the position of the foot and ankle at the time of injury. The goals of treating ankle fractures are to restore the anatomy of the ankle joint and to restore the length and rotation of the fibula. Initial treatment includes closed reduction and placement of a well-padded splint in order to protect the skin. Swelling can be a significant problem, so elevation of the foot is encouraged. Surgery may be delayed until the skin condition permits. Fractures of the ankle may be:

1. Isolated malleolar fractures, usually the lateral malleolus or the medial malleolus.
2. Bimalleolar fractures that involve the lateral and medial malleolus. Please note that the deltoid ligament may be injured instead of the medial malleolus.
3. Trimalleolar fractures that involve the lateral malleolus, medial malleolus, and posterior malleolus.

**Lateral Malleolus Fractures.** Isolated fractures of the lateral malleolus with less than 3 mm displacement and no talar shift may be stable. An external rotation stress radiograph or a gravity test is used to assess the competency of the deltoid ligament, with the goal of the test to exclude deltoid injury. If the patient has a deltoid injury, in addition to the fibular fracture, then the patient will need surgery. The fracture will require anatomic reduction in order to restore normal ankle joint congruity. The talus can sublux laterally following lateral malleolus fractures, and even 1 millimeter of talar shift decreases the surface contact between the talus and the tibia by 40%, increasing the risk of developing arthritis. Open reduction and internal fixation of the fibula is usually done with plate and screws.

**Medial Malleolar Fractures.** An isolated fracture of the medial malleolus is usually an avulsion-type injury. Minimally displaced fractures can be treated with a cast or walking boot, while displaced fractures are usually fixed with screws.

**Bimalleolar Fractures.** Fractures to both the medial and lateral malleoli usually require surgery. These injuries are more unstable, and the talus will often sublux or completely dislocate laterally. They are treated by reducing and fixing both malleoli during surgery. Occasionally, the posterior articular surface of the distal tibia, or posterior malleolus, can be fractured as well, resulting in a trimalleolar ankle fracture. Often it is a small fragment and does not need to be fixed; however, if it involves a significant amount of the articular surface, the posterior malleolus should be fixed with screws placed either anteriorly, posteriorly, or with an antiglide plate. In all ankle fractures, especially in ones associated with deltoid rupture, a syndesmotic injury should be considered. After the fixation of the fractures, an intraoperative external rotation stress test will diagnose syndesmotic injury.

**Syndesmosis Injuries.** The syndesmosis is comprised of several ligaments between the distal tibia and fibula that provide stability to the ankle joint by resisting axial, rotational, and translational forces. The syndesmosis can be disrupted at the time of ankle fractures and requires special attention (Fig. 43-20). Widening of the space between the distal tibia and fibula after fixing the fractures is indicative of a syndesmosis injury, and an intraoperative external rotation stress radiograph can be helpful for evaluation. Such injuries are treated with one or two screws placed laterally from the fibula into the tibia, parallel to the ankle joint. Patients are kept non–weight-bearing for several weeks. The screws are often removed after 3 to 6 months, though they can be left in place and are typically asymptomatic.

**Maisonneuve Fractures**

A Maisonneuve fracture is a fracture of the proximal fibula associated with fracture of the medial malleolus or rupture of the deltoid ligament. There is always an associated syndesmotic injury (Fig. 43-21). Diagnosis may be difficult, and the injury may be missed. These injuries require surgical treatment with fixation of the syndesmosis by screws from the fibula to the tibia.

**Calcaneal Fractures**

Calcaneal fractures usually occur following a fall from a height and are often associated with other injuries, including lumbar spine fractures. There is a high incidence of compartment...
syndrome of the foot associated with calcaneal fractures. These injuries are often intra-articular and can result in collapse of the weight-bearing posterior facet of the calcaneus. The Bohler angle, which is normally between 20° and 40°, would be reduced or even flattened. CT scans are useful to better visualize the fracture pattern. Some fractures can be treated nonoperatively in a well-padded splint with patients being kept non-weight-bearing for up to 12 weeks. Displaced intraarticular fractures can be treated surgically once the swelling subsides and the wrinkle test is positive. Surgery can be done with lag screws or with plate and screw fixation. Despite adequate treatment, calcaneal fractures can be debilitating injuries, leading to significant heel pain and arthritis. The outcome of surgery depends on the comminution of the fracture and degree of articular involvement. The more fragments seen on a CT scan, the worse the outcome for the patient. Wound complications are a problem for calcaneal fractures with the risk being even higher in diabetics, smokers, and in open fractures. Open calcaneal fractures have a high rate of amputation.

Talus Fractures
Fractures of the talus commonly result from forced dorsiflexion of the ankle, causing the talar neck to impact on the anterior distal tibia. The dominant blood supply is the artery of the tarsal canal. The blood supply to the talus can be jeopardized after a displaced fracture and may lead to osteonecrosis (Fig. 43-22), which is an unfortunately common complication following talus fractures. The incidence of osteonecrosis depends on the degree of displacement of the fracture. The Hawkins sign is a subchondral lucency that is seen on the mortise X-ray at 6 weeks and indicates that there is vascularity of the talus. This indicates that there is no avascular necrosis. Nondisplaced fractures are treated with a cast and have a 15% risk of osteonecrosis, while displaced fractures are often treated surgically with screw fixation. There is a high risk of osteonecrosis, ranging from 30% to 100%. Subtalar arthritis is the most common complication from this injury. Varus malunion that results from inadequate reduction is the most preventable complication. This complication leads to a decreased subtalar range of motion and eversion. With varus malunion, the patient will walk with the foot internally rotated.

Foot Fractures
The tarsal bones, including the navicular, the cuboid, and the three cuneiform bones, link the hind foot to the metatarsals and provide mechanical stability to the arch of the foot. Isolated fractures to these bones are rare and are often treated nonoperatively with a cast or boot. Cuboid fractures are also known as “nutcracker fractures” and may indicate a Lisfranc injury. Stress fractures of the navicular can be occult, and a diagnosis may be challenging to make in a patient that complains of foot pain. An MRI may be needed for diagnosis. Treatment is often a short leg non-weight-bearing cast. The Lisfranc ligament, which connects the second metatarsal base to the medial cuneiform, is an important stabilizer of the midfoot. Lisfranc injuries can be seen following torsional forces to the foot or from crush injuries. These injuries can be missed and often require surgery because anatomic reduction is extremely important for a successful outcome. Open reduction and internal fixation is the technique used for Lisfranc fractures. Primary arthrodesis is often used for purely ligamentous injuries. The main complication of Lisfranc injuries is posttraumatic arthritis. Metatarsal fractures similarly result from twisting or crush injuries and most can be treated.

Figure 43-21. Maisonneuve fracture showing proximal fibular fracture and associated ankle injury with disruption of syndesmosis.

Figure 43-22. Displaced talar neck fracture. Notice the interruption of the blood supply in the talus.
nonoperatively with a hard-soled shoe and weight-bearing as tolerated. The base of the fifth metatarsal, however, warrants close attention. Fifth metatarsal fractures at the metaphyseal-diaphyseal junction (fourth and fifth metatarsal articulation) are called Jones fractures. These fractures can jeopardize blood flow, are at risk for nonunion, and may be associated with cavovarus hindfoot. Jones fractures need close follow-up to assess for healing if treated by short-leg cast and non-weight-bearing. In athletes and active young patients, screw fixation is usually used to stabilize the fracture. Injuries to the metatarsal-phalangeal joints and phalangeal fractures can be treated symptomatically or with buddy taping with weight-bearing as tolerated in a hard-soled shoe.

SPORTS MEDICINE

Introduction
Sports medicine deals with the prevention and treatment of injuries related to sports and exercise. These injuries encompass various areas in the musculoskeletal system. In recent years, sports-related injuries have increased, and the sports medicine field has been expanding. There are multiple factors leading to this increase in sports-related injuries. They include athletes participating in one sport year-round, more “weekend warriors” participating in sporting activity, and increased expectations for higher performance.

The orthopedic subspecialty of sports medicine treats a broad spectrum of patients, ranging from children who have just started participating in their first sports to the specialized care of professional athletes. Medical treatment of athletes, recreational or professional, can be complex as short- and long-term outcomes are influenced by the higher demand that athletes put on their bodies. Additionally, the orthopedic sports medicine specialist does not only treat the patient’s injuries but also has to consider the patient’s attempted return to his or her previous level of activity. “Getting back in the game” is sometimes subject to pressure and competing interests from third parties (e.g., team members, coaches, parents, fans). This can make the athlete’s treatment and the rehabilitation a challenging process.

Surgical intervention for ligament and cartilage injuries in sports medicine patients is usually done using arthroscopic techniques. The most frequently injured joints are the shoulder, knee, and hip. Therefore, treatment of common injuries in these joints will be the scope of this section.

SHOULDER

Rotator Cuff
Rotator cuff injuries are among the most common reasons to visit an orthopedic sports specialist. Often, these injuries are associated with forceful or repeated overhead and pulling movements. The rotator cuff provides shoulder movement and glenohumeral joint stability, and injuries can typically lead to pain, weakness, and restricted movement of the arm. Over recent years, improvement of surgical indications, operative techniques, and rehabilitation protocols has led to better outcomes. Studies suggest that arthroscopic techniques are equal or superior to open techniques for most indications. Controversies surrounding rotator cuff repair remain and include use of acromioplasty, enhancement of healing with orthobiologics (Fig. 43-23), single- vs. double-row fixation, and the treatment

Figure 43-23. Imaging and treatment of rotator cuff tears. A. Magnetic resonance imaging coronal T2 image showing a full-thickness and moderately retracted tear (arrow) of the supraspinatus tendon. B. Arthroscopic image showing the supraspinatus tendon tear as viewed from a posterior portal during the surgery. C. Arthroscopic image showing completion of repair of the supraspinatus tendon tear using suture anchors imbedded in the greater tuberosity of the humerus and attached sutures that capture and reduce the torn tendon to its native insertion site.
of massive or large tears. Rehabilitation after surgery plays an important role to restore strength, motion, and function and, ultimately, to return the patient to his or her previous level of activity. The standard rehabilitation protocol is made up of three consecutive stages: immobilization, passive exercise, and active exercise. Immobilization can be established by using a sling, and passive exercises should be initiated by the therapist in the first 4 to 6 weeks after surgery. The therapist moves the arm in different positions to improve range of motion (ROM) while providing support. After 4 to 6 weeks, active exercises can be gradually introduced. At 8 to 12 weeks, muscle strength and improvement of arm control are increased by starting a strengthening exercise program.

Shoulder Instability

The most common etiology for shoulder instability is related to trauma, especially shoulder dislocation. After a shoulder has dislocated, it becomes vulnerable to repeat episodes of instability and may develop into a chronic problem. Most of the shoulder’s stability is provided dynamically by the rotator cuff and statically by the shoulder capsule and ligaments. The most common dislocation is in the anterior-inferior direction. Typically, patients with an anterior dislocation present with pain and an externally rotated shoulder. Younger patients are more susceptible to suffer from repeat dislocations than older patients. The position of the humeral head with respect to the glenoid and other bony pathology can be identified with radiographs. Views from different angles should be obtained to thoroughly evaluate the patient; an anterior-posterior (AP) view, a glenoid (axillary) view, and a “Y” view of the shoulder are recommended in assessing this injury. Immediate reduction of the glenohumeral joint is paramount to the initial treatment of this injury. Repeat radiographs should be obtained to ensure that the humeral head is appropriately positioned. As soft tissue structures are typically damaged in these injuries, an MRI can be obtained to evaluate these structures.

Relocation of the shoulder is generally accomplished with the patient in supine position and the arm under gentle traction and slight abduction. Some sedation is helpful as it relaxes the patient’s musculature. Whether or not to immobilize a first-time-dislocated shoulder remains controversial, as does the position of immobilization. Additionally, some surgeons argue that early surgical repair of capsulolabral structures is appropriate as the recurrence rate in the young population is high and may lead to more extensive bony involvement and ultimately more invasive open procedures.

Prolonged immobilization is not recommended because this will often lead to substantial stiffness in the shoulder and does not appreciably decrease the redislocation rate. Unfortunately, many patients experience recurrent dislocations, in which case surgical stabilization of the shoulder should be considered. Arthroscopic stabilization procedures have been the gold standard treatment for the majority of injuries related to shoulder dislocations, typically a tear of the anteroinferior capsulolabral complex (Bankart lesion). There are a subset of injuries, typically involving large glenoid bony lesions, that require more extensive intervention with a Latarjet procedure or bone grafting. After surgery, the shoulder is temporarily immobilized with a sling. When the sling is removed, exercises will be started to rehabilitate the shoulder, improve ROM, and prevent scarring. Strengthening exercises will gradually be added to the rehabilitation plan.

Posterior Dislocation of the Shoulder

Posterior dislocations of the shoulder are rare and could be missed. This dislocation can occur due to electric shock or seizures. Examination of the patient will show limitation in external rotation of the shoulder. The shoulder will be locked in an internally rotated position. The posterior dislocation could be missed on the AP view of the shoulder, and an axillary view should be obtained to avoid missing the injury.

Superior Labrum and Biceps Tendon

The labrum is a structure that helps to deepen the shoulder socket and stabilize the glenohumeral joint. Additionally, it serves as an attachment point for many of the shoulder ligaments, as well as the long head of the biceps tendon. A superior labrum anterior and posterior (SLAP) lesion may occur in the superior part of the labrum, usually anterior and posterior to the attachment of the biceps tendon, with occasional involvement of the biceps tendon in certain cases. Injuries to the superior labrum can be caused by trauma or by repetitive shoulder motion, such as in throwing athletes. Radiographs are generally obtained to evaluate for concomitant bony injuries or osteoarthritic changes. The labrum itself, and other soft tissue, is better visualized with MRI with addition of a gadolinium arthrogram adding sensitivity for labral injury detection. A coronal view MRI will clearly show the condition of the superior labrum.

Conservative and operative treatments have had mixed results depending on the patient’s age, activity level, type of tear, and presence of concomitant injuries. If symptoms do not improve with adequate physical therapy and/or nonsteroidal anti-inflammatory drugs (NSAIDs), surgical intervention is usually indicated. Some SLAP injuries involve the biceps tendon, which may require either tenotomy or tenodesis.

After surgical repair, the shoulder needs to be immobilized to protect the repair and allow for healing. Usually a sling is used for 4 weeks after surgery. Then a physical therapy program will gradually start improving range of motion and prevent scar formation and stiffness from developing. As healing progresses, exercises to strengthen the shoulder muscles and the rotator cuff will gradually be added to the program typically around 4 to 6 weeks after surgery. Return to early interval throwing can generally be allowed around 3 to 4 months after surgery.

Impingement Syndromes

After minor trauma or repetitive injury, patients may experience pain and discomfort which can be due to irritation of the tissues in the subacromial space. In many cases these shoulder impingement syndromes are caused by simple bursitis or tendinitis of the long head of the biceps or supraspinatus tendon. Occasionally, impingement syndromes can progress to tears of the supraspinatus tendon, which can be confirmed by MRI or ultrasound.

The goal of treatment is to reduce pain and restore function. Initial treatment is generally nonsurgical and involves rest, NSAIDs, and physical therapy. If pain is not adequately relieved, an injection of a local anesthetic and corticosteroid may be helpful, for both therapeutic and diagnostic purposes.

If conservative treatment does not relieve pain, surgery is recommended, with the goal to excise the bursa and create more subacromial space. Generally, surgery is performed arthroscopically and encompasses bursectomy and subacromial decompression via acromioplasty. If the rotator cuff (supraspinatus tendon) is also injured, arthroscopic repair is usually indicated to restore
function and can be accompanied by a bony resection of the inferior portion of the acromion.

The Acromioclavicular Joint
The acromioclavicular joint is a gliding synovial joint comprised of the lateral end of the clavicle and medial facet of the acromion, and it has limited mobility. The joint is stabilized by three ligaments: the superior acromioclavicular ligament, the inferior acromioclavicular ligament, and the coracoclavicular ligament. Injuries to these ligaments are commonly sustained by a lateral blow to the shoulder while playing contact sports such as football and ice hockey and may cause displacement of the joint. An acromioclavicular sprain is also referred to as a shoulder separation. The least severe, types I and II, are typically treated conservatively. Treatment of type III injuries, where the clavicle is displaced up to 100%, is controversial. Some advocate for early surgical intervention, while others recommend symptomatic treatment followed by ligament reconstruction if symptoms persist. Types IV to VI, where the coracoclavicular ligaments are completely torn and the clavicle is significantly displaced, are often treated surgically.

KNEE
The knee is the largest joint in the human body and is a pivotal hinge joint, which allows flexion and extension as well as some medial and lateral rotation. The knee bears tremendous axial loads as well as torsional and sheer forces, making it vulnerable to both acute injury and the development of osteoarthritis. In sports, the major stabilizing structures such as the ACL and the medial collateral ligament (MCL) are frequently injured. Other common knee injuries involve the menisci, posterolateral corner (PLC), posterior cruciate ligament (PCL), and patellofemoral joint.

Menisci
The menisci are crescent-shaped pieces of fibrocartilage that provide joint stability, shock absorption, load distribution, and proprioception to the knee. Sudden meniscal tears often happen during sports, usually during contact or twisting injuries to the knee. Symptoms associated with a meniscus tear include pain, swelling, stiffness, catching, and locking of the knee. Radiographs are typically obtained to assess for a concomitant bony injury, the presence of (early) osteoarthritis, and leg alignment abnormalities. However, since menisci are radiolucent and are not seen on radiographs, an MRI is obtained to assess the status of the menisci and the soft tissue surrounding the knee joint (Fig. 43-25). Small tears on the outer edge of the meniscus may not cause symptoms, and provided the knee is stable, nonsurgical treatment may be sufficient.

The most commonly performed surgical procedure for meniscus tears is partial (subtotal) meniscectomy. However, it has become increasingly clear over recent years that preservation of the load-distributing function of the meniscus is important in preventing the development of early osteoarthritis. Research into the use of orthobiologics (e.g., microfracture of the notch, fibrin clot) for meniscal repairs has expanded the indications for repair rather than excising the torn fragment, especially if the fragment is large. Tears have been reported in virtually all portions of the meniscus, with radial and longitudinal tears being the most common. Meniscal root tears are less common, but they are increasingly being recognized as devastating injuries that cause serious alterations of knee contact forces. Surgical techniques are developing to repair the root to restore its function. Meniscus transplantation may be an option for young patients with a largely deficient meniscus.

Figure 43-24. Imaging and treatment of a shoulder glenoid labrum tear. A. Magnetic resonance imaging axial T1 image showing a tear of the posterior superior labrum (arrow). B. Arthroscopic image with the patient in the lateral decubitus position showing detachment of the torn labrum away from the glenoid. C. Arthroscopic image demonstrating repair of the labrum to its attachment site using anchors in the glenoid and sutures that fixes the labrum to the glenoid.
The paradigm of treatment of torn menisci is shifting thanks to the development of superior surgical techniques, use of orthobiologics, and promising first results with root repair and meniscus transplantations. Nowadays, physicians are well informed on the significance of meniscal preservation when there is potential for healing.

Directly after surgery, the knee is immobilized with a brace and weight-bearing is protected to allow the meniscus to heal. When healing is complete, range of motion and strength will need to be regained. Physical therapy is an integral component of healing and return to play, which usually is allowed between 4 and 6 months after surgery.

### Collateral Ligaments

The MCL is the most frequently injured knee ligament, which usually occurs after excessive valgus stress of the knee. In more severe injuries, tearing of the meniscus and ACL can also occur, which is known as the “unhappy triad.” This is most often seen in contact sports with a lateral blow to a planted leg, causing a significant valgus force.

The MCL has good healing potential, and grade I and II injuries usually improve with bracing and activity modification. Grade III injuries may also improve with conservative treatment, and often these injuries are initially treated non-operatively. The majority of MCL injuries occur in the mid-substance or at the femoral insertion side. There is a small subset of tibial-sided grade III tears that are associated with worse clinical outcome following conservative treatment, and therefore surgical repair is often recommended. Reconstruction is rare because surgical repair is usually effective in restoring the MCL. LCL injuries are much less common than MCL ligament injuries, but, similarly, most are managed conservatively.

With return of range of motion and normal gait pattern, patients are functionally progressed towards return to sports. A functional brace during sports is often advised.

### Cruciate Ligaments

The cruciate ligaments are situated centrally within the intercondylar notch of the knee. The biomechanical function of both the ACL and the PCL is complex and three-dimensional, but both play an important role in providing anteroposterior and rotational stability of the knee.

ACL tears are a common sports injury, especially in sports involving sudden stopping and cutting (e.g., soccer, basketball) or contact (e.g., football). A torn ACL will result in altered knee biomechanics and kinematics and thus can potentially lead to the early development of degenerative changes. Since a torn ACL will not heal on its own, surgical ACL reconstruction is generally the treatment of choice in patients who are young and active. Patients with a more sedentary lifestyle and who experience no persisting or disabling instability in daily life may be effectively treated with conservative management (i.e., bracing and physical therapy).
A patient with an ACL tear typically presents with pain and swelling, instability, loss of ROM, joint line tenderness (if there is an associated meniscus injury), and discomfort while walking. The Lachman’s exam is the best clinical test for an ACL tear. Radiographs are obtained to evaluate joint condition and possible associated osseous injuries. To visualize the ACL and other soft tissue in the knee, an MRI should be obtained. Although an MRI is not required to make the diagnosis, the information it provides is invaluable with regard to objectifying anatomic characteristics by taking measurements, assessing concomitant injuries, and presurgical planning in general (Fig. 43-26).

Reconstruction is performed with use of a tendon-graft that will replace the native ACL. Commonly used graft sources include the patellar, hamstrings, and quadriceps tendons. These tendons can be harvested from the same knee (i.e., autografts) during the same procedure. Alternatively, a cadaver graft (i.e., allograft) can be used. Both have their associated benefits, including the absence of donor site morbidity with an allograft and better healing potential with an autograft. As such, it is important to have a discussion with the patient and provide the necessary information for them to make an informed decision regarding graft type.

Injuries of the PCL are less common than other knee ligaments. Frequently seen causes are a bent knee hitting a dashboard in a car accident or falling on a knee that is bent during running. A rupture of the PCL is usually better tolerated than ACL rupture, since many tears (i.e., grades I and II) have the potential to heal on their own and do not result in much knee instability. Most grade I and II injuries are treated non-operatively. Combined PCL/PLC, PCL/MCL, and grade III PCL injuries do present a challenge with regard to appropriate management. Chronic PCL-deficient (grade III) knees have an increased incidence of osteoarthrosis, particularly in the patellofemoral and medial knee compartments. Indication for surgery is influenced by age, activity level, and the presence of concomitant injuries. Different surgical techniques have been proposed; the most common are the “inlay” technique and the transstitial technique.

The goal of cruciate ligament (both ACL and PCL) reconstruction is to restore native knee kinematics, provide the patient with the best potential for a successful outcome, and to prevent the development of long-term complications, such as osteoarthrosis.

Posterolateral Corner
Critical structures of the posterolateral corner are the LCL, popliteus tendon, and popliteofibular ligament. These structures each contribute to the static and dynamic stability of the knee and are commonly seen in combination with other ligamentous injuries, most notably the ACL. It is important to evaluate the PLC after any knee injury as a deficient PLC causes altered knee biomechanics and subsequently increases the stress on surrounding stabilizing structures. As such, it has been shown that a deficient PLC is a primary cause of graft failure in cruciate ligament reconstruction.

Acute high-grade injuries of the PLC with obvious deficient structures require surgical intervention. Since primary repair becomes increasingly difficult as time between injury and surgery increases, a cut-off of 2 to 3 weeks is usually the limit to repair the deficient structures. With more chronic PLC injuries or mid substance tears, reconstruction is recommended to restore knee stability.

HIP
Femoroacetabular Impingement
Femoroacetabular impingement (FAI) is a pathologic condition that refers to impingement of the anterior femoral head-neck junction against the anterosuperior acetabular labrum. This is frequently caused by abnormal bony offset at the femoral head-neck junction and is called CAM impingement, which usually affects young males. On the other hand, a Pincer lesion usually occurs due to abnormal acetabular version and excessive anterolateral acetabular bony rim coverage, or a combination of these, which usually occurs in females. Recognition of FAI can be clinically and radiologically difficult. However, familiarity with this disorder is essential.
because FAI can lead to labral tears, cartilage delamination, and, if untreated, osteoarthritis.

Commonly, patients present with anterior groin pain exacerbated by activities involving hip flexion or pain over the greater trochanter, as well as grinding or popping. Patients report pain with flexion and internal rotation, and after prolonged sitting. On examination, there is a decrease in internal rotation that appears out of proportion to the loss of the other ranges of motion, and flexion can also be limited. The impingement test, elicited by 90° of flexion and adduction and internal rotation of the hip, is almost always positive, signified by pain in the groin region.

The imaging findings of FAI can be seen on plain radiographs, CT scan, MRI, and magnetic resonance angiography. Some of the abnormalities seen include abnormal lateral femoral head/neck offset seen as a lateral femoral neck bump, os acetabuli, synovial herniation pits, acetabular over-coverage, hyaline cartilage abnormalities, and labral tears.

Treatment of FAI has traditionally been surgical and has evolved from open surgical treatment with acetabuloplasty, to combined open-arthroscopic-assisted techniques, to all arthroscopic techniques. Hip arthroscopy is becoming increasingly popular and is being more frequently applied for this indication. This popularity is largely the result of studies reporting an improvement of functional outcome measures with follow-up of 10 years in some studies and with relatively low complication rates.

**SPINE**

**Spinal Trauma**

In spinal injury, spinal stability must be assessed and the patient immobilized until the spine is cleared. CT scan is more reliable in assessing spine injury than plain radiographs. In patients with ankylosing spondylitis, an MRI is the best study to rule out occult fracture and epidural bleeding. When neurologic deficits are present a decompressive procedure may be indicated. In spinal cord compression, prompt decompression should be performed. Spinal cord injuries should be triaged to trauma centers since trauma center care is associated with reduced paralysis.

**Occipital Cervical Dislocation**

Motor vehicle accidents can cause dislocation of the occiput on the condyles of the atlas (C1). Most patients with this injury suffer cervical cord injury and do not survive. Traction on the spine is contraindicated. Treatment consists of stabilization and fusion in situ using a screw plate from the mid cervical spine to the occiput.

**Fractures of C1 (Jefferson Fracture)**

Fracture of the C1 ring was described by Jefferson in 1920. The thin anterior and posterior rings of the C1 vertebra fracture with axial loads. C1 fracture causes the lateral masses of C1 to spread, which can be visible on an open-mouth view. A lateral view of the C-spine may show the fracture; however, this injury could be missed due to inadequate visualization of the occipitocervical junction. CT scan is the ideal study for a Jefferson Fracture. The transverse ligament may be ruptured with a Jefferson fracture, and this will render the fracture unstable, which can cause injury to the spinal cord (Fig. 43-27). Jefferson fractures may be associated with other spine fractures. This injury is rarely associated with neurologic injury. The treatment of a Jefferson fracture is based on the integrity of the transverse ligament. The integrity of the transverse ligament is assessed by the amount of C1 lateral mass displacement determined by open-mouth radiograph and CT scan. Significantly displaced fractures (less than 7 mm) indicate disruption of the transverse ligament. An increase in the atlanto-dense interval (ADI) may indicate a transverse ligament injury. Normally, the ADI is less than 3 mm, as seen on the lateral view. An unstable injury with a rupture of the transverse ligament may need a posterior C1-C2 fusion.

Bracing with a cervicothoracic orthosis or a halo ring and vest is the recommended treatment for nondisplaced and minimally displaced fractures; significantly displaced unstable fractures require more definitive surgical treatment.

**Fractures of C2 (Odontoid Fracture)**

Half of normal cervical rotation occurs at the atlantoaxial joint. The odontoid (Dens) is a small bony process which arises from the body of C2, and articulates with the body of C1 (the Atlas). There are three types of odontoid fractures (Fig. 43-28). Type I fractures are the most common and are avulsion fractures off the tip of the dens. Type I fractures occur when there is tension applied to the alar ligaments (which span from the tip of the odontoid to the skull bypassing the C1 vertebra). Type I fractures are stable and managed nonoperatively.

Type II fractures, at the base of the odontoid, results from lateral loading forces. Operative stabilization in patients with a high risk of fracture nonunion is the preferred treatment since immobilization in a halo vest results in nonunion rates ranging
from 20% to 80%. The risk of nonunion includes displacement greater than 5 mm, angulation greater than 10 degrees, age over 50 years, smoking, and delayed diagnosis more than 4 weeks. Nonunion occurs due to interruption of the blood supply. Transfixing the odontoid fracture with a screw maintains rotational movement. Posterior fusion of C1 on C2 is another option, but this results in decreased cervical spine rotation; 50% of rotation of the cervical spine comes from C1 and C2 joint.

Type III fractures extend into the body of C2, below the origin of the odontoid process. The cancellous bone is rich in blood supply and usually heals well. Type III fractures are generally treated with a halo brace.

**Hangman’s Fractures of C2**

Hangman’s Fractures are a bilateral fracture of the pars interarticularis (Fig. 43-29). The spinal canal is usually widened, and neurological deficits rarely occur (Fig. 43-30). It results from sudden extension forces on the neck. Treatment is simple immobilization in a halo vest. Higher energy injuries causing severe extension forces can dislocate the C2-3 facet complex and damage the C2-3 disc. Significantly displaced Hangman’s fractures are managed by internal fixation and bone grafting between C2 and C3. When the fracture is severely angulated, it may indicate a flexion distraction injury, and traction on the C-spine may exacerbate the injury.

**Compression Fracture of the Cervical Spine**

In C3 to C7 an axial load can cause fracture of the endplate while preserving the posterior cortex of the vertebral body. These fractures generally heal well and are treated nonoperatively with analgesics and a cervical brace.

**Burst Fractures of the Cervical Spine**

Burst Fractures of the cervical spine usually result from axial loads such as in diving and motor vehicle accidents. The injury results in displacement of bony fragments into the canal, injuring the spinal cord. Burst fractures are treated surgically by anterior decompression (corpectomy) and reconstruction using a bone graft strut stabilized with a plate and screws.

**Unilateral and Bilateral Facet Dislocation**

This injury is usually associated with motor vehicle accidents. A restrained passenger can suffer forced flexion with distraction resulting in dislocation of the facets. The diagnosis can be made on lateral radiographs. Unilateral facet dislocation can be missed on an X-ray. It usually shows less than 25% subluxation on an X-ray, and it affects the nerve roots. Bilateral facet dislocation will have more than 50% subluxation on an X-ray and may cause severe spinal cord injury. Treatment consists of closed reduction with axial traction utilizing cranial tongs, graduated application of weight, and periodic X-rays. The patient is kept awake for safety concerns. A closed reduction should not be done if the patient is not awake. Facet dislocations could be associated with disc herniation. An MRI is the study of choice to rule out disc herniations and should be done prior to reduction in an unconscious patient or prior to open reduction and/or surgical fixation. When a reduction is obtained, the patient is taken to surgery for fusion, which may be performed anteriorly.
or posteriorly. Anterior surgery is necessary if the patient has an associated herniated disc.

**Clay-Shoveler’s Injury**
Clay-shoveler’s injury can result from a motor vehicle accident or from shoveling soil or clay. The injury (of C6, C7, T1, and T2) is the result of avulsion fracture of the spinous process by the paraspinal muscle forces (Fig. 43-31). The fracture is treated nonoperatively with analgesics and a soft collar.

**FRACTURES OF THE THORACIC AND LUMBAR SPINE**

**Thoracic Lumbar Spine Injury**
The ribs stabilize fractures of the thoracic spine, making these fractures more stable than similar fractures of the lumbar spine. Neurologic injuries are more common in the thoracic and proximal lumbar spine because of the presence of the spinal cord, which ends at the L2 level, as well as the small spinal canal diameter of the thoracic spine.

**Compression Fracture**
Compression fractures result from osteoporosis as well as trauma. Compression fractures involve a fracture of the anterior part of the vertebral body without associated posterior cortex fracture. Thoracolumbar compression fractures are treated nonoperatively with braces and analgesics.

**Burst Fracture**
Burst fractures are caused by falls and high-energy automobile accidents. The posterior cortex fracture (middle column involvement) differentiates the burst fracture from a compression fracture. The injury may be associated with neurological deficits due to retropulsion of bone into the canal. A vertical lamina fracture may contain an invaginated segment of the dura mater with accompanying nerve root injury and dural tear. Widening of the pedicle in an AP view of the spine will indicate a burst fracture. CT scan will define the bony injury, and an MRI will show compression of the neural elements and any injury to the posterior ligaments.

Treatment is nonoperative with an orthoses and mobilization of the patient if the fracture is stable. Surgery is done for decompression and destabilization of the spine if the patient has neurologic deficits or if the fracture is unstable.

**Seatbelt Injuries (Flexion Distraction Injuries)**
A seatbelt injury occurs when there is acute forward flexion of the trunk and anterior (i.e., seatbelt) restraint. The pelvis and upper torso move forward, and failure of the spine under tension begins with the posterior elements. Tearing of the dorsal fascia, the interspinous ligament, dislocation of the facets, and tearing of the discs occurs. Seatbelt injury may be bony or ligamentous (Fig. 43-32). The bone of the spinous process, the lamina, the pedicles, and the vertebral body fail in tension ("chance
fracture"). The bony injury could be stable. Flexion distraction injuries involving the soft tissue, with injury to the posterior spine elements, are usually unstable. This unstable injury will require internal fixation and fusion with bone grafting. This spine injury may be associated with a colon injury, especially in children.

**Fracture Dislocations of the Spine**

Fracture dislocations of the spine displace the bony elements by translation or rotation, resulting in canal narrowing and nerve injury.

Reduction of the displaced bones is the best way to improve the canal dimensions.

Patients with fracture dislocations of the spine and partial nerve function can recover. Fracture dislocations are treated operatively with surgical stabilization.

**Disc Herniation**

Disc herniation, most common between ages of 20 and 50, can occur in the cervical, thoracic, or the lumbar spine, and consists of a tear of the annulus allowing the nucleus pulposus material to extrude through the annulus and enter the canal, pressing on the exiting nerve or the "traversing" nerve roots. In the cervical spine, spinal cord compression can occur.

Symptoms of most disc herniations resolve within 8 weeks as the nerve root accommodates and inflammation recedes. The bulk of the extruded nucleus pulposus resorbs over time. When symptoms persist beyond 6 to 8 weeks, surgery with excision of the involved disc and decompression of the nerve roots may be indicated.

In cervical disc herniation, an anterior approach to the spine is performed with dissection through a transverse incision on the neck. Dissection is carried between the trachea, esophagus medially, and the carotid sheath laterally. The disc is then removed. The disc space is usually filled with bone graft to fuse the vertebrae. A locking screw low profile titanium plate is then attached to the vertebrae.

Posterior decompression and laminotomy exposes the posterior elements of the spine. A portion of the lamina is removed to allow access to the canal to correct foraminal impingement or to remove lateral disc herniations. While the posterior approach does not require fusion with plates and screws, central disc herniation cannot be managed through a posterior approach since the spinal cord cannot be safely retracted.

In thoracic spine disc herniation, the posterior approach is contraindicated because it may lead to paralysis.

For lumbar disc herniation, a midline incision is used, and laminotomy allows visualization of the lateral recess. Retraction of the dura allows visualization of the traversing nerve roots as well as of the disc fragment.

**Cauda Equina Syndrome**

Cauda equina syndrome is uncommon and occurs from a central disc herniation (Fig. 43-33). This can be a difficult diagnosis to make; however, it is a true emergency, and a delay in diagnosis can lead to permanent impairment. The patient will complain of back pain with bilateral leg pain. Bladder and bowel difficulty such as incontinence and frequency, saddle anesthesia, decreased perianal sensation, impotence, diminished rectal tone, and motor deficits. MRI will show a central disc herniation. Treatment is with urgent diagnosis and urgent surgical decompression. The results are better if the decompression is done within 48 hours of onset of symptoms. A central disc herniation causing cauda equina should be differentiated from a posterolateral disc herniation. The posterolateral disc herniation usually affects a nerve root and can be treated conservatively, at least initially. In the case of central disc herniation, it affects the cauda equina (the lumbosacral nerve roots), and this is a surgical emergency. Spontaneous recovery does not occur, and the outcome is catastrophic, including permanent loss of bowel and bladder control as well as the ability to have an erection if treatment is delayed.

**Spinal Stenosis**

A loss of hydration of the discs causes loss of disc height and bulging of annular tissue and the ligamentum flavum, which effectively narrows the canal (spinal stenosis). Osteophyte formation on the facet joints can also cause nerve impingement. Cervical stenosis can cause myelopathic symptoms (hyperreflexia, problems with fine hand dexterity, balance problems resulting in gait disturbance, weakness, and pain). In patients with low back pain and gait disturbance, obtain an MRI of the cervical spine to rule out cervical myelopathy. Pathology of the lumbo-sacral spine does not cause gait disturbances.

Lumbar stenosis causes neurogenic claudication (progressive pain, weakness, and numbness in the legs). The claudication symptoms result from standing and walking, which increases lumbar lordosis. Extension of the spine decreases the spinal canal diameter as well as the foramen size and worsens the condition. The symptoms resolve with sitting and bending forward (i.e., over a shopping cart) (Fig. 43-34). Flexion of the spine increases the spinal canal diameter as well as the foramen size and decrease the symptoms. The patient may have a normal neurologic exam, and it is important to study the vascular status of the patient and differentiate between neurologic and vascular claudication. Examine the pulses, prescribe a noninvasive vascular study if necessary. In general, walking causes the symptoms for both conditions, and standing relieves the vascular claudication symptoms. Spinal stenosis is treated with NSAIDs, epidural steroid injections, and physical therapy. Resistant cases may require surgical decompression.

Spinal stenosis usually occurs in patients over 50 years of age. With degenerative spondylolisthesis or scoliosis, fusion with instrumentation is usually required to prevent progression of the deformity.
Back Pain and Degenerative Disc Disease
Back pain occurs in the majority of adults but is usually self-limited, resolving in 1 to 2 weeks. Chronic unremitting back pain may suggest the possibility of infection, malignancy, or metastatic disease.

While radiographs are one option in the management of disabling low back pain, they are ineffective at ruling out malignancy, and radiographic findings correlate poorly with symptoms. Patients with severe degenerative symptoms may have no pain, while others with mild degenerative findings complain of severe pain. The potential for secondary gain and psychiatric problems and the unpredictable results of spine fusion add to the difficulty of diagnosis and choosing a treatment plan.

Intervertebral disc replacement prostheses are experimental in the treatment of degenerative disc disease. The potential for loosening, creation of wear debris, and bone loss complicating revision surgery are concerns, as are the proximity of the device to the spinal canal and the great vessels.

Scoliosis
Scoliosis is a lateral curvature of the spine. Lateral bending of the spine is always accompanied by rotational deformity (coupling).

In order to measure the severity of scoliosis, lines are drawn along the endplates of the vertebral bodies at either end of the curve, and the angle formed when these lines intersect determines the magnitude of the curve.

Scoliotic curves are classified as congenital, degenerative, metabolic (mucopolysaccharidoses), neurogenic (cerebral palsy), and myogenic curves (muscular dystrophy). Idiopathic scoliosis is the most common form and represents a spectrum of genetic disease.

Adults with scoliosis may present with axial pain and imbalance in posture. Treatment for scoliosis may include anti-inflammatory medications, therapy, and activity modification. In severe cases with objective deformity, surgical correction of the deformity may be indicated.

Idiopathic Scoliosis
The majority of idiopathic scoliosis curves become apparent during adolescence and progress during skeletal growth. Initial management consists of observation. Rapidly progressing curves are treated with braces. Brace treatment is recommended for curves between 20 and 40 degrees. For patients with large curves, surgical intervention may be needed using rods with grafting and fusion.

Neuromuscular Scoliosis
Neurologic conditions such as polio and cerebral palsy can lead to “uncompensated” scoliosis curves where the patient is unable to lean with his upper body to restore balance. Scoliosis correction surgery may be needed to facilitate sitting balance and to avoid skin breakdown caused by pelvic obliquity.

Introduction to Arthritis
Arthritis refers to a large number of medical conditions, including osteoarthritis, rheumatoid arthritis, septic arthritis, and post-traumatic arthritis. Each has the potential to lead to loss of articular cartilage lining the joints. According to the CDC and the National Health Interview Survey, approximately 55 million adults (22% of the U.S. population) have been diagnosed with some form of arthritis. This number is projected to grow to an astounding 67 million adults by 2030 (or 25% of the U.S. population).

Arthritis causes pain, loss of range of motion, decreased ability to perform work duties or participate in social functions, and decreased quality of life. The number of individuals suffering from arthritic conditions will continue to rise as the “baby boomer” generation enters old age and the prevalence of obesity rises in the U.S. population, as age and obesity are two major factors in the onset of arthritis.

Examination of the Patient
A thorough history and physical examination is indicated for all orthopedic patients. Patient history should include location, quality, severity, timing, and radiation of pain along with any referred pain, associated signs and symptoms, modifying factors, or prior treatments, including both conservative and surgical measures. Other details within the history and physical examination are equally important in establishing a diagnosis and successfully developing a treatment plan. If you listen carefully to your patients, they will often tell you their diagnosis.

For example, location of “hip pain” can narrow a differential diagnosis. Patients with activity-related groin pain often are found to have hip arthritis, whereas patients with

Figure 43-34. Person is seen bending over shopping cart to improve symptoms.
peritrochanteric pain (lateral hip pain) may be suffering from trochanteric bursitis. The importance of listening and focusing on the patient’s description of location and type of pain cannot be overemphasized.

Physical examination should begin by observing the patient’s gait, both with and without assistive devices if possible. This demonstrates the extent of the patient’s functional deficit and the effect of the patient’s pain. Typical gait patterns include antalgic gait due to pain, or a “Trendelenburg gait” (Fig. 43-35) where abductor weakness may lead to a poor outcome following total hip arthroplasty. Other aspects of the exam include assessment of leg length discrepancy, joint contractions, skin changes, assessment for prior surgical incisions to identify prior treatments or plan future surgical approaches, neurovascular status, and strength, as well as range of motion. These details document functional status and help to formulate a differential diagnosis. Patients with “hip pain” may have lumbar spinal stenosis, radiculopathy, or vascular disease that may play a large role in their presentation. Once an appropriate physical examination is performed, weight-bearing radiographs are needed. Advanced imaging, including CT and MRI, are rarely indicated in the initial workup. Once a diagnosis is made, specific treatment directed towards the patient’s condition can be initiated. The goals of treatment are to improve pain, preserve motion, and maximize patient function, independence, and quality of life.

**Nonoperative Management and Prevention of Arthritis**

Nonoperative measures to treat arthritis include weight loss, activity modification, rest, physical therapy, NSAIDs, bracing, and assistive devices such as a cane or walker. These treatments have the potential to decrease symptoms and improve function and quality of life. For example, holding a cane on the opposite side of the symptomatic extremity reduces the forces across the hip joint and subsequently decreases hip pain (Fig. 43-36). In nearly all cases, it is best to treat patients nonoperatively prior to recommending surgery.

Health and exercise can also play a role in the prevention of arthritis. Weight loss of as little as 11 lbs (5 kg) has been shown to decrease the risk of developing knee osteoarthritis in women by 50%. Similarly, patients who engage in regular physical activity have a lower incidence of arthritis. However, despite nonoperative treatment, surgical intervention may be required to effectively manage patient symptoms.

**Injections**

Joint injections are commonly performed into the knee and shoulder. Common injections into the knee include corticosteroids and hyaluronic-acid gels. Corticosteroid injections can decrease inflammation within the joint. These injections are usually administered in combination with a local anesthetic, such as lidocaine, in order to provide more immediate relief for both diagnostic and therapeutic purposes. If the patient has immediate relief of pain with injection of the joint, this localizes the source of the patient’s pain to the joint and may assist with diagnosis. Diagnostic hip injections are particularly helpful in distinguishing pain resulting from hip versus lumbar spine pathology. Any benefit received is therapeutic for the patient. Hyaluronic acid injections in the knee are frequently used and are commonly referred to as “viscosupplementation.” The viscosity of the synovial fluid is increased by hyaluronic acid, but its role and mechanism are not well defined yet. There is a risk of joint infection, cartilage injury from the needle, hemarthrosis, and failure to receive benefit. Short-term altered glucose metabolism in diabetic patients is common with corticosteroid injections. The efficacy of hyaluronic acid injections has been questioned by recent evidence summarized in the American Academy of Orthopaedic Surgeons’ Clinical Practice Guidelines.
Surgical Management of Arthritis

The most commonly performed procedure for arthritis of a major joint is arthroplasty, or joint replacement. Joint replacements, including hip and knee arthroplasty, are considered two of the most successful procedures performed in all of surgery. However, nonarthroplasty options exist and are typically performed for certain indications and goals.

Osteotomy. Osteotomy is cutting of the bone to change the position of the fragments, thereby improving rotation, alignment, or angulation. Osteotomy can be performed for both congenital and acquired deformities that contribute to the patient’s pain or development or progression of disease. Pelvic and femoral osteotomy can be utilized in the treatment of developmental dysplasia of the hip. The position of the acetabulum can be altered with pelvic osteotomies in order to provide more appropriate coverage of the femoral head, which is typically deficient anteriorly and laterally. Femoral osteotomies can be performed to correct version and varus/valgus deformity of the femoral neck. Osteotomies are performed to obtain more normal alignment and coverage of the femoral head within the acetabulum to prevent or delay future disease.

An osteotomy commonly used in the knee is a proximal tibia osteotomy. An adult patient who presents with isolated medial compartment knee arthritis and associated varus deformity would be a candidate for a valgus-producing (high tibial) osteotomy. An osteotomy that realigns the knee into slight valgus has the potential to off-load the medial compartment, slow disease progression, and prevent or delay the need for further procedures (unicompartmental or total knee arthroplasties).

Arthrodesis. Arthrodesis is a treatment option for severe arthritis where the overlying articular cartilage is removed and two opposing bones heal together with the use of hardware (internal or temporary external fixation) often supplemented by bone graft. After successful arthrodesis, no motion is possible through the joint and the source of pain is removed. Arthrodesis of large joints, such as the knee, shoulder, or hip, are typically explored as an option in the face of infection, in older adult, low-demand patients or in young, active patients who are considered too young for a joint replacement (out of concern for component wear and the need for early revision). Arthrodesis can also serve as a “last resort” procedure in orthopedics when joint preserving treatments fail due to fracture or infection. Ankle arthrodesis is the primary procedure performed in adult patients with traumatic arthritis of the ankle.

Joint Arthroplasty/Joint Replacement. Joint arthroplasty is the most common option for patients suffering from pain associated with arthritis in a joint. The surfaces of the bones are replaced after removing the damaged articular cartilage. The amount of bone and the determination of how to make the bone cuts is made based on preoperative radiographs and templating, cutting guides, anatomic measurements, and soft tissue/ligament balancing. The cut bony surfaces are covered with new components, usually made of metal, ceramic, or polyethylene. These new components are sized to appropriately match the patient, based on templated preoperative radiographs, intraoperative measurements, and examination for stability, leg length, alignment, and range of motion.

If all compartments or surfaces of the joint are replaced, the arthroplasty is referred to as a total joint arthroplasty. In comparison, if only one surface or compartment of the joint is replaced, it is referred to as hemiarthroplasty (hip, shoulder) or unicompartmental arthroplasty (knee). Total hip and knee arthroplasties are considered among the most successful of all surgical procedures performed in terms of patient outcome and improvement in pain.

Hip Arthroplasty

Background Hip arthroplasty is utilized for end stage arthritis in the hip that has failed a reasonable trial of nonoperative measures (Fig. 43-37). Conventional hip arthroplasty commonly refers to total hip arthroplasty where both the femoral head and acetabulum are replaced or resurfaced, respectively. Finally, hemiarthroplasty describes the replacement of the femoral head and neck with a stemmed femoral component in isolation. The acetabulum is not addressed surgically.

History of Hip Arthroplasty The history of hip arthroplasty (hip replacement) may be broken down into “Pre-Charnley” and “Post-Charnley” eras, referring to the significant contributions of Sir John Charnley to the evolution of hip arthroplasty. Prior to Charnley’s contributions, hip arthroplasty consisted of a variety of procedures with highly variable results. Early attempts at relieving hip pain were made with interpositional arthroplasty, where tissue layers, plastic, or metal were placed between the worn articular surfaces. Fracture of the interposed material or loosening of components often led to failure.

Later attempts introduced stemmed components to improve fixation. One of the earliest femoral components was designed by Austin-Moore. This prosthesis replaced the femoral head and neck with a metal component secured into the femoral shaft with a stem extending down the diaphysis. This prosthesis was utilized in hemiarthroplasty for many years and served as
a step in the development of total hip arthroplasty with the later addition of the acetabular component.

Surgical Approaches to the Hip A variety of approaches to the hip joint have been utilized in joint arthroplasty, including anterior approach (Smith Petersen), anterolateral approach (Watson-Jones), lateral approach (Hardinge), and posterior approach (Kocher-Langenbach). Each approach contains a unique set of advantages and disadvantages. The following is a brief summary of the most common approaches that are utilized in total hip arthroplasty.

Anterior approach (Smith Petersen): This approach is an inter nervously and intermuscular approach. It utilizes the inter
ervous plane between the femoral nerve and superior gluteal nerve. Superficially, the plane between the sartorius (femoral nerve) and tensor fasciae lata (superior gluteal nerve) is dissected in the deep layer and the plane between the rectus femoris (femoral nerve) and gluteus medius (superior gluteal nerve) is dissected. Advantages to this approach include supine positioning, use of intraoperative fluoroscopy for acetabular component positioning, and discontinuation of all ambulatory assistive devices 1 week earlier than other approaches. Downsides include difficult preparation and placement of the femoral component with higher rate of femoral fracture/femoral component revision, higher rate of wound complications, and lack of a true extensile approach.

Posterior approach (Kocher–Langenbach): The posterior approach is a muscle-splitting approach without an inter
ervous plane. After incising the skin and subcutaneous fat, the fascia lata is incised along with the gluteus maximus. The short external rotators are exposed and dissected, including the piriformis, superior and inferior gemelli, obturator internus and externus, and quadratus femoris. This allows internal rotation of the hip along with flexion and adduction to dislocate the hip. The posterior approach with posterior soft tissue repair has no increased rate of dislocation compared to the anterior approach. The posterior approach is extensile and provides excellent exposure of both the femur and acetabulum for complex and revision cases.

Lateral approach (Hardinge): While there have been many modifications to the original Hardinge approach, first described in 1982, most involve releasing the anterior one-third of the gluteus medius, underlying minimus, abductor tendon, and vastus lateralis distally in one sleeve off of the greater trochanter. The capsule is then incised to expose the hip joint. Care must be taken to protect the superior gluteal nerve during this exposure, which lies 5 cm proximal to the tip of the greater trochanter.

Exposure of the acetabulum is excellent with the modified Hardinge approach, which is extensile. However, access to the posterior column is limited compared to the posterior approach. The increased risk of postoperative Trendelenburg gait, other pathologic gait, and heterotopic ossification compared to all other approaches to the hip have made it far less commonly performed than the posterior approach. Minimally invasive total hip arthroplasty is associated with decreased visualization intraoperatively and associated risks of component malposition, intraoperative fracture, and nerve or vascular injury. In fact, the only documented benefit of minimally invasive techniques appears to be a smaller incision, but with increased soft tissue tension intraoperatively comes the risk of compromised wound healing and periprosthetic joint infection.

Bearing Surfaces in Hip Arthroplasty The most common combination of bearing surfaces used in total hip arthroplasty is a metal (generally cobalt chrome) or ceramic prosthetic head, articulating with a polyethylene liner. Metal on metal (MOM) articulations have largely been abandoned in total hip arthroplasty as they are associated with production of metal ions that deposit in solid organs, pseudotumors that are locally destructive to soft tissue/bone, and risk of early failure (Fig. 43-38). Ceramic on ceramic articulations have the lowest friction of all current bearing combinations. However, ceramic may fracture or squeak in ceramic on ceramic total hip arthroplasties.

Alignment of Hip Arthroplasty Components Proper alignment of hip arthroplasty components is vital to a successful procedure and patient outcome. Surgeons aim for appropriate alignment of components to restore a functional and stable range of motion. This is accomplished with combined version of the femoral and acetabular components, appropriate abduction of the acetabular components, and staying true to Sir John Charnley’s principles: establishing a low friction articulation,

![Figure 43-38. Failed ceramic on metal hip arthroplasty components. Note the metallic staining on the ceramic femoral head.](image-url)
medializing the acetabular component and center of rotation and restoring abductor length and tension with restoration of appropriate length and femoral offset. Inappropriate placement of components can lead to early failure, accelerated component wear, dislocation, need for revision surgery, as well as poor patient outcomes and satisfaction.

**Knee Arthroplasty**

**Background** Knee arthroplasty is indicated for end-stage arthritis that has failed a reasonable trial of nonoperative measures (Figs. 43-39 and 43-40). Knee arthroplasty commonly refers to total knee arthroplasty where the distal femur, tibia, and patella are resurfaced after any remaining articular cartilage and a layer of subchondral bone are resected. A unicompartmental knee arthroplasty consists of replacing one compartment of the knee, most commonly the medial compartment.

**Surgical Approach to the Knee** Total knee arthroplasty is generally accomplished through a medial parapatellar approach. This approach utilizes a longitudinal skin incision extending, on average, 5 cm proximal to the patella to the medial aspect of the tibial tubercle distally. Dissection is carried down to the capsule. To gain access to the joint, an arthrotomy is performed medial to the patella extending proximally along the most medial aspect of the quadriceps tendon and distally just medial to the patellar tendon. This approach provides excellent exposure to all three compartments of the knee after patellar dislocation.

Once the joint surfaces are adequately exposed, remaining articular cartilage and a thin layer of underlying bone are removed prior to placement of prosthetic components. Bone cuts are made based on preoperative templating, cutting guides, ligament balancing, and anatomic measurements (Figs. 43-41 and 43-42).

**Bearing Surfaces in Knee Arthroplasty** The femoral component consists of a metal prosthetic cap sized to fit the normal shape of the distal femur. The tibia is cut perpendicular to the anatomic and mechanical axis, and a flat, stemmed, metal tray is placed that serves as a base plate for a polyethylene bearing surface. The patella is usually resurfaced with a polyethylene component.

Two types of primary total knee arthroplasty systems exist, including cruciate retaining and posterior stabilized systems. As the name implies, with cruciate retaining systems, the PCL is retained in hopes of preserving more normal knee structures and minimizing bone loss, while in posterior stabilized systems the ligament is sacrificed and the components are designed to accommodate for the loss. These two systems have equivalent results in knee arthroplasty.

**Alignment and Balancing in Knee Arthroplasty** Appropriate sizing and positioning of the components and balancing of the size and geometry of bony gaps in flexion and extension are essential for a successful knee arthroplasty. Inappropriate component position can lead to early wear and failure, instability, pain, and stiffness.

**Computer Navigation, Robotics, and Joint Arthroplasty**

Computer-navigated joint arthroplasty has the theoretical benefit of more accurate and consistent placement of arthroplasty components through intraoperative feedback to the surgeon regarding component position, planned bone cuts, and alignment. Disadvantages include increased costs of the technology, prolonged operative times, and risk of infection/fracture at the sites of intraoperative sensor placement within bone. Use of
Figure 43-41. A. Varus knee with osteoarthritis. B. Right total knee replacement.

Figure 43-42. Computer-assisted robotic targeting arm for total knee replacement.
computer navigation in total joint arthroplasty has been shown to minimize outliers in alignment, but there has been no proven benefit in survival or function secondary to computer-navigated or robotic-assisted joint replacement.

**Fixation Options in Joint Arthroplasty**

Components in hip and knee arthroplasty can be secured with cement or biologic fixation. The cement most commonly used is polymethylmethacrylate (PMMA). PMMA serves as a grout between the component and the bone surface. Components secured without cement are grit blasted or porous coated to allow bony ongrowth or ingrowth, respectively. Hydroxyapatite can also be utilized on implant surfaces to promote bone ingrowth or ongrowth through osteoconductive properties. A majority of hip joint arthroplasty components are now secured without cement, where initial fixation of components is accomplished through press fit techniques. In knee arthroplasty, cement utilization is generally preferred. In hip replacement patients where biologic fixation is unreliable, such as older adults, osteoporotic or previously irradiated cement may be a better option. With revision total hip arthroplasty, cement fixation of components has been shown to lead to earlier mechanical failure.

**Osteolysis and Aseptic Loosening.** Osteolysis is a term used to describe abnormal resorption of bone. Osteolysis can be caused by underlying infection, metastatic disease, or in case of joint replacement, the production of wear debris. Even with appropriately positioned components, some wear of the bearing surfaces is expected over time. However, wear rates as well as the size and amount of wear debris differs with the bearing surface. Friction in ceramic on ceramic articulations is the lowest of all bearing surfaces; however, there is increased risk of component fracture and postoperative “squeaking.” In metal or ceramic on polyethylene articulations, wear debris is produced, and polyethylene particles are phagocytized by local macrophages. Activated macrophages lead to an osteolytic process and bone resorption. Particulate methylmethacrylate cement debris can also play a role in osteolysis by damaging the polyethylene bearing surface. Osteolysis has been shown to be significantly decreased with the advent and use of highly cross-linked polyethylene. Improperly positioned components or patient-related factors such as high impact activities can lead to increased wear. A substantial osteolytic response may occur and lead to component micromotion and aseptic loosening. Patients who present to clinic with pain following joint arthroplasty and an increasing zone of osteolysis in the periprosthetic region frequently need revision surgery (Fig. 43-43). Alternative bearing surfaces continue to be explored in hopes of decreasing component wear, associated osteolysis, and aseptic loosening.

**Complications in Joint Arthroplasty**

The risk of any complication following joint arthroplasty procedures falls in the range of 5% to 10%. Risks shared by hip and knee arthroplasties include infection, intraoperative or postoperative fracture, vascular injury, need for intraoperative or postoperative blood transfusion, nerve injury or nerve palsy (most commonly involving the deep peroneal nerve and loss of ankle dorsiflexion), stress shielding, component fracture or wear, and medical complications, including venous thromboembolic disease (DVT and PE), myocardial infarction, or cerebrovascular accident. Complications unique to total hip arthroplasty include dislocation, leg length discrepancy, and iliopsoas impingement or tendinitis.

**Dislocation Following Hip Arthroplasty.** Dislocation can result from malpositioned components (inadequate combined version of the femoral stem and acetabular component; extremes of inclination of the acetabular component), noncompliance, cognitive or neuromuscular disorders, compromised soft tissue envelope from revision surgery, fracture, or insufficient restoration of length and/or offset. Comparable dislocation rates have been found with anterolateral, lateral, anterior, and posterior with soft tissue repair (approximately 0.5%) approaches. History, physical examination, and radiographs are vital to proper treatment of dislocation. Closed reduction can usually be performed with conscious sedation and gentle traction or manipulation. Rarely, open reduction may be necessary. Component position should be assessed in patients with multiple dislocations. Patients with recurrent dislocations and suboptimally positioned components may require component revision. Patients with recurrent dislocations and properly positioned components should be considered for conversion to a device with a larger prosthetic head (dual mobility construct) or a constrained total hip arthroplasty implant that provides improved stability.
blue cell lesions are most likely neuroblastoma in a 5-year-old, Ewing’s sarcoma in a 10-year-old, lymphoma in a 20-year-old, and myeloma in a 60-year-old. Gender also aids in the differential. For instance, giant cell tumor is more common in females, while osteosarcoma is more common in males. Multiple bone involvement may suggest enchondromas (Ollier disease, Maffucci’s syndrome) or osteochondromas (multiple hereditary exostoses).

**Laboratory Tests.** Laboratory tests determine the level of cellular turnover (lactate dehydrogenase [LDH]) or of bone destruction (calcium, alkaline phosphatase). Elevated prostate-specific antigen (PSA) suggests prostate cancer.

**Imaging.** Radiographic studies are critical in the diagnosis of bony tumors. Radiographs can help assess the aggressiveness of the tumor. Four questions should be addressed when assessing radiographs: (a) Where is the tumor—in which bone (Table 43-1) and in which part of the bone is the lesion? (Table 43-2) (b) What is the tumor doing to the bone (clinical behavior)? (c) What is the bone doing to the tumor (biologic response)? and (d) What is the matrix pattern? Matrix is the acellular interstitial substance produced by tumor cells. Particular attention should be paid to the junction between the tumor and the host bone since this margin can also indicate the aggressiveness of the tumor. Ewing’s sarcoma has a characteristic “onion skin” periosteal reaction pattern. This reaction pattern also occurs in other tumors and infections.

**OSTEOSARCOMA**

The most common primary malignant bone tumor is osteosarcoma (Fig. 43-44). Osteosarcomas are classified as osteoblastic, chondroblastic, fibroblastic, telangiectatic, round cell, or MFH-like, according to the predominant cell type. Most osteosarcomas present in patients between 10 and 20 years of age. Secondary osteosarcomas occur in older patients in abnormal bone affected by Paget’s disease, radiation, or bone infarct.

**Intramedullary Osteosarcoma**

This is the most common primary sarcoma of the bone. It usually occurs in the distal femur or the proximal tibia in young people. This condition may also occur at the proximal humerus, proximal femur, or pelvis. It usually presents itself as a high-grade extracompartamental disease. It can metastasize to the bone, which is called a “skip lesion,” but the lung is the primary site of metastases. Long term survival is 75% with adequate treatment. The response to chemotherapy (98% necrosis of the

---

**Table 43-1**

<table>
<thead>
<tr>
<th>Common locations of bone tumors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FEMUR</strong></td>
</tr>
<tr>
<td>Distal posterior</td>
</tr>
<tr>
<td>Distal anterior</td>
</tr>
<tr>
<td><strong>TIBIA</strong></td>
</tr>
<tr>
<td>Adamantinoma, chondromyxoid, fibroma</td>
</tr>
<tr>
<td><strong>HANDS AND FEET</strong></td>
</tr>
<tr>
<td>Enchondroma, exostosis</td>
</tr>
<tr>
<td>Calcaneus</td>
</tr>
<tr>
<td><strong>SPINE</strong></td>
</tr>
<tr>
<td>Anterior</td>
</tr>
<tr>
<td>Posterior</td>
</tr>
<tr>
<td><strong>PELVIS</strong></td>
</tr>
<tr>
<td>Metastatic, myeloma, chondrosarcoma, giant cell tumor, aneurysmal bone cyst, Paget’s disease, Ewing’s Sarcoma</td>
</tr>
<tr>
<td><strong>SACRUM</strong></td>
</tr>
<tr>
<td>Chordoma (midline), chondrosarcoma, giant cell tumor, aneurysmal bone cyst, lymphoma</td>
</tr>
<tr>
<td><strong>RIBS</strong></td>
</tr>
<tr>
<td>Metastatic, myeloma, fibrous dysplasia, chondrosarcoma</td>
</tr>
</tbody>
</table>

---

**Table 43-2**

<table>
<thead>
<tr>
<th>Tumor location in bone</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Epiphysis</strong></td>
</tr>
<tr>
<td>Chondroblastoma, clear cell chondrosarcoma, giant cell tumor (GCT), infection, dysplasia epiphysealis hemimelica (DEH)</td>
</tr>
<tr>
<td><strong>Metaphysis</strong></td>
</tr>
<tr>
<td>Most common site of involvement</td>
</tr>
<tr>
<td><strong>Diaphysis</strong></td>
</tr>
</tbody>
</table>

---

**Figure 43-44.** Osteosarcoma.
tumor after chemotherapy is a good sign) and the stage of the disease determines the prognosis. Young patients may present with pain and swelling, with X-rays showing plastic lesions in some areas of destruction with periosteal reaction called “Codman’s Triangle.” X-rays may show bone formation with a sunburst appearance. MRI should involve the entire bone to diagnose the skip metastases. CT scans of the chest are usually done to find primary metastases. Alkaline phosphatase is usually high. Diagnosis is typically confirmed with a biopsy, which is done after staging the tumor. Proper biopsy technique should be employed, which includes longitudinal incisions. As a general rule, the biopsy should be done by the same surgeon who will provide the definitive treatment for the patient. Treatment of osteosarcoma will be preoperative chemotherapy and wide resection, followed by postoperative chemotherapy.

**Parosteal Osteosarcoma**

Parosteal osteosarcoma is a low-grade surface osteosarcoma that appears as if it were stuck on the bone, especially in the posterior distal femoral metaphysis (80%). The differential diagnosis includes osteochondroma and myositis ossificans. Treatment consists of wide excision. The prognosis is 95% 5-year survival as it is a low-grade tumor.

**Periosteal Osteosarcoma**

Periosteal osteosarcoma is a high-grade tumor. It occurs on the anterior surface of the distal femur or proximal tibia. The lesion appears chondroblastic on histology. Radiographs show scalloping of the underlying cortex with a “sunburst” periosteal reaction. Treatment is chemotherapy and wide surgical excision. The 5-year survival rate is 80%.

**Paget’s Sarcoma**

Paget’s sarcoma is a rare complication of Paget’s disease. In Paget’s disease with multiple bone involvement, osteogenic sarcoma, fibrosarcoma, chondrosarcoma, and MFH have occurred, most often in the pelvis, but also in the femur, humerus, spine, and skull. This malignant transformation occurs in less than 1% of patients. The patient will complain of new onset pain and swelling. The physician must have a high index of suspicion in patients with Paget’s who previously had no pain. Imaging may demonstrate osteolytic areas and loss of normal fatty marrow and multifocal lesions. Treatment of Paget’s sarcoma is chemotherapy and wide surgical excision. The prognosis is poor, and the 5-year survival rate is less than 10%.

**Radiation-Induced Sarcoma**

The three criteria for diagnosis of radiation-induced sarcoma are (a) histology different from the original lesion, (b) sarcoma develops in the irradiated field, and (c) a 3- to 5-year latent period between radiation and sarcoma development. Radiation for carcinoma of the breast and cervix can result in osteosarcoma, chondrosarcoma, fibrosarcoma, or MFH. Treatment is a combination of chemotherapy and surgery.

**EWING’S SARCOMA**

Ewing’s sarcoma is the second most common primary bone tumor in patients under 25 years of age. The typical presentation is a tumor in the diaphysis of long bones, especially the femur. It can also be seen in the pelvis, the proximal tibia, and proximal humerus, and it is usually seen in young white males. It has a t(11:22) translocation and positive CD99. The patient may have pain and fever with an elevated sedimentation rate and WBC count; the condition may be confused with an infection. An “onion skin” periosteal reaction may be seen on radiographs. A large soft-tissue extension from the primary bone tumor may be seen, and histology reveals a small, round, blue cell tumor (Fig. 43-45). Diagnosis is confirmed with bone marrow biopsy specimen. Bone scan can identify multiple lesions. Treatment is chemotherapy and surgery or radiation therapy for spine or pelvic lesions.

**CARTILAGE-FORMING TUMORS**

**Chondrosarcomas**

Chondrosarcomas typically occur in male patients over 40 years of age, and they are the third most common primary bone malignancies. Primary chondrosarcomas can form clear cell, mesenchymal, or dedifferentiated neoplastic cartilage. Secondary chondrosarcomas may also develop in preexisting lesions such as exostoses or enchondromas. Pelvis, shoulder, and ribs are common locations. Chondroid or “popcorn” calcifications are typical on radiographs. Clear cell chondrosarcoma and mesenchymal chondrosarcoma occur in younger patients (second to fifth decades of life). Clear cell chondrosarcomas are low-grade lesions that often affect the epiphyses. The dedifferentiated chondrosarcoma is a high-grade chondrosarcoma with a less than 10% survival rate. It has a biomorphic histology, with a chondroid component and a high-grade spindle cell component.

The treatment of chondrosarcoma is surgical excision, since cells are not chemosensitive or radiosensitive. For high-grade lesions, wide or radical resection is recommended. Pelvic and scapular chondrosarcomas have a high recurrence rate, and adjuvant chemotherapy does not improve survival rates.

**FIBROUS LESIONS OF BONE**

**Desmoplastic Fibroma**

Desmoplastic fibroma is a rare tumor occurring in the mandible, femur, pelvis, radius, or tibia in young adults. It presents as a painful lesion. Radiographs show a metadiaphyseal “soap bubble” appearance and endosteal scalloping. Histology resembles
desmoid tumors or fibromatosis. Recommended treatment is wide excision to avoid recurrence.

**Malignant Fibrous Histiocytoma of Bone**

MFH occurs in the metadiaphysis of long bones after conditions like nonossifying fibromas and bone infarcts. It may present with pain or by a pathologic fracture. Radiographs typically show destructive lesions with soft-tissue extension. Histology resembles osteosarcoma with pleomorphic spindle cells, histiocytes, and giant cells, but no neoplastic osteoid formation. Treatment is chemotherapy and wide surgical excision.

**Malignant Vascular Tumors**

**Hemangioendothelioma.** Hemangioendothelioma is a malignant neoplasm arising from vascular endothelium in long bones and most often occurs in the lower extremity. Radiographs show a metadiaphyseal lytic lesion with a “soap bubble” appearance. Histology reveals eosinophilic cells in a basophilic stroma. Lesions may be multifocal. Treatment consists of curettage for low-grade lesions and wide excision +/- radiation therapy for high-grade lesions.

**Hemangiopericytoma.** Hemangiopericytoma is usually a solitary lesion occurring in the soft tissues or the axial skeleton and proximal long bones in middle-aged or older adult males. Histology reveals branching “staghorn” vascular spaces. The tumor cells resemble cells normally seen adjacent to capillaries. Treatment is wide excision.

**Angiosarcoma of Bone.** Angiosarcoma is a soft tissue malignancy usually seen in older adult males; chronic vascular stasis is a risk factor. Histology reveals vascular channels with anaplasia. Treatment is wide excision, or if the tumor is surgically inaccessible, radiation.

**MISCELLANEOUS TUMORS**

**Giant Cell Tumor of Bone**

Giant cell tumor is a benign aggressive tumor. Fifty percent of these tumors occur around the knee, especially at the distal femur and the proximal tibial. Giant cell tumors may also occur in the distal radius, proximal humerus, and pelvis (especially the sacrum ala) in women 20 to 40 years of age. Presenting complaints include pain and pathologic fracture. Imaging reveals eccentric, epimetaphyseal lytic lesions eroding the subchondral bone. Histology reveals multinucleate giant cells and mononuclear stromal cells. An abundance of giant cells in the field can help establish the diagnosis, and the nuclei of giant cells appear the same as the stroma cells (all nuclei look similar) (Fig. 43-46). Giant cell tumors must be differentiated from the Brown tumor of hyperparathyroidism. While both have giant cells, hyperparathyroidism affects multiple areas, and the serum calcium is not normal. Epiphyseal lesions such as chondroblastoma should also be part of the differential diagnosis. These tumors can occasionally metastasize to the chest. Primary malignant giant cell tumor has a poor prognosis. Treatment of giant cell tumors is with curettage and high-speed burr. Recurrence rates are high with simple curettage, and the use of adjuvants such as cryosurgery, phenol, or polymethylmethacrylate bone cement may help decrease recurrence rates. After pathologic fractures, wide excision with reconstruction or amputation may be required.

**Adamantinoma and Osteofibrous Dysplasia**

Adamantinomas are low-grade malignant tumors usually seen in the tibia (Fig. 43-47). Adamantinomas are capable of metastasizing to the lung. The patient may present with pain and/or bowing of the tibia. X-ray reveals multiple lucent lesions on the cortex of the tibia. Histology reveals a biphasic tumor with nests of epithelial cells and fibrous stroma (see Fig. 43-46). Osteofibrous dysplasia is considered the precursor to adamantinoma and should be part of the differential diagnosis. Osteofibrous dysplasia is a benign lesion, usually occurring in children, at the anterior tibia, which is treated with observation. The treatment of adamantinoma is with wide surgical excision.
Primary Lymphoma of Bone

Primary lymphoma accounts for about 5% of all neoplasms of bone. Long bone involvement is more frequent than spine. Lymphoma of bone typically occurs in males in their forties. Histology reveals large B cell lymphomas. Treatment is a combination of chemotherapy and radiation. Surgery may be required for stabilization of pathologic fractures.

Chordoma

Chordoma arises from notochordal remnants in the sacrum. It is usually midline in location. These tumors are found in middle-aged to older men and presents with bladder and bowel symptoms due to involvement of the cauda equina. Visualization of the lesion may be difficult because of the bowel gas shadow. Diagnosis may be delayed. An MRI shows a destructive extensive midline lesion with a large soft tissue mass. Histology shows epithelioid cells arranged in cords with vacuolated foamy physaliferous cells. These cells are keratin positive. Treatment is surgical excision and muscle flaps and a mesh for reconstruction. Urinary diversion and colostomy may be needed for loss of bladder and bowel control. Local recurrence is common.

Multiple Myeloma

Myeloma, the most common primary bone malignancy, is a proliferative disorder of B cells with plasma cells producing immunoglobins. These plasma cells have a classic eccentric nucleus giving a “signet ring” appearance (Fig. 43-48). Evidence of monoclonal protein in the serum and/or urine (Bence Jones proteinuria), and hypercalcemia, renal insufficiency, anemia, or bone disease are usually present. Presenting symptoms in myeloma range from bone pain and osteopenia to focal lytic lesions with pathologic fractures and hypercalcemia. Myeloma protein 1-α stimulates osteoclast formation. Osteoclast activating factors increase receptor activator of nuclear factor κB ligand (RANKL) in the bone marrow. RANKL induces osteoclast differentiation and activation. Myeloma cells inhibit osteoblast differentiation and activity. Serum and urine electrophoresis detect the M protein. Workup also includes complete blood cell count, erythrocyte sedimentation rate, calcium levels, renal function assessment, β2-microglobulin levels, and a skeletal survey. X-ray will show multiple punched out lytic lesions. Bone scans may be cold in about 30% of cases. The SPEP, UPEP, and bone marrow biopsy are helpful in diagnosis. Histology will show atypical plasma cells with eccentric nuclei, its appearance resembles a “signet ring (Fig. 43-49).” Plasmacytoma is a solitary tumor with a negative bone marrow biopsy, usually treated with radiation to the lesion. Myeloma is treated with bisphosphonates, chemotherapy, stem cell transplantation, and radiation therapy. Surgical stabilization and irradiation is done for pathologic fractures or impending fractures. Many patients with myeloma develop a vertebral compression fracture. Kyphoplasty can be useful in providing pain relief. The risks of cement extravasation and related complications are lower with kyphoplasty than with vertebroplasty. If there is instability or if there is neural compression, surgical stabilization may be required.

METASTATIC BONE TUMORS

Metastatic bone tumors are more common than primary bone tumors. Metastatic tumors affect the lung, liver, and bone. Cancers that commonly metastasize to bone are breast, lung, thyroid, kidney, and prostate. In patients older than 40 years of age, metastases and myeloma are the most common causes of destructive lesions in bone. The most common site of involvement is the axial skeleton, especially the thoracic spine, and proximal ends of long bones, especially the proximal femur. Lung and renal cell carcinomas can metastasize distal to the knee and elbow. Malignant cells are able to detach from one location and set up a focus at a distant site. The tumor activates osteoclasts and causes destruction of the bone, a mechanism that involves the RANK/RANKL pathway. The patient may present with pain, pathologic fractures, or the manifestation of hypercalcemia. Workup of a patient with a suspected metastatic disease to bone and an unknown primary tumor should include CT of the chest, abdomen, and pelvis. The extent of the disease is evaluated by bone scans (myeloma and thyroid are usually cold in bone scans), mammography, tumor markers, serum, and urine electrophoresis (SPEP and UPEP). A biopsy may be necessary to rule out primary bone lesions if the primary site is not identified. Treatment of bone tumors depends on the diagnosis, as metastatic tumors are treated differently than primary bone tumors. Metastatic tumors are usually treated by bisphosphonates and by surgical stabilization with postoperative radiation if warranted. Primary bone tumors are usually treated by wide excision with chemotherapy in high-grade tumors (chondrosarcoma are treated only with wide excision). Radiation therapy can be used in Ewing’s.
Birth Injuries

Neonatal Brachial Plexus Palsy. Injury of the brachial plexus during delivery occurs in 2 births in every 1000. Large birth weight, forceps delivery, breech presentation, and prolonged second stage of labor with shoulder dystocia are risk factors. Brachial plexus injury usually represents a stretch injury on the nerve roots of the upper or lower plexus.

Upper plexus injuries (Erb-Duchenne) are lesions manifested by weakness of shoulder abductors and external rotators as well as the elbow flexors (Fig. 43-50). The hand is not involved. It has a good prognosis, if the biceps function is present early.

In lower plexus injury, the hand is involved, with deformity of the fingers. An ipsilateral Horner’s Syndrome consisting of ptosis, myosis, anhidrosis, and enophthalmos may occur indicating a preganglionic injury of the T1 cervical sympathetic nerve. This condition has a poor prognosis.

Management is therapy and gentle, passive range-of-motion exercises to preserve motion in the shoulder and prevent muscle contractures and joint incongruency in the early neonatal period while awaiting return of neurologic function and motor reinnervation. Early surgical intervention for the brachial plexus is indicated in infants who did not recover elbow flexion by 3 months of age, as they are anticipated to have a poor chance of full recovery.

Surgical intervention includes microsurgical repair procedures in the form of neurolysis, nerve transfer, or nerve grafts. Later orthopedic reconstruction such as muscle rebalancing procedures may be considered to improve function around the shoulder.

Cerebral Palsy. Cerebral palsy results from an injury to the brain, which may be associated with mental impairment. Cerebral palsy is classified as spastic, athetotic, or ataxic and may present with spasticity, hemiplegia, diplegia, or scoliosis. The typical cerebral palsy patient is hyperreflexic with increased muscle tone and spasm. Treatment includes tendon lengthening procedures, release of contractures, and tendon transfers to maintain motion and function.

Hip dislocation or subluxation results from unbalanced muscle forces in many cerebral palsy patients. Early treatment consists of soft tissue releases in the form of adductor tendon releases, iliofemoral releases, and immobilization in an abduction brace.

In older children with severe deformity, bony procedures in the form of open reduction and femoral or acetabular osteotomies are usually required. Femoral head resection is considered to be a salvage procedure in nonambulatory patients with painful dislocated hips.

Knee flexion contractures are treated with hamstring muscle lengthenings and immobilization in knee extension braces.

Foot and ankle deformities are treated even in nonambulatory patients to facilitate shoe wear. The most common foot deformity in cerebral palsy is an equinovarus foot caused by heel cord contracture and peroneal spasm. Tendon balancing is usually necessary, and bony reconstruction may also be needed in severe cases.

Skeletal Growth

Injury, inflammatory disease, and developmental disorders in actively growing bones requires special attention to preserve the growth plates. The pediatric skeleton is incompletely ossified making the diagnosis of an injury difficult, since significant portions of the skeleton are invisible on radiographs. The epiphysis, generally containing an articular surface, is found at the ends of the long bone. The physis, or growth plate, is found beneath the epiphysis. The physis is divided into specific zones: the reserve zone, the zone of proliferation, and the hypertrophic zone. The hypertrophic zone has three phases: the maturation zone, the degenerative zone, and the zone of calcification (Fig. 43-51).

Injury or insult to the growth plate can lead to premature growth arrest or angular deformity of the limb. Surrounding the metaphyseal and diaphyseal bone is the periosteum. This metabolically active layer of tissue synthesizes new bone onto the diaphyseal and metaphyseal bone and provides circumferential growth of the bones.

Ossification centers in the epiphysis appear in a predictable order and can help determine “bone age.”

---

Figure 43-50. Erb’s point.

Figure 43-51. Different zones of the growth plate.
**Pediatric Fractures**

In a pediatric patient, the epiphyseal growth plate is unossified and weak and is at risk of fracture. Reduction and stabilization of epiphyseal fractures is critical to minimize permanent growth disturbances and deformity. Fractures near the growth plate have significant potential to remodel. For example, 80% of the growth of the humerus occurs from the proximal humeral growth plate; therefore, severely displaced proximal humeral fracture can remodel in the younger age group.

**Classification of Growth Plate Injuries**

Salter and Harris described a useful classification for epiphyseal fractures (Fig. 43-52). A Salter-Harris type I injury is a simple transverse fracture through the physis. A Salter-Harris type II fracture contains a component of fracture through the growth plate in continuity with a fracture of the metaphysis. A Salter-Harris type III fracture occurs through the epiphysis and exits through the growth plate, while a Salter-Harris type IV fracture extends through the physis from the metaphysis into the epiphysis. A Salter-Harris type V fracture is a crushing injury to the physis. Type III and type IV involve the joint.

Treatment of growth plate fracture requires anatomical reduction of the fragments, closed or open. If internal fixation is used, avoid placing the hardware across the growth plate to minimize the chance of injury and premature growth plate closure. When hardware needs to be placed across the physis, it should be limited to smooth K-wires. The most common complication is a physeal arrest resulting in leg length discrepancy (LLD) and/or angular deformity. Complete arrest will lead to LLD. Partial arrest will result in angulation in the area of the bar, bridge, and fusion. If less than 50% of the physis is involved and the patient has two years of growth remaining, the bar is usually resected with interposition of fat graft. If the bar (fusion area) is more than 50%, the surgeon will complete the arrest on the same side and will do a contralateral epiphysiodesis on the other extremity.

Distal femur physeal fractures are known to have a high rate of leg length discrepancy and angular deformity. The injury will need an anatomical reduction and close follow-up. Parents need to be counseled about the poor prognosis associated with these fractures.

**Diaphyseal Injuries in a Pediatric Patient**

Long bone diaphyseal fractures are generally treated closed. Pediatric patients are capable of extensive remodeling so that an angular deformity within the plane of an adjacent joint is often completely remodeled by the growth of the child. Older children do not remodel as well. A 10° angulation in both bones of the forearm in a child over an age of 10 years may cause significant limitation of rotation of the forearm. When internal fixation of a diaphyseal fracture is required, fixation through the physis is avoided.

**Fractures of the Pediatric Hip**

Fractures of the pediatric hip can occur with high-energy trauma, and there is a high rate of avascular necrosis. Pediatric patients with hip fractures may be treated with a spica cast. The spica cast includes the abdomen, lower back, pelvis, and lower limb, and derives its name from the resemblance of the plaster wrap over the hip to wheat “spica.” Closed or open reduction and internal fixation is done in fractures with severe displacement. Avascular necrosis is the most common complication after hip fractures. The incidence depends on the age of the patient and the type of the fracture. Children between 3 and 8 years old with very proximal fractures such as transphyseal fractures have the highest incidence of avascular necrosis.

**Fractures of the Femoral Shaft**

Consider child abuse if a femoral shaft fracture occurs before the walking age. Femoral shaft fractures in a child younger than 6 months are usually treated by a Pavlik harness or spica cast. A child between 6 months and 5 years with an acceptable shortening of the femur is usually treated by an immediate spica cast. The child between 5 years and 11 years is usually treated by surgery. If the fracture is transverse, flexible IM nails may be used, especially if the child weighs less than 100 lbs (45 kg). If the fracture is too proximal or too distal, or if the fracture is comminuted and unstable, a submuscular bridge plate is usually used; alternatively, an external fixator may also be used, especially in multiple trauma patients. If the patient is older than 11 years, an interlocking IM rod with a lateral trochanteric entry is used. Insertion of IM rod in younger children can cause avascular necrosis of the femoral head due to interruption of the blood supply. Refracture of the femur is a risk after using an external fixator. Overgrowth of the injured femur with leg length discrepancy can occur in children between 2 and 10 years of age.
Pediatric Ankle Fractures
Pediatric ankle fractures include several types. Salter-Harris type I and type II usually involve the fibula, and the fracture may not be apparent. The patient may present with pain and swelling. Salter-Harris type III usually involves fracture of the medial malleolus or avulsion of the anterior inferior tibiofibular ligament from the tibia. It is called a Tillaux fracture. Tillaux fractures occur because the lateral part of the ankle is not fused and it is weak. Triplane fractures are complex ankle fractures in older children as a result of partial closure of the growth plate, and they appear as a Salter II in the lateral view and as a Salter III in an AP view (Fig. 43-53). Salter-Harris I and II fractures are usually managed with casting. Salter-Harris III or IV fractures are usually managed by closed or open reduction and internal fixation. Smooth percutaneous pins or screws are utilized, avoiding the physis.

Pediatric Elbow Fractures
Management of pediatric elbow fractures is complex. Familiarity with the timing of the ossification centers’ appearance aids in diagnosis. Distal humeral physeal separation can occur from child abuse and can be mistaken for an elbow dislocation. A lateral condylar fracture of the elbow is a significant injury, and when it is displaced it will need anatomical surgical reduction. Medial epicondyle fractures of the elbow are usually treated conservatively unless they are severely displaced. It is associated with elbow dislocation in 50% of cases. When the elbow is reduced, the fragment may lodge in the joint itself and must be removed and fixed. In supracondylar fractures of the humerus (Fig. 43-54), the neurovascular status of the extremity must be assessed carefully before, during, and after treatment. The anterior interosseous nerve could be injured, and the patient may not be able to make an “OK sign” (Fig. 43-55). The brachial artery may also be injured. Closed reduction, possible open reduction, and percutaneous pinning is usually done for these fractures. The procedure should be done emergently if there is concern about the vascular status of the extremity. Close follow-up for maintenance of reduction and neurovascular status is needed.

DEVELOPMENTAL DISEASE
Developmental Dysplasia of the Hip
Developmental dysplasia of the hip (DDH) involves a spectrum of disease that includes dysplasia, subluxation, or dislocation of the hip. Teratologic hip dislocation is a different entity in which the hip is dislocated in utero and irreducible on neonatal examination, usually associated with neuromuscular conditions and genetic syndromes. Developmental dysplasia of the hip is most often seen in firstborn females with a positive family history or with breech birth.

Untreated hip dislocations can lead to a dysplastic acetabulum, and they should be recognized and treated early. Newborns are examined for hip instability within the first 72 hours
of life. Ortolani’s test consists of gentle elevation and abduction of the femur causing a palpable click in the relocation of a dislocated hip. Barlow’s test is gentle adduction and depression of the femur, which causes a palpable click as the hip slips into a dislocated position. In older infants (older than 3 months), limited abduction of the involved hip is an important finding. Infants with a dislocated or dislocatable hip will have apparent length discrepancies of the femur when the hip is positioned at 90° (Galeazzi test).

Since the bones are not ossified at birth, X-ray images of the acetabulum and femoral head are not reliable for diagnosis. Ultrasound is the imaging modality of choice in the neonatal period and can often demonstrate a dislocated or dislocatable hip.

**Treatment of DDH**
The main goal in the treatment of DDH is to achieve stable concentric reduction of the hip.

- **Neonate to 6 months:** Early treatment with abduction and flexion in a Pavlik harness for 6 to 12 weeks is usually sufficient. Avoid severe abduction and flexion in the Pavlik harness to avoid the risk of avascular necrosis of the femoral head and femoral nerve palsy.
- **Children 6 to 18 months:** Closed reduction and application of hip spica cast is indicated in this age group and in those children who failed Pavlik harness treatment.
- **Children older than 18 months:** Open reduction and capsulorrhaphy is indicated in this age group. A variety of procedures, including femoral shortening and pelvic osteotomies, are done in older age groups and in more severe cases. Osteonecrosis of the femoral head is a possible complication of treatment and can result in pain and decreased range of motion.

**Legg-Calvé-Perthes Disease**
Osteonecrosis of the proximal femoral epiphysis can cause flattening of the femoral head called Legg-Calvé Perthes disease. The age at presentation is between 4 and 8 years of age and occurs more in males, usually affecting one side. Younger age at presentation (less than 6 years old) will have a better prognosis. The patient presents with groin or knee pain, decreased hip motion, and a limp. Treatment includes traction, physical therapy, abduction exercises, and crutches. Restoration of range of motion is important. Femoral and pelvic osteotomies may be needed in extreme cases and in older children.

**Slipped Capital Femoral Epiphysis**
Children ages 10 to 16 years can develop displacement of the epiphysis on the femoral neck with no history of injury. The slippage occurs through the weak zone (hypertrophic zone) of the growth plate. When slippage occurs in young patients, check for endocrine disorders such as hypothyroidism, renal osteodystrophy, and growth hormone deficiency. Slipped capital femoral epiphysis (SCFE) is associated with African-American heritage and obesity, and it is more common in boys than in girls. One-quarter of cases are bilateral. In patients with endocrine etiology, the condition is usually bilateral. Patients generally present with groin and anterior thigh pain, and the patient may have antalgic gait and a limp. Patient may present with knee pain that can lead to missing the diagnosis. In pediatric patients with knee pain, the ipsilateral hip should be assessed as well.

Examination of the patient will show obligatory external rotation with flexion and loss of internal rotation of the hip. Obtain AP and frog leg lateral views of both the hips. Slipped epiphysis is classified as either stable or unstable on the basis of the patient’s ability to bear weight. It is classified as stable if the patient is able to bear weight and the risk of osteonecrosis is less than 10%. It is classified as unstable if the patient is unable to bear weight even with crutches, and the incidence of avascular necrosis is high.

Treatment for slipped capital femoral epiphysis patients is percutaneous screw fixation through the femoral neck to engage the epiphysis, causing the growth plate to close. Reduction of the slipped epiphysis is not recommended because of an increased risk of avascular necrosis. One screw is usually adequate to prevent further slip.

**Lower Extremity Rotational Abnormalities**
Intoeing can result from femoral anteversion, tibial torsion, and metatarsus adductus. Mild degree of intoeing is normal in young children 3 to 5 years of age.

Excessive internal rotation of the femur will usually correct by age 8. Severe rotation with functional impairment that does not correct by age 10 or 11 may require rotational femoral osteotomy.

Tibial torsion is the most common cause of intoeing in toddlers and could be bilateral. The condition usually resolves without treatment.

Metatarsus adductus in infants will also resolve spontaneously in most cases.

**Congenital Talipes Equinovarus (Clubfoot)**
Clubfoot is a congenital disorder, and its etiology is not known. Clubfoot is a common problem associated with contractures of the medial tendons of the foot, a tight Achilles tendon, and contractures of the ankle, hindfoot, and midfoot. The foot is usually small, and it is in the equinus, varus, cavus, and adduction position. Talipes equinovarus can be corrected by sequential corrective casting of the foot. The serial manipulation and the casting technique is called the Ponseti technique, and it has a high success rate. A successful program of casting may be complete in 1 to 5 months. In patients with severe disease or who initiate treatment after 9 months of age, surgical release of contracted soft tissues may be necessary. The procedure is called posteromedial soft tissue release and tendon lengthening.

**Osgood-Schlatter Disease**
Osgood-Schlatter disease is a common problem most often seen in athletically active adolescents, especially in sprinters and jumpers. It is a traction apophysitis of tibial tubercle (Figs. 43-56 and 43-57). One must know the difference between the epiphysis, apophysis, and physes. This disorder is characterized by ossification in the distal patellar tendon at the point of its tibial insertion, and it is thought to result from mechanical stress on the tendinous insertion. The disease presents with severe local pain and tenderness in the area of the tibial tubercle. Radiographs may show calcified ossicles within the tendon at its insertion.
Treatment for the disease is activity restriction and anti-inflammatory drugs. The majority of patients improve with conservative treatment, and athletic participation can be resumed. Usually, symptoms regress after skeletal maturity or after activity modification. In refractory cases, surgery in the form of ossicle excision is rarely done.

BIBLIOGRAPHY

Entries highlighted in bright blue are key references.

Trauma


SPECIFIC CONSIDERATIONS

PART II


**Sports**


Spine


### Joint Reconstruction


Maratt JD, Gagnier JJ, Butler PD, Hallstrom BR, Urquhart AG, Roberts KC. No difference in dislocation seen in anterior vs posterior approach total hip arthroplasty. *J Arthroplasty.* 2016;31(9):127-130.


**Pediatric Orthopedics**


## INTRODUCTION

The highly mobile, functional, and strong hand is a major distinguishing point between humans and the nonhuman primates. The hand is an essential participant for activities of daily living, vocation, and recreational activities. The hand is even adaptable enough to read for the blind and speak for the mute. The underlying goal of all aspects of hand surgery is to maximize mobility, sensibility, stability, and strength while minimizing pain. These goals are then maximized to the extent possible given the patient’s particular pathology. Hand surgery is a regional specialty.

Hand surgeons integrate components of neurologic, orthopedic, plastic, and vascular surgery in the care of patients with disorders of the upper extremities.

## ANATOMY OF THE HAND AND WRIST

In order to understand any disorder of the hand, one must understand the anatomy of the underlying structures. Examination of the hand is based on demonstrating the function or lack thereof of each of these structures.

### Bones

The hand is highly mobile in space to allow maximum flexibility in function. As such, a number of directions particular to the hand are necessary in order to properly describe position, motion, and so on. Palmar (or volar) refers to the anterior surface of the hand in the anatomic position; dorsal refers to the posterior surface in the anatomic position. The hand can rotate at the wrist level; rotation to bring the palm down is called...
pronation, and rotation to bring the palm up is called supination. Because the hand can rotate in space, the terms medial and lateral are avoided. Radial and ulnar are used instead as these terms do not vary with respect to the rotational position of the hand. Abduction and adduction, when used on the hand, refer to movement of the digits away from and toward the middle finger, respectively (Fig. 44-1).

The hand is comprised of 19 bones arranged in five rays. A ray is defined as a digit (finger or thumb) from the metacarpal base to the tip of the digit (Fig. 44-2A). The rays are numbered 1 to 5, beginning with the thumb. By convention, however, they are referred to by name: thumb, index, middle, ring, and small. There are five metacarpals, comprising the visible palm of the hand. Each digit has a proximal and a distal phalanx, but only the fingers have a middle phalanx as well. The metacarpophalangeal (MP) joint typically allows 90° of flexion with a small amount of hyperextension. In addition, the fingers can actively abduct (move away from the middle finger) and adduct (move toward the middle finger). The thumb, in contrast, moves principally in the flexion-extension arc at the MP joint. Although there can be laxity in the radial and ulnar direction, the thumb cannot actively move in these directions at the MP level. The proximal interphalangeal joint (PIP) is the critical joint for finger mobility. Normal motion is 0° to 90° (full extension to flexion). The distal interphalangeal joint (DIP) also moves only in a flexion-extension plane from 0° to 90° on average. The thumb interphalangeal joint (IP) also moves only in a flexion-extension plane. Its normal motion is highly variable between individuals, but averages 0° to 80°.

Each of the MP and IP joints has a radial and ulnar collateral ligament to support it. The IP joint collateral ligaments are on tension with the joint fully extended. For the fingers, the MP joint collateral ligaments are on tension with the joint bent 90°. Collateral ligaments have a tendency to contract when not placed on tension; this becomes relevant when splinting the hand (see later “Trauma” section on splinting).

The wrist consists of eight carpal bones divided into two rows (see Fig. 44-2B). The proximal row consists of the scaphoid, lunate, and triquetrum. The lunate is the principle axis of motion of the hand onto the forearm. It bears approximately 35% of the load of the wrist onto the forearm. The scaphoid is shaped like the keel of a boat and bears 55% of the load of the hand onto the forearm, but it also serves as the principle link between the proximal and distal rows, allowing for motion while maintaining stability. Both the scaphoid and the lunate articulate with the radius. The triquetrum resides ulnar to the lunate. It does not interact with the ulna proximally; rather, it interacts with a cartilage suspended between the ulnar styloid and the distal radius called with triangular fibrocartilage complex (TFCC) (see Fig. 44-2B). The remaining 10% of load of the hand onto the forearm is transmitted through the TFCC.

The distal row consists of four bones. The trapezium resides between the scaphoid and the thumb metacarpal. Distally, it has a saddle-shaped surface, which interacts with a reciprocally saddle-shaped base of the thumb metacarpal to allow for high mobility of the thumb carpometacarpal (CMC) joint in radial-ulnar and palmar-dorsal directions and opposition (Fig. 44-1B). The trapezoid rests between the scaphoid and the index finger metacarpal. The capitate, the largest carpal bone and first to ossify in a child, lies between the lunate and the middle finger metacarpal, but it also interacts with the scaphoid on its proximal radial surface. The index and middle finger CMC joints are highly stable and have minimal mobility. The hamate is the ulnar-most bone in the distal row, sitting between the triquetrum proximally and the ring and small finger metacarpals distally. The ring and small finger CMC joints are mobile, principally in the flexion-extension direction.

The pisiform is a carpal bone only by geography. It is a sesamoid bone within the FCU tendon (see following section). It does not bear load and can be excised, when necessary, without consequence.

Muscles Affecting the Hand and Wrist

The wrist is moved by multiple tendons that originate from the forearm and elbow. The digits of the hand are moved by both intrinsic (originating within the hand) and extrinsic (originating in the forearm) muscles. All of these muscles are innervated by the median, radial, or ulnar nerves (or their branches) (Fig. 44-3).

Three muscles flex the wrist, all of which originate from the medial epicondyle of the humerus. The flexor carpi radialis (FCR, median nerve) inserts on the volar base of the index finger metacarpal. The flexor carpi ulnaris (FCU, ulnar nerve) also originates from the proximal ulna and inserts on the volar base of the small finger metacarpal. The palmaris longus (PL) tendon does not insert on a bone; it inserts on the palmar fascia, located deep to the skin in the central proximal palm, and is absent in up to 15% of patients. The FCR also deviates the wrist radially, whereas the FCU deviates the wrist ulnarly.

All three wrist extensors are innervated by the radial nerve or its branches. The extensor carpi radialis longus (ECRL)
originates from the distal shaft of the humerus and inserts on the dorsal base of the index finger metacarpal. The extensor carpi radialis brevis (ECRB) originates from the lateral epicondyle of the humerus and inserts on the dorsal base of the middle finger metacarpal. The extensor carpi ulnaris (ECU) also originates from the lateral epicondyle of the humerus and inserts on the dorsal base of the small finger metacarpal. The ECRL deviates the wrist radially, whereas the ECU deviates the wrist ulnarly.

The long flexors of the fingers all originate from the medial epicondyle of the humerus. The flexor digitorum superficialis (FDS) inserts on the base of the middle phalanx of each finger and primarily flexes the PIP joint. The flexor digitorum profundus (FDP) inserts on the base of the distal phalanx and primarily flexes the DIP joint. The flexor pollicis longus (FPL) originates more distally, from the ulna, radius, and interosseous membrane between them in the forearm. It inserts on the base of the distal phalanx of the thumb and primarily flexes the IP joint. All of these tendons can also flex the more proximal joint(s) in their respective rays. All of these muscles are innervated by the median nerve (or its branches) except the FDP to the ring and small fingers, which are innervated by the ulnar nerve.

The extrinsic extensors of the fingers and thumb are all innervated by the posterior interosseous nerve (PIN, branch of the radial nerve). The extensor digitorum communis (EDC) originates from the lateral epicondyle of the humerus and extends the MP joints of the fingers. Unlike most tendons that attach directly into a bone, the EDC tendons do not insert on the dorsal base of the proximal phalanx, but rather into a soft tissue sling called the sagittal hood, which surrounds the proximal phalanx base and pulls up on the volar surface in a

Figure 44-1. Directions of finger, hand, and wrist motion. A. Finger abduction (white arrows) and adduction (black arrows). B. Thumb radial (black arrow) and palmar (white arrow) abduction. C. Thumb and small finger opposition. D. Hand/wrist pronation (black arrow) and supination (white arrow).
hammock-like manner. More distally in the dorsal forearm, the extensor indices proprius (EIP) and extensor digiti quinti (EDQ) originate from the ulna, radius, and posterior interosseous membrane and insert on the sagittal hood of the index and small fingers, respectively.

The thumb has three separate extrinsic extensors. All of these originate from the dorsal ulna in the mid-forearm and are innervated by the PIN. The abductor pollicis longus (APL) inserts on the radial base of the thumb metacarpal to produce some extension, but mostly abduction. The extensor pollicis...
The intrinsic muscles of the hand are what allow humans fine, subtle movements of the hand. Microsurgery, typing, and even video gaming would be difficult, if not impossible, without them.

The thenar muscles originate from the volar radial surface of the scaphoid and trapezium and the flexor retinaculum. The abductor pollicis brevis (APB) inserts on the radial base of the thumb proximal phalanx and abducts the thumb in a radial and volar direction. The opponens pollicis (OP) inserts on the radial distal aspect of the thumb metacarpal and draws the thumb across the palm toward the small finger. The flexor pollicis brevis (FPB) inserts on the base of the thumb proximal phalanx and flexes the thumb MP joint. The APB, OP, and superficial head of the FPB are all innervated by the thenar motor branch of the median nerve.

The lumbrical muscles are unique in the body in that they originate from a tendon. Each finger’s lumbral originates from the FDP tendon in the palm. The lumbral tendon passes along the radial aspect of the digit to flex the MP and extend the IP joints. The index and middle lumbricals are median nerve innervated, and the ring and small finger lumbricals are ulnar nerve innervated.

The hypothenar muscles originate from the pisiform, hamate, and flexor retinaculum and insert on the ulnar base of the small finger proximal phalanx. The adductor digiti quinti (ADQ) abducts the small finger. The opponens digiti quinti (ODQ) brings the small finger across the palm in reciprocal motion to the OP. The flexor digiti quinti (FDQ) flexes the small finger metacarpal. All of these muscles are innervated by the ulnar nerve.

The interosseous muscles occupy the space between the metacarpal bones. Their tendons insert on the bases of the proximal phalanges. All act to flex the MP joints and extend the IP joints. The three palmar interosseous muscles adduct the fingers. The four dorsal interosseous muscles abduct the fingers. The adductor pollicis originates from the middle finger metacarpal and inserts on the ulnar base of the thumb proximal phalanx. It acts to adduct the thumb. All of these muscles, as well as the deep head of the FPB, are innervated by the ulnar nerve.

**Tendons and Pulleys**

Multiple pulleys pass over or surround the extrinsic tendons en route to or within the hand. Their purpose is to maintain tendon position near the bone, allowing maximal translation of tendon excursion into joint motion.

The most well known of the wrist-level pulleys is the flexor retinaculum, also known as the transverse carpal ligament. It attaches to the scaphoid tubercle and trapezium radially and the hook of the hamate bone and pisiform ulnarly. Deep to this ligament, between the scaphoid (radially) and the hamate (ulnarly), pass the FDS, FDP, and FPL tendons as well as the median nerve. This area is also known as the carpal tunnel (see Fig. 44-3).

On the dorsum of the wrist, the extensor retinaculum is divided into six compartments. Beginning on the radial aspect of the radius, the first compartment contains the APL and EPB tendons. The second holds the ECRL and ECRB tendons. The EPL passes through the third compartment. The fourth compartment contains the EIP and EDC tendons, the fifth the EDQ, and the sixth the ECU. The sixth compartment is located on the ulnar aspect of the distal ulna. Although the compartments end at the radiocarpal/ulnocarpal joints, the relative geography of the tendons is preserved over the carpal bones (see Fig. 44-3).

In the hand, the pulleys maintain the long flexor tendons in close apposition to the fingers and thumb. There are no extensor pulleys within the hand. Each finger has five annular and three cruciate pulleys (Fig. 44-4). The second and fourth (A2 and A4) pulleys are the critical structures to prevent bowstringing of the finger. The remaining pulleys can be divided as needed for surgical exposure or to relieve a stricture area.

**Vascular**

Two major arteries serve the hand. The radial artery travels under the brachioradialis muscle in the forearm. At the junction of the middle and distal thirds of the forearm, the artery becomes superficial and palpatable, passing just radial to the FCU tendon. At the wrist level, the artery splits into two branches. The smaller, superficial branch passes volarly into the palm to contribute to the superficial palmar arch. The larger branch passes dorsally over the scaphoid bone, under the EPL and EPB tendons (known as the anatomic snuffbox) and back volarly between the proximal thumb and index finger metacarpals to form the superficial palmar arch.

The ulnar artery travels deep to the FCU muscle in the forearm. When the FCU becomes tendinous, the ulnar artery resides deep and slightly radial to it. At the wrist, the artery travels between the hamate and pisiform bones superficial to the transverse carpal ligament (known as Guyon’s canal) into the palm. The larger, superficial branch forms the superficial...
palmar arch. The deeper branch contributes to the deep palmar arch (Fig. 44-5A). In 97% of patients, at least one of the deep or superficial palmar arches is intact, allowing for the entire hand to survive on the radial or ulnar artery.5

Each digit receives a radial and ulnar digital artery. For the thumb, the radial digital artery may come from the deep palmar arch or the main body of the radial artery. The larger ulnar digital artery comes off the deep arch as either a discrete unit, the princeps pollicis artery, or less frequently as the first common digital artery, which then splits into the radial digital artery to the index finger and the ulnar digital artery to the thumb. The second, third, and fourth digital arteries typically branch off the superficial palmar arch and pass over the similarly named interosseous spaces respectively, ultimately dividing into two proper digital arteries each. The ulnar digital artery of the small finger comes off as a separate branch from the superficial arch. Within the finger, the proper digital arteries travel lateral to the bones and tendons, just palmar to the midaxis of the digit, but dorsal to the proper digital nerves (Fig. 44-5B).

Nerve
Three principal nerves serve the forearm, wrist, and hand: the median, radial, and ulnar nerves. The most critical of these from a sensory standpoint is the median nerve. The median nerve begins as a terminal branch of the medial and lateral cords of the brachial plexus. It receives fibers from C5–T1. The palmar cutaneous branch of the median nerve separates from the main body of the nerve 6 cm proximal to the volar wrist crease and serves the proximal, radial-sided palm. The main body of the median nerve splits into several branches after the carpal tunnel: a radial digital branch to the thumb, an ulnar digital nerve to the thumb, and a radial digital nerve to the index finger (sometimes beginning as a single first common digital nerve); the second common digital nerve that branches into the ulnar digital nerve to the index finger and the radial digital nerve to the middle finger; and a third common digital nerve that branches into the ulnar digital nerve to the middle finger and a radial digital nerve to the ring finger. The digital nerves provide volar-sided sensation from the metacarpal head level to the tip of the digit. They also, through their dorsal branches, provide dorsal-sided sensation to the digits from the midportion of the middle phalanx distally via dorsal branches. The thenar motor branch of the median nerve most commonly passes through the carpal tunnel and then travels in a recurrent fashion back to the thenar muscles. Less commonly, the nerve passes through or proximal to the transverse carpal ligament en route to its muscles.

In the forearm, the median nerve gives motor branches to all of the flexor muscles except the FCU, and the ring and small finger portions of the FDP. Distal median motor fibers (with the exception of those to the thenar muscles) are carried through a large branch called the anterior interosseous nerve.

The ulnar nerve is a terminal branch of the medial cord of the brachial plexus. It receives innervation from C8 and T1 roots. The FCU and FDP (ring/small) receive motor fibers from the ulnar nerve. In the distal forearm, 5 cm above the head of the ulna, the nerve gives off a dorsal sensory branch. Once in the hand, the nerve splits into the motor branch and sensory branches. The motor branch curves radially at the hook of the hamate bone to innervate the intrinsic muscles, as described earlier. The sensory branches become the ulnar digital nerve to the small finger and the fourth common digital nerve, which splits into the ulnar digital nerve to the ring finger and the radial digital nerve to the small finger. The sensory nerves provide distal dorsal sensation similar to the median nerve branches.

The radial nerve is the larger of two terminal branches of the posterior cord of the brachial plexus. It receives fibers from C5–T1 nerve roots. It innervates all of the extensor muscles of the forearm and wrist through the PIN branch except for the ECRL, which is innervated by the main body of the radial nerve in the distal upper arm. There is no ulnar nerve contribution to extension of the wrist, thumb, or finger MP joints. As noted earlier, the ulnar innervated intrinsic hand muscles are the principle
extensors of the finger IP joints, although the long finger extensors (EDC, EIP, EDQ) make a secondary contribution to this function.

In the proximal dorsal forearm, the superficial radial nerve (SRN) is the other terminal branch of the radial nerve. It travels deep to the brachioradialis muscle until 6 cm proximal to the radial styloid, where it becomes superficial. The SRN provides sensation to the dorsal hand and the radial three and a half digits up to the level of the mid-middle phalanx (where the dorsal branches of the proper digital nerves take over, as described earlier). The dorsal branch of the ulnar nerve provides sensation to the ulnar one and a half digits and dorsal hand in complement to the SRN.

**HAND EXAMINATION**

**Emergency Department/Inpatient Consultation**

A common scenario in which the hand surgeon will be introduced to the patient is in trauma or other acute situations. The patient is evaluated by inspection, palpation, and provocative testing.

On inspection, one should first note the position of the hand. The resting hand has a normal cascade of the fingers, with the small finger flexed most and the index finger least (Fig. 44-6). Disturbance of this suggests a tendon or skeletal problem. Also note any gross deformities or wounds and what deeper structures, if any, are visible in such wounds. Observe for abnormal coloration of a portion or all of the hand (this can be confounded by ambient temperature or other injuries), edema, and/or clubbing of the fingertips.

Palpation typically begins with the radial and ulnar artery pulses at the wrist level. Pencil Doppler examination can supplement this and evaluate distal vessels. A pulsatile signal is normally detectable by pencil Doppler in the pad of the finger at the center of the whorl of creases. Discrepancies between digits should be noted. If all other tests are inconclusive, pricking the involved digit with a 25-gauge needle should produce bright red capillary bleeding. If an attached digit demonstrates inadequate or absent blood flow (warm ischemia), the urgency of completing the evaluation and initiating treatment markedly increases.

Sensation must be evaluated prior to any administration of local anesthetic. At a minimum, light and sharp touch sensation should be documented for the radial and ulnar aspects of the tip of each digit. Beware of writing “sensation intact” at the conclusion of this evaluation. Rather, one should document what was tested (e.g., “light and sharp touch sensation present and symmetric to the tips of all digits of the injured hand”). For a more detailed evaluation of hand sensation, two-point discrimination may be assessed using a bent paperclip or monofilament. In the setting of a sharp injury, sensory deficit implies a lacerated structure until proven otherwise. Once sensation has been evaluated and documented, the injured hand can be anesthetized for patient comfort during the remainder of the examination (see below).

Ability to flex and extend the wrist and digital joints is typically examined next. At the wrist level, the FCR and FCU tendons should be palpable during flexion. The wrist extensors are not as readily palpated due to the extensor retinaculum. Ability to flex the DIP joint (FDP) is tested by blocking the finger at the middle phalanx level. To test the FDS to each finger, hold the remaining three fingers in slight hyperextension and ask the patient to flex the involved digit (Fig. 44-7). This maneuver makes use of the fact that the FDP tendons share a common muscle belly. Placing the remaining fingers in extension prevents the FDP from firing, and allows the FDS, which has a separate muscle belly for each tendon, to fire. Strength in grip, finger abduction, and thumb opposition is tested and compared to the uninjured side. Range of motion for the wrist, MP, and IP joints should be noted and compared to the opposite side.

If there is suspicion for closed space infection, the hand should be evaluated for erythema, swelling, fluctuance, and localized tenderness. The dorsum of the hand does not have fascial septae; thus, dorsal infections can spread more widely than palmar ones. The epitrochlear and axillary nodes should be palpated for enlargement and tenderness. Findings for specific infectious processes will be discussed in the “Infections” section.
Additional exam maneuvers and findings, such as those for office consultations, will be discussed with each disease process covered later in this chapter.

HAND IMAGING

Plain X-Rays
Almost every hand evaluation should include plain X-rays of the injured or affected part. A standard, anteroposterior, lateral, and oblique view of the hand or wrist (as appropriate) is rapid, inexpensive, and usually provides sufficient information about the bony structures to achieve a diagnosis in conjunction with the symptoms and findings.6

Lucencies within the bone should be noted. Most commonly, these represent fractures, but they can on occasion represent neoplastic or degenerative processes. Great care should be taken to evaluate the entire X-ray, typically beginning away from the area of the patient’s complaint. Additional injuries can be missed, which might affect the treatment plan selected and eventual outcome.

Congruency of adjacent joints should also be noted. The MP and IP joints of the fingers should all be in the same plain on any given view. Incongruency of the joint(s) of one finger implies fracture with rotation. At the wrist level, the proximal and distal edge of the proximal row and proximal edge of the distal row should be smooth arcs, known as Gilula’s arcs (Fig. 44-8A). Disruption of these implies ligamentous injury or possibly dislocation (Fig. 44-8B).7

Computed Tomography
Computed tomography (CT) scanning of the hand and wrist can provide additional bony information when plain X-rays are insufficient. Comminuted fractures of the distal radius can be better visualized for number and orientation of fragments. Scaphoid fractures can be evaluated for displacement and comminution preoperatively as well as for the presence of bony bridging postoperatively (Fig. 44-9). Recent studies have suggested that in the setting of suspected scaphoid fractures with negative radiographs, the use of CT scans may decrease the healthcare costs and patient morbidity.8 CT scans are also useful for CMC fractures of the hand where overlap on a plain X-ray lateral view may make diagnosis difficult.

Unlike the trunk and more proximal extremities, CT scans with contrast are less useful to demonstrate abscess cavities due to the small area of these spaces.

Ultrasoundography
Ultrasoundography has the advantages of being able to demonstrate soft tissue structures and being available on nights and weekends. Unfortunately, it is also highly operator dependent. In the middle of the night when magnetic resonance imaging (MRI) is not available, ultrasound may be able to demonstrate a

Figure 44-7. The examiner holds the untested fingers in full extension, preventing contracture of the flexor digitorum profundus. In this position, the patient is asked to flex the finger, and only the flexor digitorum superficialis will be able to fire.

Figure 44-8. Gilula’s arcs are seen shown in this normal patient (A) and in a patient with a scaphoid fracture and perilunate dislocation (B).
large deep infection in the hand but is rarely more useful than a thorough clinical examination. Additionally, the use of dynamic ultrasound may be used to evaluate tendon motion and aid in the diagnosis of tendon pathology or injury.9

Magnetic Resonance Imaging
MRI provides the best noninvasive visualization of the soft tissue structures. With contrast, MRI can demonstrate an occult abscess. Unfortunately, it is often not available on an urgent basis for hand issues when this information is often needed. MRI can also demonstrate soft tissue injuries such as cartilage or ligament tears or tendonitis (usually by demonstrating edema in the area in question). It can demonstrate occult fractures that are not sufficiently displaced to be seen on X-ray or CT (again, by demonstrating edema). MRI can also demonstrate vascular disturbance of a bone, as in a patient with avascular necrosis of the scaphoid (Fig. 44-10).

Angiography
Angiography of the upper extremity is rarely used. In many centers, MRI and CT angiography provide sufficient resolution of the vascular structures to make traditional angiography unnecessary. Also, primary vascular disease of the upper extremity is relatively uncommon. In the trauma setting, vascular disturbance usually mandates exploration and direct visualization of the structures in question, and angiography is thus obviated.

For a patient with vascular disease of the upper extremity, angiography of the upper extremity is usually performed through a femoral access much like with the leg. An arterial catheter can be used to deliver thrombolytic drugs to treat a thrombotic process.

TRAUMA
The upper extremity–injured patient may have additional injuries to other parts of the body. All injured patients should receive an appropriate trauma survey to look for additional injuries.

The patient with upper extremity trauma is evaluated as described in the “Hand Examination” section. Sensory examination should be performed early. Once sensory status has been documented, administration of local anesthesia can provide comfort to the patient during the remainder of the evaluation.
Fractures and Dislocations
For dislocations and displaced fractures, a visible deformity is often present. Nondisplaced fractures may not show a gross deformity but will have edema and tenderness to palpation at the fracture site. A fracture is described by its displacement, rotation, and angulation. A fracture is also described in terms of comminution and the number and complexity of fracture fragments. Displacement is described as a percentage of the diameter of the bone; rotation is described in degrees of supination or pronation with respect to the rest of the hand; angulation is described in degrees. To avoid confusion, it is useful to describe which direction the angle of the fracture points. All injuries should be evaluated for nearby wounds (open) that may introduce bacteria into the fracture site or joint space.

Once the initial force on the fracture ceases, the tendons passing beyond the fracture site provide the principal deforming force. Their force is directed proximally and, to a lesser extent, volarly. Based on this, the stability of a fracture can be determined by the orientation of the fracture with respect to the shaft of the bone. Transverse fractures are typically stable. Oblique fractures typically shorten. Spiral fractures typically rotate as they shorten and thus require surgical treatment.

Fractures of the tuft of the distal phalanx are common. Catching of a finger in a closing door is a common causative mechanism. These fractures are often nondisplaced and do not require treatment beyond protection of the distal phalanx from additional trauma while the fracture heals.

Displaced transverse fractures of the phalanges can usually be reduced with distraction. The distal part is pulled away from the main body of the hand and then pushed in the direction of the proximal shaft of the finger, and then distraction is released. Postreduction X-rays should routinely be performed to document satisfactory reduction. Oblique and spiral fractures usually are unstable after reduction. The involved digit(s) should be splinted until appropriate surgical intervention can be performed.

Articular fractures of the IP and MP joints are worrisome because they may compromise motion. Chip fractures must be evaluated for instability of the collateral ligaments. If the joint is stable, the patient should initially be splinted for comfort. Motion therapy should be instituted early (ideally within the first week) to prevent stiffness. For larger fractures, the patient should be splinted until surgical treatment can be performed. In surgery, the fracture is typically internally fixated to allow for early motion, again with the goal of preventing stiffness.

Dislocations of the PIP joints produce traction on the neurovascular structures but usually do not lacerate them. In general, the patient should not be sent home with a joint that remains dislocated. Most commonly, the distal part is dorsal to the proximal shaft and sits in a hyperextended position. For this patient, the examiner gently applies pressure to the base of the distal part until it passes beyond the head of the proximal phalanx. Once there, the relocated PIP joint is gently flexed, confirming the joint is in fact reduced. The joint is splinted in slight flexion to prevent redislocation. On occasion, the head of the proximal phalanx may pass between the two slips of the FDS tendon. For these patients, the joint may not be reducible in a closed fashion.

Angulated fractures of the small finger metacarpal neck (“boxer’s fracture”) are another common injury seen in the ER. Typical history is that the patient struck another individual or rigid object with a hook punch. These are often stable after reduction using the Jahss maneuver (Fig. 44-11).

Fractures of the thumb metacarpal base are often unstable. The Bennett fracture displaces the volar-ulnar base of the bone. The remainder of the articular surface and the shaft typically dislocate dorsoradially and shorten. The thumb often appears grossly shortened, and the proximal shaft of the metacarpal may reside at the level of the trapezium or even the scaphoid on X-ray. In a Rolando fracture, a second fracture line occurs between the remaining articular surface and the shaft. These fractures nearly always require open reduction and internal fixation.

Most nondisplaced fractures do not require surgical treatment. The scaphoid bone of the wrist is a notable exception to this rule. Due to peculiarities in its vascular supply, particularly vulnerable at its proximal end, nondisplaced scaphoid fractures can fail to unite in up to 20% of patients even with appropriate immobilization. Recent developments in hardware and surgical technique have allowed stabilization of the fracture with minimal surgical exposure. One prospective randomized series of scaphoid wrist fractures demonstrated shortening of time to union by up to 6 weeks in the surgically treated group, but no difference in rate of union. Surgery may be useful in the younger, more active patient who would benefit from an earlier return to full activity.

Ligament injuries of the wrist can be difficult to recognize. Patients often present late and may not be able to localize their pain. In severe cases, the ligaments of the wrist can rupture to the point of dislocation of the capitale off the lunate or even the lunate off the radius. Mayfield and colleagues classified the progression of this injury into four groups. In the most severe group, the lunate dislocates off the radius into the carpal tunnel. In some circumstances, the scaphoid bone may break rather than
the scapholunate ligament rupturing. Attention to the congruency or disruption of Gilula’s arcs will help the examiner to recognize this injury. For patients with type 4 (most severe) and some with type 3 injury, the examiner should also evaluate for sensory disturbance in the median nerve distribution because this may indicate acute carpal tunnel syndrome and necessitate more urgent intervention. Although the Mayfield pattern of injury is most common, force can also transmit along alternate paths through the carpus.¹⁶

After reduction of fractures and dislocations (as well as after surgical repair of these and many other injuries), the hand must be splinted in a protected position. For the fingers, MP joints should be splinted 90°, and the IP joints at 0° (called the intrinsic plus position). The wrist is generally splinted at 20° extension because this puts the hand in a more functional position. This keeps the collateral ligaments on tension and helps prevent secondary contracture. In general, one of three splints should be used for the emergency department (ED) patient (Fig. 44-12). The ulnar gutter splint places plaster around the ulnar border of the hand. It is generally appropriate for small finger injuries only. Dorsal plaster splints can be used for injuries of any of the fingers. Plaster is more readily contoured to the dorsal surface of the hand than the volar surface, particularly in the setting of trauma-associated edema. For thumb injuries, the thumb spica splint is used to keep the thumb radially and palmarly abducted from the hand. Lastly, sugar tong splints include a volar and dorsal slab that includes the elbow in order to prevent supination and pronation. Sugar tong splints are most often used in the setting of acute distal radius or ulna fractures.

Tendons
Injuries to the flexor and extensor tendons compromise the mobility and strength of the digits. On inspection, injury is normally suspected by loss of the normal cascade of the fingers. The patient should be examined as described earlier to evaluate for which tendon motion is deficient. If the patient is unable to cooperate, extension of the wrist will produce passive flexion of the fingers and also demonstrate a deficit. This is referred to at the tenodesis maneuver.

Flexor tendon injuries are described based on zones (Fig. 44-13). Up until 40 years ago, zone 2 injuries were always reconstructed and never repaired primarily due to concern that the bulk of repair within the flexor sheath would prevent tendon glide. The work of Dr. Kleinert and colleagues at the University of Louisville changed this “axiom” and established the principle of primary repair and early controlled mobilization postoperatively.¹⁷ Flexor tendon injuries should always be repaired in the operating room. Although they do not need to be repaired on the day

Figure 44-12. Commons splints used for hand injuries/surgeries. A. Ulnar gutter splint. The ring and small fingers are included and maintain an interphalangeal (IP) joint extension and metacarpophalangeal (MP) joint flexion to 90°. B. Dorsal four-finger splint. As with the ulnar gutter splint, finger MP joints are flexed to 90° with IP joints kept fully extended. C. Thumb spica splint. One easy method to fabricate is to place one slab of plaster radially over the wrist and thumb with a second square of plaster over the thenar eminence, which joins the first. D. Sugar tong splint. This dorsal and volar slab splints immobilizes the wrist and elbow in neutral and 90° positions, respectively.

Figure 44-13. The zones of flexor tendon injury. I. Flexor digitorum superficialis insertion to the flexor digitorum profundus insertion. II. Start of the A1 pulley to the flexor digitorum superficialis insertion. III. End of the carpal tunnel to the start of the A1 pulley. IV. Within the carpal tunnel. V. Proximal to the carpal tunnel.
of injury, the closer to the day of injury they are repaired, the easier it will be to retrieve the retracted proximal end in surgery. The laceration should be washed out and closed at the skin level only using permanent sutures. The hand should be splinted as described earlier; one notable difference is that the wrist should be splinted at slight flexion (about 20°) to help decrease the retracting force on the proximal cut tendon end.

Extensor tendons do not pass through a sheath in the fingers. As such, bulkiness of repair is less of a concern. With proper supervision/experience and equipment, primary extensor tendon repair can be performed in the ED.

Very distal extensor injuries near the insertion on the dorsal base of the distal phalanx may not have sufficient distal tendon to hold a suture. Closed injuries, called mallet fingers, can be treated with extension splinting of the DIP joint for 6 continuous weeks. For patients with open injuries, a dermatotendolysis suture is performed. A 2-0 or 3-0 suture is passed through the distal skin, tendon remnant, and proximal tendon as a mattress suture. Using a suture of a different color than the skin closing sutures will help prevent removing the dermatotendolysis suture(s) too soon. The DIP joint is splinted in extension.

More proximal injuries are typically repaired with a 3-0 braided permanent suture. Horizontal mattress or figure-of-eight sutures should be used, two per tendon if possible. Great care should be used to ensure matching the appropriate proximal and distal tendon ends. The patient is splinted with IP joints in extension and the wrist in extension per usual. MP joints should be splinted in 45° flexion, sometimes less. Although this position is not ideal for MP collateral ligaments, it is important for taking tension off of the tendon repairs.

Nerve Injuries
In the setting of a sharp injury, a sensory deficit implies a nerve laceration until proven otherwise. For blunt injuries, even displaced fractures and dislocations, nerves are often contused but not lacerated and are managed expectantly. Nerve repairs require appropriate microsurgical equipment and suture; they should not be performed in the ED. As with tendons, nerve injuries do not require immediate exploration. However, earlier exploration will allow for easier identification of structures and less scar tissue to be present. The nerve must be resected back to healthy nerve fascicle prior to repair. Delay between injury and repair can thus make a difference between the ability to repair a nerve primarily or the need to use a graft. The injured hand should be splinted with MPs at 90° and IPs at 0°, as described earlier.

Vascular Injuries
Vascular injuries have the potential to be limb or digit threatening. A partial laceration of an artery at the wrist level can potentially cause exsanguinating hemorrhage. Consultations for these injuries must be evaluated urgently.

Initial treatment for an actively bleeding wound should be direct local pressure for no less than 10 continuous minutes. If this is unsuccessful, an upper extremity tourniquet inflated to 100 mmHg above the systolic pressure should be used. One should keep this tourniquet time to less than 2 hours to avoid tissue necrosis. Once bleeding is controlled well enough to evaluate the wound, it may be cautiously explored to evaluate for bleeding points. One must be very cautious if attempting to ligate these to ensure that adjacent structures such as nerves are not included in the ligature.

The hand must be evaluated for adequacy of perfusion to the hand as a whole as well as the individual digits. Capillary refill, turgor, Doppler signal, and bleeding to pinprick all provide useful information regarding vascular status. The finger or hand with vascular compromise requires urgent operative exploration. Unlike the complete amputation, in which the amputated part can be cold preserved (see later section, “Amputations and Replantation”), devascularization without amputation produces warm ischemia, which is tolerated only for a matter of hours.

For the noncritical vascular injury, two treatment options exist. Simple ligation will control hemorrhage. At least one of the palmar arterial arches is intact in 97% of patients, so this will usually not compromise hand perfusion. Each digit also has two arterial inflows and can survive on one (see “Amputations and Replantation” section). In the academic hospital setting, however, consideration should be given to repairing all vascular injuries. Instructing a resident in vascular repair in the noncritical setting will produce a more skilled and prepared resident for when a critical vascular injury does arise.

ANESTHESIA

Local Anesthesia
Anesthetic blockade can be administered at the wrist level, digital level, or with local infiltration as needed. Keep in mind that all local anesthetics are less effective in areas of inflammation.

The agents most commonly used are lidocaine and bupivacaine. Lidocaine has the advantage of rapid onset, whereas bupivacaine has the advantage of long duration (average 6–8 hours). Although bupivacaine can produce irreversible heart block in high doses, this is rarely an issue with the amounts typically used in the hand. For pediatric patients, the tolerated dose is 2.5 mg/kg. This can be easily remembered by noting that when using 0.25% bupivacaine, 1 mL/kg is acceptable dosing.

A commonly held axiom is that epinephrine is unacceptable to be used in the hand. Several recent large series have dispelled this myth. Epinephrine should not be used in the fingertip and not in concentrations higher than 1:100,000 (i.e., what is present in commercially available local anesthetic with epinephrine). Beyond that, its use is acceptable and may be useful in an ED where tourniquet control may not be available. Also, because most ED procedures are done under pure local anesthesia, many patients will not tolerate the discomfort of the tourniquet beyond 30 minutes. Epinephrine will provide hemostasis and also prolong the effect of the local anesthetic.

Studies have reported that the addition of sodium bicarbonate (NaHCO₃) in order to buffer local anesthetic solutions and decrease the pain experienced during the administration of local anesthetic. This decrease in pain has been attributed to decreasing the acidity of local anesthetic solutions. In the clinical setting, the mixing of 8.4% sodium bicarbonate with 1% lidocaine with 1:100,000 epinephrine in a 1:9 ratio is adequate to provide a decrease in pain during the injection of local anesthetic.

Simple lacerations, particularly on the dorsum of the hand, can be anesthetized with local infiltration. This is performed in the standard fashion.

Blocking of the digital nerves at the metacarpal head level is useful for volar injuries distal to this point and for dorsal injuries beyond the midpoint of the middle phalanx (via dorsal branches of the proper digital nerves). Fingertip injuries are particularly well anesthetized by this technique. A digit can be anesthetized via a flexor sheath approach or via the dorsal web space (Fig. 44-14A,B).
Figure 44-14. Local anesthesia can be administered at the digital or the wrist level. A. A single injection into the flexor tendon sheath at the metacarpal head level provides complete anesthesia for the digit. B. Alternatively, one can inject from a dorsal approach into the web space on either side. C. The superficial radial nerve is blocked by infiltrating subcutaneously over the distal radius from the radial artery pulse to the distal radioulnar joint. The dorsal sensory branch of the ulnar nerve is blocked in similar fashion over the distal ulna. D. To block the ulnar nerve, insert the needle parallel to the plane of the palm and deep to the flexor carpi ulnaris tendon; aspirate to confirm the needle is not in the adjacent ulnar artery. E. To block the median nerve, insert the needle just ulnar to the palmaris longus tendon into the carpal tunnel. One should feel two points of resistance: one when piercing the skin, the second when piercing the antebrachial fascia.
Blocking one or more nerves as they cross the wrist can provide several advantages: anesthesia for multiple injured digits, avoiding areas of inflammation where the local anesthetic agent may be less effective, and avoiding injection where the volume of fluid injected may make treatment harder (such as fracture reduction). Four major nerves cross the wrist: the median nerve, SRN, ulnar nerve, and dorsal sensory branch of the ulnar nerve (Fig. 44-14C–E). When blocking the median and ulnar nerves, beware of intraneural injection, which can cause irreversible neural scarring. If the patient complains of severe paresthesias with injection or high resistance is encountered, the needle should be repositioned.

Hand Surgery Under Local Anesthesia
Wide awake hand surgery is surgery that is performed under surgeon-administered local anesthesia with field sterility but without the use of sedation or a tourniquet. A major benefit of this approach is the reduction of healthcare costs due to the elimination of an anesthesia provider and postoperative monitoring because only local anesthesia is used. Further benefits of sedation-free surgery include decreased time spent in the hospital for surgery and the ability of patients to follow instructions during surgery. This advantage is evident during flexor tendon repairs, where intraoperative active movement allows direct visualization of the tendon repair under active movement.23 Perceived weaknesses of sedation-free surgery include patient intraoperative anxiety and fear of pain during the administration of local anesthetic. A study by Davison et al, however, found that patients undergoing carpal tunnel release under wide awake local had no difference in anxiety or pain compared to patients undergoing carpal tunnel release with sedation.24

Postoperative Pain Management
Since the recognition of pain as the fifth vital sign in the early 2000s, the number of opioid prescriptions has risen dramatically. Accordingly, over the last decade, the United States has seen an increase in the number of deaths due to prescription opioid overdose. Deaths due to opioid overdose now exceed the number of deaths caused by heroin and cocaine combined. As healthcare providers, it is essential that we adequately treat postoperative pain with the minimal amount of narcotics necessary. A recent study by Rodgers et al identified that the majority of patients undergoing elective hand surgery used prescription pain medication for only 2 or fewer days after surgery. Many patients achieved adequate pain control with over-the-counter pain medication and were often left with unused opioid analgesics.25 Accordingly, there has been increased emphasis on educating prescribers on the recognition of opioid abuse and guidelines for appropriate opioid prescribing. Approaches such as multimodal pain management and opioid prescription protocols have shown to achieve adequate pain control while also reducing excess opioid prescriptions.26

SPECIAL CONSIDERATIONS

Amputations and Replantation
After replantation was first reported, replantation was attempted for nearly all amputations.27 Over the ensuing decades, more stringent guidelines have been established regarding what should be replanted. Indications for replantation include amputations of the thumb, multiple digit amputations, and amputations in children. Relative contraindications to replantation include crush injuries, injuries to a single digit distal to the PIP joint, and patients who are unable to tolerate a long surgical procedure. As with all guidelines, one should evaluate the particular needs of the injured patient.

In preparation for replantation, the amputated part and proximal stump should be appropriately treated. The amputated part should be wrapped in moistened gauze and placed in a sealed plastic bag. This bag should then be placed in an ice water bath. Do not use dry ice, and do not allow the part to contact ice directly; frostbite can occur in the amputated part, which will decrease its chance of survival after replantation. Bleeding should be controlled in the proximal stump by as minimal a means necessary, and the stump should be dressed with a nonadherent gauze and bulky dressing.

For digital amputations deemed unsalvageable, revision amputation can be performed in the ED if appropriate equipment is available. Bone prominences should be smoothed off with a rongeur and/or rasp. Great care must be taken to identify the digital nerves and resect them back as far proximally in the wound as possible; this helps decrease the chance of painful neuroma in the skin closure. Skin may be closed with permanent or absorbable sutures; absorbable sutures will spare the patient the discomfort of suture removal several weeks later. For more proximal unsalvageable amputations, revision should be performed in the operating room to maximize vascular and neural control.

Prostheses can be made for amputated parts. The more proximal the amputation, the more important to function the prosthesis is likely to be. Although finger-level prostheses are generally considered cosmetic, patients with multiple finger amputations proximal to the DIP have demonstrable functional benefit from their prosthesis as well.28

Fingertip Injuries
Fingertip injuries are among the most common pathologies seen in an ED. The usual history is that a door closed on the finger (commonly the middle, due to its increased length) or something heavy fell on the finger.

Initial evaluation should include: wound(s) including the nail bed, perfusion, sensation, and presence and severity of fractures. For the common scenario, complex lacerations with minimally displaced fracture(s) and no loss of perfusion, the wound is cleansed, sutured, and splinted in the ED. To properly assess the nail bed, the nail plate (hard part of the nail) should be removed. A Freer periosteal elevator is well suited for this purpose. Lacerations are repaired with 6-0 fast gut suture. Great care must be taken when suturing because excessive traction with the needle can further lacerate the tissue. After repair, the nail folds are splinted with the patient’s own nail plate (if available) or with aluminum foil from the suture pack. This is done to prevent scarring from the nail folds down to the nail bed that would further compromise healing of the nail.

In some situations, tissue may have been avulsed in the injury and be unavailable for repair. Choice of treatment options depends on the amount and location of tissue loss (Fig. 44-15). Historically, wounds less than 1 cm² with no exposed bone can be treated with local wound care and secondary intention. Recently, studies have reported that wounds with an average size of 1.75 cm² have healed well with excellent functional and aesthetic results.29 For larger wounds or wounds with bone exposed, one must decide if the finger is worth preserving at the current length or if shortening to allow for primary closure is a
better solution. A useful guideline is the amount of fingernail still present; if greater than 50% is present, local or regional flap coverage may be a good solution.

If sufficient local tissue is present, homodigital flaps can be considered. A wide range of antegrade and retrograde homodigital flaps can be mobilized to cover the defect. Some carry sensation or can receive nerve coaptation to recover sensation over time. For the thumb only, the entire volar skin including both neurovascular bundles can be raised and advanced distally up to 1.5 cm. The thumb receives separate vascularity to its dorsal skin from the radial artery. This flap is not appropriate for the fingers. Patients retain full sensibility in the advanced skin and can be mobilized within days of surgery (Fig. 44-16A–C).

For wounds too large to cover with homodigital tissue, regional flaps can be considered. The skin from the distal radial thenar eminence can be raised as a random pattern flap (Fig. 44-16D–F). The finger is maintained in flexion for 14 to 21 days until division of the flap pedicle and inset of the flap. Some authors have reported prolonged stiffness in patients over 30 years old, but careful flap design helps minimize this complication. Alternatively, the skin from the dorsum of the middle phalanx of an adjacent digit can be raised as a flap to cover the volar P3 (Fig. 44-16G–I). The flap is inset at 14 to 21 days. Long-term studies have shown this flap develops sensation over time.

Patients with fingertip injuries must be assessed for the possibility of salvage of the injured digit(s) taken within the context of the patient’s recovery needs and goals. The surgeon then matches the available options to the particular patient needs.

**High-Pressure Injection Injuries**

High-pressure devices are commonly used for cleaning and applications of liquids such as lubricants and paint. Most commonly, the inexperienced worker accidentally discharges the device into his nondominant hand at the base of the digit. Severity of injury depends on the amount and type of liquid injected; hydrophobic compounds cause greater damage.

These injuries are typically quite innocuous to inspection. They are, however, digit-threatening emergencies. The patient should be informed of the severity of the injury, and exploration is ideally performed within 6 hours of injury. Up to 50% of such injuries result in loss of the digit, but early recognition and treatment are associated with increased chance of digit survival. Early frank discussion with the patient and initiation of appropriate treatment produce the best results and medicolegal protection.

**Compartment Syndrome**

Compartment syndromes can occur in the forearm and/or the hand. As in other locations, these are potentially limb-threatening issues. Principle symptoms are pain in the affected compartments, tense swelling, tenderness to palpation over the compartment, and pain with passive stretch of the muscles of the compartment. Pulse changes are a late finding; normal pulses do not rule out compartment syndrome.

There are three compartments in the forearm and four groups of compartments in the hand. The volar forearm is one compartment. On the dorsum of the forearm, there is the dorsal compartment as well as the mobile wad compartment, beginning proximally over the lateral epicondyle. In the hand, the thenar and hypothenar eminences each represent a compartment. The seven interosseous muscles each behave as a separate compartment.

Compartment syndrome can be caused by intrinsic and extrinsic causes. Intrinsic causes include edema and hematoma due to fracture. Extrinsic causes include splints and dressings that are circumferentially too tight and intravenous infiltrations. Infiltrations with hyperosmolar fluids such as X-ray contrast are particularly dangerous, because additional water will be drawn in to neutralize the hyperosmolality.

Measurement of compartment pressures can be a useful adjunct to assessment of the patient. The Stryker pressure measurement device or similar device is kept in many operating rooms for this purpose. The needle is inserted into the compartment in question, a gentle flush with 0.1 to 0.2 cc of saline clears the measurement chamber, and a reading is obtained. Studies have disagreed about whether the criterion is a measured pressure (30–45 mmHg, depending on the series) or within a certain amount of the diastolic blood pressure.

Compartment releases are performed in the operating room under tourniquet control. Release of the volar forearm compartment includes release of the carpal tunnel. As the incision travels distally, it should pass ulnar and then curve back radially just before the carpal tunnel. This avoids a linear incision across a flexion crease and also decreases the chance of injury to the palmar cutaneous branch of the median nerve. One dorsal forearm incision can release the dorsal compartment and the mobile wad. In the hand, the thenar and hypothenar compartments are released each with a single incision. The interosseous compartments are released with incisions over the index and ring metacarpal shafts. Dissection then continues radial and ulnar to each of these bones and provides release of all the muscle compartments. Any dead muscle is debrided. Incisions are left open and covered with a nonadherent dressing. The wounds are reexplored in 2 to 3 days to assess for muscle viability. Often the incisions can be closed primarily, but a skin graft may be needed for the forearm.
Figure 44-16. Local flaps for digital tip coverage. A–C. For thumb injuries, Moberg described elevation of the entire volar skin with both neurovascular bundles for distal advancement. Sensation to the advanced skin is maintained. D–F. An 8-year-old girl underwent fingertip replantation that did not survive. A thenar flap was transferred to cover the defect. Some authors advise against its use in patients over 30 years old. G–I. In this 45-year-old man, the entire skin of P3 of the long finger was avulsed and unrecoverable. A cross-finger flap was transferred and provides excellent, durable coverage. The border of the flap and surrounding skin is still apparent 4.5 months after surgery.
Figure 44-16. (Continued)
Figure 44-16. (Continued)
If the examiner feels the patient does not have a compartment syndrome, elevation and serial examination are mandatory. When in doubt, it is safer to release an early compartment syndrome than to wait and risk muscle necrosis. Progression of compartment syndrome can lead to Volkmann’s ischemic contracture with muscle loss and scarring that may compress nerves and other critical structures. Medically, it is far easier to defend releasing an early compartment syndrome than delaying treatment until the process has progressed to necrosis and/or deeper scarring.

**COMPLICATIONS**

**Nonunion**

Any fractured bone has the risk of failing to heal. Fortunately, in the fingers and hand, this is a rare problem. Tuft injuries, where soft tissue interposes between the fracture fragments, can have relatively higher risk of this problem. The nonunited tuft can be treated with debridement and bone grafting or revision amputation depending on the needs and goals of the patient. Phalangeal and metacarpal nonunions are also quite rare. They can similarly be treated with debridement of the nonunion, grafting, and rigid fixation. More proximally, the scaphoid bone of the wrist has a significant risk of nonunion even if nondisplaced (see Fig. 44-9A). Any patient suspected of a scaphoid injury, namely those with tenderness at the anatomic snuffbox, should be placed in a thumb spica splint and reevaluated within 2 weeks even if initial X-rays show no fracture. Scaphoid nonunions can be quite challenging to repair, and immobilization at the time of injury in a thumb spica splint is essentially always warranted.

**Stiffness**

The desired outcome of any hand injury is a painless, mobile, functional hand. Multiple factors can contribute to decreased mobility, including complex injuries of soft tissue and bone, noncompliance of the patient with postoperative therapy, and inappropriate splinting. The surgeon performing the initial evaluation can greatly impact this last factor. The goal of splinting is to keep the collateral ligaments on tension (MPs at 90°, IP joints straight). For severe cases of stiffness, mobilization surgeries such as tenolysis and capsulotomies can be performed, but these rarely produce normal range of motion. Prevention of joint contractures with appropriate splinting and early, protected mobilization is the best option to maximize mobility at the end of healing. Healing of an injured or diseased structure in the hand is not the endpoint of treatment; the goal of any intervention must be to obtain structure healing, relief of pain, and maximization of function.

**Neuroma**

Any lacerated nerve will form a neuroma. A neuroma consists of a ball of scar and axon sprouts at the end of the injured nerve. In unfavorable circumstances, this neuroma can become painful. The SRN is particularly notorious for this problem. By providing proximal axon sprouts a target, nerve repair is an excellent preventive technique. In some circumstances, such as injuries requiring amputation, this is not possible. As mentioned earlier, the surgeon should resect the nerve stump as far proximally in the wound as possible to avoid the nerve stump healing in the cutaneous scar to minimize this risk.

For the patient who develops a painful neuroma, nonsurgical treatments are initiated first. The neuroma can be identified by the presence of a Tinel’s sign. Therapy techniques of desensitization, ultrasound, and electrical stimulation have all proven useful. Corticosteroid injection to the neuroma has also proven useful in some hands.

When these techniques fail, surgery is contemplated. The neuroma can be resected, but a new one will form to replace it. The nerve ending can be buried in muscle or even bone to prevent the neuroma from residing in a superficial location where it may be impacted frequently.

**Regional Pain Syndromes**

Injuries to the upper extremity can occasionally result in the patient experiencing pain beyond the area of initial injury. Reflex sympathetic dystrophy and sympathetic mediated pain are two terms that have been used in the past to describe this phenomenon. Both are inaccurate, as the sympathetic nervous system is not always involved. Current terminology for this condition is complex regional pain syndrome (CRPS). Type I occurs in the absence of a documented nerve injury; type II occurs in the presence of one. CRPSs manifest as pain beyond the area of initial injuries. There is often associated edema and changes in hair and/or sweat distribution. Comparison to the unaffected side is useful to better appreciate these findings. There are currently no imaging studies that can be considered diagnostic for CRPS.

For the patient in whom the diagnosis of CRPS is not clear, no definitive diagnostic study exists. Patients suspected of CRPS should be referred for aggressive hand therapy. Brief trials of oral corticosteroids have been successful in some series. Referral to a pain management specialist including a trial of stellate ganglion blocks is also frequently employed.

**NERVE COMPRESSION**

Nerves conduct signals along their axonal membranes toward their end organs. Sensory axons carry signals from distal to proximal; motor axons from proximal to distal. Myelin from Schwann cells allows faster conduction of signals. Signals jump from the start of one Schwann cell to the end of the cell (a location called a gap junction) and only require the slower membrane depolarization in these locations.

Nerve compression creates a mechanical disturbance of the nerve. In early disease, the conduction signal is slowed across the area of compression. When compression occurs to a sufficient degree for a sufficient time, individual axons may die. On a nerve conduction study, this manifests as a decrease in amplitude. Muscles receiving motor axons may show electrical disturbance on electromyogram (EMG) when sufficiently deprived of their axonal input.

Compression of sensory nerves typically produces a combination of numbness, paresthesias (pins and needles), and pain. Knowledge of the anatomic distribution of the peripheral nerves can aid in diagnosis. Sensory disturbance outside an area of distribution of a particular nerve (e.g., volar and dorsal radial-sided hand numbness for median nerve) makes compression of that nerve less likely. Diseases that cause systemic neuropathy (e.g., diabetes) can make diagnosis more difficult.

Nerve compression can theoretically occur anywhere along a peripheral nerve’s course. The most common sites of nerve compression in the upper extremity are the median nerve at the carpal tunnel, ulnar nerve at the cubital tunnel, and ulnar nerve at Guyon’s canal. Other, less common locations of nerve...
Carpal Tunnel Syndrome

The most common location of upper extremity nerve compression is the median nerve at the carpal tunnel, called carpal tunnel syndrome (CTS). The carpal tunnel is bordered by the scaphoid bone radially, the lunate and capitate bones dorsally, and the hook of the hamate bone ulnarly (see Fig. 44-3). The transverse carpal ligament, also called the flexor retinaculum, is its superficial border. The FPL, four FDS, and four FDP tendons pass through the carpal tunnel along with the median nerve. Of these 10 structures, the median nerve is relatively superficial and radial to the other nine.

An estimated 53 per 10,000 working adults have evidence of CTS. The National Institute for Occupational Safety and Health website asserts, “There is strong evidence of a positive association between exposure to a combination of risk factors (e.g., force and repetition, force and posture) and CTS.” There is disagreement among hand surgeons regarding whether occurrence of CTS in a patient who does repetitive activities at work represents a work-related injury.

Initial evaluation of the patient consists of symptom inventory: location and character of the symptoms, sleep disturbance due to symptoms, history of dropping objects, and difficulty manipulating small objects such as buttons, coins, or jewelry clasps.

Physical examination should begin with inspection. Look for evidence of wasting of the thenar muscles. Tinel’s sign should be tested over the median nerve from the volar wrist flexion crease to the proximal palm, although this test has significant interexaminer variability. Applying pressure over the carpal tunnel while flexing the wrist has been shown in one series to have the highest sensitivity when compared to Phalen’s and Tinel’s signs. Strength of the thumb in opposition should also be tested.

Early treatment of CTS consists of conservative management. The patient is given a splint to keep the wrist at 20° extension worn at nighttime. Many patients can have years of symptom relief with this management. As a treatment and diagnostic modality, corticosteroid injection of the carpal tunnel is often employed. Mixing local anesthetic into the solution provides the benefit of early symptom relief (corticosteroids often take 3–7 days to provide noticeable benefit), and report of postinjection anesthesia in the median nerve distribution confirms the injection went into the correct location. Multiple authors have shown a strong correlation to relief of symptoms with corticosteroid injection and good response to carpal tunnel release.

When lesser measures fail or are no longer effective, carpal tunnel release is indicated. Open carpal tunnel release is a time-tested procedure with documented long-term relief of symptoms. A direct incision is made over the carpal tunnel, typically in line with where the ring finger pad touches the proximal palm in flexion. Skin is divided followed by palmar fascia. The carpal tunnel contents are visualized as they exit the carpal tunnel. The transverse carpal ligament is divided with the median nerve visualized and protected at all times. Improvement in symptoms is typically noted by the first postoperative visit, although symptom relief may be incomplete for patients with long-standing disease or systemic nerve-affecting diseases such as diabetes.

Endoscopic techniques have been devised to address CTS. All involve avoidance of incising the skin directly over the carpal tunnel. In experienced hands, endoscopic carpal tunnel release provides the same relief of CTS with less intense and shorter lasting postoperative pain. After 3 months, however, the results are equivalent to open release. In inexperienced hands, there may be a higher risk of injury to the median nerve with the endoscopic techniques; this procedure is not for the occasional carpal tunnel surgeon.

Cubital Tunnel Syndrome

The second most common location of upper extremity nerve compression is the ulnar nerve where it passes behind the elbow at the cubital tunnel. The cubital tunnel retinaculum passes between the medial epicondyle of the humerus and the olecranon process of the ulna. It stabilizes the ulnar nerve in this location during elbow motion. Over time, or sometimes after trauma, the ulnar nerve can become less stabilized in this area. Motion of the elbow then produces trauma to the nerve as it impacts the retinaculum and medial epicondyle.

Cubital tunnel syndrome may produce sensory and motor symptoms. The small finger and ulnar half of the ring fingers may have numbness, paresthesias, and/or pain. The ulnar nerve also innervates the dorsal surface of the small finger and ulnar side of the ring finger, so numbness in these areas can be explained by cubital tunnel syndrome. The patient may also report weakness in grip due to effects on the FDP tendons to the ring and small fingers and the intrinsic hand muscles. Patients with advanced disease may complain of inability to fully extend the ring and small finger IP joints.

Physical examination for cubital tunnel syndrome begins with inspection. Look for wasting in the hypothenar eminence and the interdigital web spaces. When the hand rests flat on the table, the small finger may rest in abduction with respect to the other fingers; this is called Wartenberg’s sign. Tinel’s sign is often present at the cubital tunnel. Elbow flexion and the shoulder internal rotation tests are affective maneuvers to aid in the diagnosis of cubital tunnel syndrome. Grip strength and finger abduction strength should be compared to the unaffected side. Froment’s sign can be tested by placing a sheet of paper between the thumb and index finger and instructing the patient to hold on to the paper while the examiner pulls it away without flexing the finger or thumb (this tests the strength of the adductor pollicis and first dorsal interosseous muscles). If the patient must flex the index finger and/or thumb (FDP-index and FPL, both median nerve supplied) to maintain traction on the paper, this is a positive response.

Early treatment of cubital tunnel syndrome begins with avoiding maximal flexion of the elbow. Splints are often used for this purpose. Corticosteroid injection is rarely done for this condition; unlike in the carpal tunnel, there is very little space within the tunnel outside of the nerve. Injection in this area runs a risk of intraneural injection, which can cause permanent scarring of the nerve and dysfunction.

When conservative management fails, surgery has been contemplated. Treatment options include releasing the cubital tunnel retinaculum with or without transposing the nerve anterior to the elbow. While some authors advocate anterior transposition into the flexor-pronator muscle group with the goal of maximizing nerve recovery, recent studies have demonstrated equivalent results between transposition and in situ release of the nerve even in advanced cases. For this reason, the simpler in situ release, either open or endoscopic, is preferred by many surgeons.
Other Sites of Nerve Compression

All nerves crossing the forearm have areas described where compression can occur. The median nerve can be compressed as it passes under the pronator teres. The ulnar nerve can be compressed as it passes through Guyon’s canal. The radial nerve, or its posterior interosseous branch, can be compressed as it passes through the radial tunnel (distal to the elbow where the nerve divides and passes under the arch of the supinator muscle). The SRN can be compressed distally in the forearm as it emerges from under the brachioradialis tendon, called Wartenberg’s syndrome. As mentioned previously, any nerve can become compressed in scar at the site of a previous trauma.

DEGENERATIVE JOINT DISEASE

As with other joints in the body, the joints of the hand and wrist can develop degenerative changes. Symptoms typically begin in the fifth decade of life. Symptoms consist of joint pain and stiffness and often are exacerbated with changes in the weather. Any of the joints can become involved. As the articular cartilage wears out, pain typically increases and range of motion decreases. The patient should always be asked to what degree symptoms are impeding activities.

Physical findings are documented in serial fashion from the initial visit and subsequent visits. Pain with axial loading of the joint may be present. Decreased range of motion may be a late finding. Instability of the collateral ligaments of the joint is uncommon in the absence of inflammatory arthritis.

Plain X-rays are typically sufficient to demonstrate arthritis. Initially, the affected joint has a narrower radiolucent space between the bones. As joint degeneration progresses, the joint space further collapses. Bone spurs, loose bodies, and cystic changes in the bone adjacent to the joint all may become apparent. X-ray findings do not always correlate with patient symptoms. Patients with advanced X-ray findings may have minimal symptoms, and vice versa. Treatment is initiated and progressed based on the patient’s symptoms regardless of imaging findings.

Initial management begins with rest of the painful joint. Splints are often useful, but may significantly impair the patient in activities and thus are frequently used at nighttime only. Oral nonsteroidal anti-inflammatory medications such as ibuprofen and naproxen are also useful. Patients on anticoagulants and antiplatelet medications may not be able to take these, and some patients simply do not tolerate the gastric irritation side effect even if they take the medication with food.

For patients with localized disease affecting only one or a few joints, corticosteroid injection may be contemplated. Needle insertion can be difficult since these joint spaces are quite narrow even before degenerative disease sets in. Also, many corticosteroid injections are suspensions, not solutions; injected corticosteroid will remain in the joint space and can be seen as a white paste if surgery is performed on a joint that has been previously injected.

Small Joints (Metacarpophalangeal and Interphalangeal)

When conservative measures fail, two principal surgical options exist: arthrodesis and arthroplasty. The surgeon and patient must decide together as to whether conservative measures have failed. Surgery for arthritis, whether arthrodesis or arthroplasty, is performed for the purpose of relieving pain. Arthrodesis, fusion of a joint can be performed with a tension band or axial compression screw techniques. Both methods provide excellent relief of pain and is durable over time. However, it comes at the price of total loss of motion.

Silicone implant arthroplasty has been available for over 40 years. Rather than a true replacement of the joint, the silicone implant acts as a spacer between the two bones adjacent to the joint. This allows for motion without bony contact that would produce pain. Long-term studies have shown that all implants fracture over time, but usually continue to preserve motion and pain relief.

In the past 15 years, resurfacing implant arthroplasties have become available for the small joints of the hand. Multiple different materials have been used to fabricate such implants. These are designed to behave as a true joint resurfacing (as knee and hip arthroplasty implants are) and have shown promising outcomes in short- and intermediate-term studies. Neither the silicone nor the resurfacing arthroplasties preserve (or restore) full motion of the MP or PIP joints.

Wrist

The CMC joint of the thumb, also called the basilar joint, is another common location of arthritis pain. Pain in this joint particularly disturbs function because the CMC joint is essential for opposition and cylindrical grasp. Patients will typically complain of pain with opening a tight jar or doorknob and strong pinch activities such as knitting. Conservative management is used first, as described earlier. Prefabricated, removable thumb spica splinting can provide excellent relief of symptoms for many patients.

Multiple surgical options exist for thumb CMC arthritis. Many resurfacing implants have been used in the past; often they have shown good short- and intermediate-term results and poor long-term results. Resection of the arthritic trapezium provides excellent relief of pain; however, many authors feel that stabilization of the thumb metacarpal base is necessary to prevent shortening and instability. Some surgeons have demonstrated excellent long-term results from resection of the trapezium without permanent stabilization of the metacarpal base. For both of these operations, the thumb base may not be sufficiently stable to withstand heavy labor. For these patients, fusion of the thumb CMC in mild opposition provides excellent pain relief and durability. The patient must be warned preoperatively that he will not be able to lay his hand flat after the surgery. This loss of motion can be problematic when the patient attempts to tuck in clothing or reach into a narrow space.

Degenerative change of the radiocarpal and midcarpal joints is often a consequence of scapholunate ligament injury. Often the initial injury goes untreated, with the patient believing it is merely a “sprain”; the patient is first diagnosed with the initial injury when he presents years later with degenerative changes.

Degenerative wrist changes associated with the scapholunate ligament follow a predictable pattern over many years, called scapholunate advanced collapse or SLAC wrist. Because of this slow progression (Fig. 44-17A), patients can usually be treated with a motion-sparing procedure. If there is truly no arthritic change present, the scapholunate ligament can be reconstructed.

If arthritis is limited to the radiocarpal joint, two motion-sparing options are available. The proximal carpal row (scaphoid, lunate, and triquetrum) can be removed (proximal row carpectomy [PRC]). The lunate facet of the radius then
articulates with the base of the capitate, whose articular surface is similar in shape to that of the base of the lunate. Studies have shown maintenance of approximately 68% of the wrist flexion-extension arc and 72% of hand strength compared to the contralateral side.61 Alternatively, the scaphoid can be excised, and four-bone fusion (lunate, capitate, hamate, and triquetrum) can be performed. This maintains the full length of the wrist and the lunate in the lunate facet of the radius. Some series have shown better strength but less mobility with this technique, others have shown equivalent results to the PRC.62 The four-bone fusion does appear to be more durable for younger patients and/or those who perform heavy labor.

If the patient presents with pancarpal arthritis or motion-sparing measures have failed to alleviate pain, total wrist fusion is the final surgical option. The distal radius is fused, through the proximal and distal carpal rows to the third metacarpal, typically with a dorsal plate and screws. Multiple long-term studies have shown excellent pain relief and durability; this comes at the cost of total loss of wrist motion. This is surprisingly well tolerated in most patients, especially if the other hand/wrist is unaffected. The only activity of daily living that cannot be done with a fused wrist is personal toileting.

**Rheumatoid Arthritis**

Rheumatoid arthritis (RA) is an inflammatory arthritis that can affect any joint in the body. Inflamed synovium causes articular cartilage breakdown with pain and decreased range of motion. The goals of hand surgery for the RA patient are relief of pain, improvement of function, slowing progression of disease, and improvement in appearance.63 In addition, swelling of the joint due to the inflammation can cause laxity and even failure of the collateral ligaments supporting the joints. Recent advances in the medical care of RA have made the need for surgical care of these patients far less common than in previous decades.

MP joints of the fingers are commonly affected. The base of the proximal phalanx progressively subluxates and eventually dislocates volarly with respect to the metacarpal head. The collateral ligaments, particularly on the radial side, stretch out and cause the ulnar deviation of the fingers characteristic of the rheumatoid hand. In more advanced cases, the joint may not be salvageable (Fig. 44-17B). For these patients, implant arthroplasty is the mainstay of surgical treatment. Silicone implants have been used for over 40 years with good results.64 The silicone implant acts as a spacer between proximal and distal bone, rather than as a true resurfacing arthroplasty. The radial collateral ligament must be repaired to appropriate length to correct the preoperative ulnar deviation of the MP joint. Extensor tendon centralization is then performed, as needed, at the end of the procedure.

For MP joint and PIP joint disease, fusion is an option. However, since RA usually affects multiple joints, fusion is typically avoided due to impaired function of adjacent joints, which would leave a severe motion deficit to the finger.

Failure of the support ligaments of the distal radioulnar joint (DRUJ) leads to the caput ulnae posture of the wrist with the ulnar head prominent dorsally. As this dorsal prominence becomes more advanced, the ulna head, denuded of its cartilage to act as a buffer, erodes into the overlying extensor tendons. Extensor tenosynovitis, followed ultimately by tendon rupture, begins ulnarly and proceeds radially. Rupture of the ECU tendon may go unnoticed due to the intact ECRL and ECRB tendons to extend the wrist. EDQ rupture may go unnoticed if a sufficiently robust EDC tendon to the small finger exists. Once the fourth compartment (EDC) tendons begin to fail, the motion deficit is unable to be ignored by the patient.

Surgical solutions must address the tendon ruptures as well as the DRUJ synovitis and instability and ulna head breakdown that led to them.65 Excision of the ulna head removes the bony prominence. The DRUJ synovitis must also be resected.
Alternatively, the DRUJ can be fused and the ulna neck resected to create a pseudoarthrosis to allow for rotation. For both procedures, the remaining distal ulna must be stabilized. Multiple techniques have been described using portions of FCU, ECU, wrist capsule, and combinations thereof.

The ruptured extensor tendons are typically degenerated over a significant length. Primary repair is almost never possible, and the frequent occurrence of multiple tendon ruptures makes repair with graft less desirable due to the need for multiple graft donors.

Strict compliance with postoperative therapy is essential to maximizing the surgical result. Due to the chronic inflammation associated with RA, tendon and ligament repairs will be slower to achieve maximal tensile strength. Prolonged nighttime splinting, usually for months, helps prevent recurrence of extensor lag. Finally, the disease may progress over time. Reconstructions that were initially adequate may stretch out or fail over time. Medical management is the key to slowing disease progression and maximizing the durability of any surgical reconstruction.

DUPUYTREN’S CONTRACTURE

In 1614, a Swiss surgeon named Felix Plater first described contraction of multiple fingers due to palpable, cord-like structures on the volar surface of the hand and fingers. The disease state described would ultimately come to be known as Dupuytren’s contraction. Dupuytren’s name came to be associated with the disease after he performed an open fasciotomy of a contracted cord before a class of medical students in 1831.

The palmar fascia consists of collagen bundles in the palm and fingers. These are primarily longitudinally oriented and reside as a layer between the overlying skin and the underlying tendons and neurovascular structures. There are also connections from this layer to the deep structures below and the skin above. Much is known about the progression of these structures from their normal state (called bands) to their contracted state (called cords), but little is known on how or why this process begins.

Increased collagen deposition leads to a palpable nodule in the palm. Over time, there is increased deposition distally into the fingers. This collagen becomes organized and linearly oriented. These collagen bundles, with the aid of myofibroblasts, contract down to form the cords, which are the hallmark of the symptomatic patient. Detail of the molecular and cell biology of Dupuytren’s disease is beyond the scope of this chapter but is available in multiple hand surgery texts.

Most nonoperative management techniques will not delay the progression of disease. Corticosteroid injections may soften nodules and decrease the discomfort associated with them but are ineffective against cords. Splinting has similarly been shown not to retard disease progression.

Recently, several minimally invasive treatment approaches have been described for the treatment of Dupuytren’s disease. Disruption of the cord with a needle is an effective means of releasing contractures, particularly at the MP joint level. Longer-term studies have demonstrated more rapid recovery from needle fasciotomy, as the procedure is called, but more durable results with fasciectomy. Injectable clostridial collagenase was approved by the U.S. Food and Drug Administration in 2009, and although it has shown good early results, treatment costs remain high.

For patients with advanced disease including contractures of the digits that limit function, surgery is the mainstay of therapy. Although rate of progression should weigh heavily in the decision of whether or not to perform surgery, general guidelines are MP contractures greater than or equal to 30° and/or PIP contractures greater than or equal to 20°.

Surgery consists of an open approach through the skin down to the involved cords. Skin is elevated off of the underlying cords. Great care must be taken to preserve as much of the subdermal vascular plexus with the elevated skin flaps to minimize postoperative skin necrosis. All nerves, tendons, and blood vessels in the operative field should be identified. Once this is done, the involved cord is resected while keeping the critical deeper structures under direct vision. The skin is then closed, with local flap transpositions as needed, to allow for full extension of the fingers that have been released (Fig. 44-18).

Alternative cord resection techniques include removal of the skin over the contracture (dermatofasciectomy). This requires a skin graft to the wound and should only be done if skin cannot be separated from the cords and local tissue cannot be rearranged with local flaps to provide closure of the wound.

Complications of surgical treatment of Dupuytren’s disease occur in as many as 24% of cases. Problems include digital nerve laceration, digital artery laceration, buttonholing of the skin, hematoma, swelling, and pain, including some patients with CRPS (see earlier section on CRPS). Digital nerve injury can be quite devastating, producing annoying numbness at best or a painful neuroma in worse situations.

Hand therapy is typically instituted within a week of surgery to begin mobilization of the fingers and edema control. The therapist can also identify any early wound problems because he or she will see the patient more frequently than the surgeon. Extension hand splinting is maintained for 4 to 6 weeks, with nighttime splinting continued for an additional 6 to 8 weeks. After this point, the patient is serially followed for evidence of recurrence or extension of disease.

INFECTIONS

Trauma is the most common cause of hand infections. Other predisposing factors include diabetes, neuropathies, and immunocompromised patients. Proper treatment consists of incision and drainage of any collections followed by debridement, obtaining wound cultures, antibiotic therapy, elevation, and immobilization. Staphylococcus and Streptococcus are the offending pathogens in about 90% of hand infections. Infections caused by intravenous drug use or human bites and those associated with diabetes will often be polymicrobial, including gram-positive and gram-negative species. Heavily contaminated injuries require anaerobic coverage. Although α-hemolytic Streptococcus and Staphylococcus aureus are the most commonly encountered pathogens in human bites, Eikenella corrodens is isolated in up to one-third of cases and should be considered when choosing antimicrobial therapy. Ziehl-Neelsen staining and cultures at 28°C to 32°C in Lowenstein-Jensen medium must be performed if there is a suspicion for atypical mycobacteria.

Cellulitis

Cellulitis is characterized by a nonpurulent diffuse spreading of inflammation characterized by erythema, warmth, pain, swelling, and induration. Skin breakdown is a frequent cause, but
often no inciting factor is identified. Group A \(\alpha\)-hemolytic *Streptococcus* is the most common offending pathogen and causes a more diffuse spread of infection. *S. aureus* is the second most common offending pathogen and will cause a more localized cellulitis. The diagnosis of cellulitis is clinical. Septic arthritis, osteomyelitis, an abscess, a deep-space infection, and necrotizing fasciitis are severe infectious processes that may initially mimic cellulitis. These must be ruled out appropriately before initiating treatment, and serial exams should be conducted to ensure proper diagnosis. Treatment of cellulitis consists of elevation, splint immobilization, and antibiotics that cover both *Streptococcus* and *Staphylococcus*. Intravenous antibiotics are usually initiated for patients with severe comorbidities and those who fail to improve on oral antibiotics after 24 to 48 hours. Failure to improve after 24 hours indicates a need to search for an underlying abscess or other infectious cause.73

**Abscess**

An abscess will present much like cellulitis, but they are two clinically separate entities. The defining difference is an area of fluctuance. Skin-puncturing trauma is the most common cause. *S. aureus* is the most common pathogen, followed by *Streptococcus*. Treatment consists of incision and drainage with appropriate debridement, wound cultures, wound packing, elevation, immobilization, and antibiotics. The packing should be removed in 12 to 24 hours or sooner if there is clinical concern, and warm soapy water soaks with fresh packing should be initiated. Most should be allowed to heal secondarily. Delayed primary closure should only be performed after repeat washouts for larger wounds where complete infection control has been achieved.

**Collar-Button Abscess**

This is a subfascial infection of a web space and is usually caused by skin trauma that becomes infected; it often occurs in

Figure 44-18. Dupuytren’s disease. A. This patient has cords affecting the thumb, middle, ring, and small fingers. B. The resected specimens are shown. C. Postoperatively, the patient went on to heal all his incisions and, with the aid of weeks of hand therapy, recover full motion.
laborers. The adherence of the palmar web space skin to the palmar fascia prevents lateral spread, so the infection courses dorsally, resulting in both palmar web space tenderness and dorsal web space swelling and tenderness. The adjacent fingers will be held in abduction with pain on adduction (Fig. 44-19). Incision and drainage, often using separate volar and dorsal incisions, is mandatory, and follows the same treatment as for any abscess or deep-space infection.

**Osteomyelitis**

Osteomyelitis in the hand usually occurs due to an open fracture with significant soft tissue injury. The presence of infected hardware, peripheral vascular disease, diabetes, and alcohol or drug abuse are also predisposing factors. Presentation includes persistent or recurrent swelling with pain, erythema, and possible drainage. It will take 2 to 3 weeks for periosteal reaction and osteopenia to be detected on radiographs. Bone scans and MRI are useful modalities to aid in diagnosis. Erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP) have low specificity but are useful for monitoring the progress of treatment, with CRP being more reliable. Treatment consists of antibiotics alone in the early stage as long as there is favorable response. All necrotic bone and soft tissue, if present, must be debrided. Initial intravenous antibiotic therapy should cover *S. aureus*, the most common pathogen, and should then be adjusted according to bone cultures. Antibiotic therapy is continued for 4 to 6 weeks once the patient clinically improves and there is no further need for debridement. For osteomyelitis in the setting of an acute fracture with internal fixation in place, the hardware should be left in place as long as it is stable and the fracture has not yet healed. If the hardware is unstable, it must be replaced. An external fixation device may be useful in this setting. If osteomyelitis occurs in a healed fracture, all hardware and necrotic bone and soft tissue must be removed.74

**Pyogenic Arthritis**

Infection of a joint will progress quickly to severe cartilage and bony destruction if not addressed quickly. Direct trauma and local spread of an infection are the most common causes. Hematogenous spread occurs most commonly in patients who are immunocompromised. *S. aureus* is the most common pathogen, followed by *Streptococcus* species. *Neisseria gonorrhoeae* is the most common cause of atraumatic septic arthritis in an adult less than 30 years of age. Presentation includes exacerbation of pain with any joint movement, severe pain on axial load, swelling, erythema, and tenderness. Radiographs may show a foreign body or fracture, with widened joint space early in the process and decreased joint space late in the process due to destruction. Joint aspiration with cell count, Gram stain, and culture is used to secure the diagnosis. Treatment of nongonococcal septic arthritis includes open arthrotomy, irrigation, debridement, and packing the joint or leaving a drain in place. Intravenous antibiotics are continued until there is clinical improvement, followed by 2 to 4 weeks of additional oral or intravenous antibiotics. Gonococcal septic arthritis is usually treated nonoperatively. Intravenous ceftriaxone is first-line therapy. Joint aspiration may be used to obtain cultures and decrease joint pressure.75

**Necrotizing Infections**

Necrotizing soft tissue infections occur when the immune system is unable to contain an infection, leading to extensive spread with death of all involved tissues. This is different from an abscess, which forms when a functioning immune system is able to “wall off” the infectious focus. Necrotizing infections can result in loss of limb or life, even with prompt medical care. Bacteria spread along the fascial layer, resulting in the death of soft tissues, which is in part due to the extensive blood vessel thrombosis that occurs. An inciting event is not always identified. Immunocompromised patients and those who abuse drugs or alcohol are at greater risk, with intravenous drug users having the highest increased risk. The infection can be monomicrobial, with group A β-hemolytic *Streptococcus* being the most common pathogen, followed by α-hemolytic *Streptococcus*, *S. aureus*, and anaerobes. Prompt clinical diagnosis and treatment are the most important factors for salvaging limbs and saving life. Patients will present with pain out of proportion with findings. Appearance of skin may range from normal to erythematous or maroon with edema, induration, and blistering. Crepitus may occur if a gas-forming organism...
"Dirty dishwater fluid" may be encountered as a scant grayish fluid, but often there is little to no discharge. There may be no appreciable leukocytosis. The infection can progress rapidly and can lead to septic shock and disseminated intravascular coagulation. Radiographs may reveal gas formation, but they must not delay emergent debridement once the diagnosis is suspected. Intravenous antibiotics should be started immediately to cover gram-positive, gram-negative, and anaerobic bacteria. Patients will require multiple debridements, and the spread of infection is normally wider than expected based on initial assessment.\(^7\)

Necrotizing myositis, or myonecrosis, is usually caused by Clostridium perfringens due to heavily contaminated wounds. Unlike necrotizing fasciitis, muscle is universally involved and found to be necrotic. Treatment includes emergent debridement of all necrotic tissue along with empirical intravenous antibiotics.

Wet gangrene is most common in diabetics with renal failure and an arteriovenous shunt. It is usually polymicrobial. Patients will present with a necrotic digit that is purulent and very malodorous, with rapidly evolving pain, swelling, skin discoloration, and systemic collapse. Emergent treatment is the same as for other necrotizing infections, and amputation of the involved digit or extremity must often be performed.

**Infectious Flexor Tenosynovitis**

Flexor tenosynovitis (FTS) is a severe pathophysiologic state causing disruption of normal flexor tendon function in the hand. A variety of etiologies are responsible for this process. Most acute cases of FTS are due to purulent infection. FTS also can occur secondary to chronic inflammation as a result of diabetes, RA, crystalline deposition, overuse syndromes, amyloidosis, psoriatic arthritis, systemic lupus erythematosus, and sarcoidosis.

The primary mechanism of infectious FTS usually is penetrating trauma. Most infections are caused by skin flora, including both *Staphylococcus* and *Streptococcus* species. Bacteria involved vary by etiology of the infection: bite wounds (*Pasteurella multocida*—cat, *E. corrodens*—human); diabetic patients (*Bacteroides, Fusobacterium, Haemophilus* species, gram-negative organisms); hematogenous spread (*Mycobacterium tuberculosis, N. gonorrhoeae*); or water-related punctures (*Vibrio vulnificus, Mycobacterium marinum*). Infection in any of the fingers may spread proximally into the wrist, carpal tunnel, and forearm, also known as Parona’s space.\(^7\)

Suppurative FTS has the ability to rapidly destroy a finger’s functional capacity and is considered a surgical emergency. Suppurative FTS results from bacteria multiplying in the closed space of the flexor tendon sheath and culture-rich synovial fluid medium causing migration of inflammatory cells and subsequent swelling. The inflammatory reaction within the closed tendon sheath quickly erodes the paratenon, leading to adhesions and scarring, as well as increase in pressures within the tendon sheath that may lead to ischemia. The ultimate consequences are tendon necrosis, disruption of the tendon sheath, and digital contracture.

Patients with infectious FTS present with pain, redness, and fever (Fig. 44-20). Physical examination reveals Kanavel’s “cardinal” signs of flexor tendon sheath infection: finger held in slight flexion, fusiform swelling, tenderness along the flexor tendon sheath, and pain over the flexor sheath with passive extension of the digit.\(^7\) Kanavel’s signs may be absent in patients who are immunocompromised, have early manifestations of
infection, have recently received antibiotics, or have a chronic, indolent infection.

If a patient presents with suspected infectious FTS, empiric intravenous antibiotics should be initiated. Prompt medical therapy in early cases may prevent the need for surgical drainage. For healthy individuals, empiric antibiotic therapy should cover *Staphylococcus* and *Streptococcus*. For immunocompromised patients (including diabetics) or infections associated with bite wounds, empiric treatment should include coverage of gram-negative organisms as well.78

Adjuncts to antibiotics include splint immobilization (intrinsic plus position preferred) and elevation until infection is under control. Hand rehabilitation (i.e., range-of-motion exercises and edema control) should be initiated once pain and inflammation are under control.

If medical treatment alone is attempted, then initial inpatient observation is indicated. Surgical intervention is necessary if no obvious improvement has occurred within 12 to 24 hours.

Several surgical approaches can be used to drain infectious FTS. The method used is based on the extent of the infection. Michon developed a classification scheme that can be useful in guiding surgical treatment (Table 44-1).79 Figure 44-20 (B and C) demonstrates drainage of a stage II FTS. A Brunner incision allows better initial exposure but may yield difficulties with tendon coverage if skin necrosis occurs. A 16-gauge catheter or 5-French pediatric feeding tube then is inserted into the tendon sheath through the proximal incision. The sheath is copiously irrigated with normal saline. Avoid excessive fluid extravasation into the soft tissue because the resulting increase in tissue pressure can lead to necrosis of the digit. The catheter is removed after irrigation. The incisions are left open. Some surgeons prefer a continuous irrigation technique for a period of 24 to 48 hours. The catheter is sewn in place, and a small drain is placed at the distal incision site. Continuous or intermittent irrigation every 2 to 4 hours with sterile saline can then be performed through the indwelling catheter.

After surgery, an intrinsic plus splint is applied, the hand is elevated, and the appropriate empiric antibiotic coverage is instituted while awaiting culture results. The hand is reexamined the following day. Whirlpool therapy and range of motion are begun. Drains are removed before discharge from the hospital. The wounds are left open to heal by secondary intention. In severe cases, repeat irrigation and operative debridement may be required.

Antibiotic therapy is guided by culture results as well as clinical improvement. Once there is no further need for debridement, a 7- to 14-day course of oral antibiotics is generally prescribed. Consultation with an infectious disease specialist should be considered early in order to maximize efficiency and efficacy of therapy.

**Felon**

A felon is a subcutaneous abscess of the fingertip and is most commonly caused by penetrating trauma. *S. aureus* is the most common pathogen. The fingertip contains multiple septa connecting the distal phalanx to the skin. These septa are poorly compliant, and presence of an abscess will increase pressure and lead to severe pain and tissue death. Patients will experience erythema, swelling, and tenderness of the volar digital pad. Oral antibiotics may resolve the infection if diagnosed very early, but incision and drainage is indicated when fluctuance is identified. A digital block should be performed, followed by a longitudinal incision over the point of maximal fluctuance (Fig. 44-21). Transverse and lateral incisions should be avoided, and the incision should never extend across the distal phalangeal joint crease. Deep incision should not be performed as this may cause seeding of bacteria into the flexor tendon sheath. The wound is irrigated and packed, with warm soapy water soaks and packing changes initiated within 24 hours and performed two to three times daily until secondarily healed. Antibiotic coverage should cover for *Staphylococcus* and *Streptococcus* species.73

**Paronychia**

Paronychia is an infection beneath the nail fold. The nail plate can be viewed as an invagination into the dorsal skin extending down to the distal phalanx periosteum. Predisposing factors include anything that causes nail trauma, such as manicures, artificial nails, or nail biting. The infection may spread around

---

### Table 44-1

**Michon’s stages of suppurative flexor tenosynovitis and appropriate treatment**

<table>
<thead>
<tr>
<th>STAGE</th>
<th>FINDINGS</th>
<th>TREATMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Increased fluid in sheath, mainly a serous exudate</td>
<td>Catheter irrigation</td>
</tr>
<tr>
<td>II</td>
<td>Purulent fluid, granulomatous synovium</td>
<td>Minimal invasive drainage ± indwelling catheter irrigation</td>
</tr>
<tr>
<td>III</td>
<td>Necrosis of the tendon, pulleys, or tendon sheath</td>
<td>Extensive open debridement and possible amputation</td>
</tr>
</tbody>
</table>

---

*Figure 44-21.* Felon. **A.** Lateral view of the digit showing fluctuance between the skin of the pad and the underlying distal phalanx bone. **B.** The authors prefer to drain felons with a longitudinal incision (dashed line) directly over the area of maximal fluctuance.
the nail plate from one side to the other, or it may extend into the pulp and result in a felon. An acute paronychia is usually caused by *S. aureus* or *Streptococcal* species. Patients report pain, erythema, swelling, and possibly purulent drainage involving the periungual tissue. Treatment consists of warm water soaks and oral antibiotics if diagnosed early. If purulence or fluctuance is present, then a freer elevator or 18-gauge needle can be passed along the involved nail fold to decompress the collection (Fig. 44-22). If the infection involves the eponychial fold, a small proximally based flap of eponychium is created by using a scalpel, followed by irrigation and packing. The nail plate must be removed if the infection extends beneath the nail plate. Packing is kept in place for 24 to 48 hours, followed by warm water soaks and local wound care. Usually, the wound cannot be repacked once the dressing is removed.\(^{73}\)

A chronic paronychia is most commonly caused by *Candida* species and is most often found in patients who perform jobs involving the submersion of their hands in water or other moist environments. These develop into thickened nails with callus-like formation along the nail folds and may occasionally become red and inflamed. They do not respond to antibiotic treatment, and nail plate removal with marsupialization of the skin proximal to the eponychial fold will allow the wound to heal secondarily. The environmental factors leading to the chronic paronychia must also be corrected in order for treatment to be successful.

All hand infections other than cellulitis will require surgical management. Clinical examination, particularly noting the area of greatest tenderness and/or inflammation, is the single most useful diagnostic tool to localize any purulence requiring drainage. Specific recommendations for differentiating among the possible locations of hand infection are included in the diagnostic algorithm shown in Fig. 44-23.

TUMORS

Tumors of the hand and upper extremity can be classified as benign soft tissue tumors; malignant soft tissue tumors (subclassified into cutaneous and noncutaneous malignancies); benign bony tumors; malignant bony tumors; and secondary metastatic tumors. Initial investigation for any mass starts with a complete

---

**Figure 44-22.** Paronychia. A. Fluctuance in the nail fold is the hallmark of this infection. B. The authors prefer to drain a paronychia using the bevel of an 18-gauge needle inserted between the nail fold and the nail plate at the location of maximal fluctuance.

**Figure 44-23.** Diagnostic algorithm. Diagnostic workup for a patient with hand inflammation to evaluate for infection. See text for details about particular infectious diagnoses. Abx = antibiotics; FTS = flexor tenosynovitis; IF MC = index finger metacarpal; MRI = magnetic resonance imaging; SF MC = small finger metacarpal.
history and physical exam. Hand and/or wrist X-rays should be obtained in every patient presenting with a mass unless clearly not indicated (e.g., a superficial skin lesion with no aggressive/malignant features). The workup proceeds in an orderly fashion until a diagnosis is obtained. Once a benign diagnosis is secured (by strong clinical suspicion in an experienced hand surgeon, radiographic evidence, or tissue biopsy), further workup is not needed; this may occur at any point in the workup of a mass.

Most hand masses are benign and can be readily diagnosed without advanced imaging or tissue biopsy. When necessary, additional workup may include baseline laboratory studies, CT and/or MRI of the involved region, and a bone scan or positron emission tomography (PET) scan. Staging of a malignant tumor may occur before biopsy if a malignancy is strongly suspected, or it may occur after formal biopsy. Staging includes a chest X-ray and CT with intravenous contrast of the chest, abdomen, and pelvis to detect possible metastasis. Biopsy of the mass is always the last step of a workup and should occur only after all other available information has been gathered. Any mass that is over 5 cm in size, is rapidly increasing in size (as judged by an experienced surgeon or oncologist), is symptomatic or painful, or has an aggressive clinical or radiographic appearance warrants workup and biopsy to rule out malignancy.

CT scans are useful for detecting bony tumor extension across planes and identifying tumors of small bones, such as the carpal bones. MRI is useful for evaluating soft tissue tumor involvement (e.g., which muscle compartments are involved) as well as intramedullary lesions. Most soft tissue tumors will appear dark on T1-weighted images and bright on T2-weighted images. Hematomas, hemangiomas, lipomas, liposarcomas, and adipose tissue will appear bright on T1-weighted images and dark on T2-weighted images. Scintigraphy uses methylene diphosphonate attached to technetium-99m. This complex will attach to hydroxyapatite. Immediate uptake is seen in areas of increased vascularity, such as infection, trauma, and neoplasia. Increased uptake 2 to 3 hours later is seen in “pooled” areas where new bone formation has occurred. This modality is useful for detecting areas of tumor invasion or metastases not otherwise seen on prior CT, MRI, or radiographs.

Biopsy is reserved for masses that cannot be diagnosed as benign based on prior clinical and radiographic exams. Needle biopsy is not reliable for primary diagnosis, but it can be useful for recurrent or metastatic disease. Open excisional (if mass is less than 5 cm in size) or incisional (if mass is greater than 5 cm in size) biopsy is the most common biopsy method. Proper surgical oncologic technique is strictly adhered to in order to prevent tumor spread into uninvolved tissues or compartments. This includes making all incisions longitudinally using sharp dissection and meticulous hemostasis; carrying the incision directly down to the tumor with no development of tissue planes (i.e., making a straight-line path from skin to tumor); incising through the fewest number of muscle compartments; and avoiding critical neurovascular structures. The CT or MRI images will help determine the best surgical approach for biopsy or resection in order to avoid uninvolved compartments and critical structures.

**Benign Soft Tissue Tumors**

**Ganglion Cyst.** This is the most common soft tissue tumor of the hand and wrist, comprising 50% to 70% of all soft tissue tumors in this region. They can occur at any age but are most common in the second to fourth decades with a slight predilection toward females. Patients may report a slow-growing soft mass that may fluctuate in size and can sometimes be associated with mild pain. Compressive neuropathies may be seen if they occur in Guyon’s canal or the carpal tunnel, but they are uncommon. There are no reports of malignant degeneration. History and physical exam are usually sufficient to establish a diagnosis. Occurrence by location is as follows: 60% to 70% occur on the dorsal wrist between the third and fourth extensor compartments and are connected by a stalk to the scapholunate ligament (Fig. 44-24); 18% to 20% occur on the volar wrist; and 10% to 12% occur in the digits as volar retinacular or flexor tendon sheath cysts. The cyst transilluminates. There is always a stalk that communicates with the underlying joint or tendon sheath. The cyst wall is composed of compressed collagen fibers with no epithelial or synovial cells present. Clear viscous mucin fills the cyst and is composed of glucosamine, albumin, globulin, and hyaluronic acid. The etiology is unclear. The most accepted theory currently is Angelides’ who proposed that repeated stress of a joint, ligament, or tendon sheath causes an increase of mucin-producing cells and subsequent mucin production. The increased mucin production dissects superficially and coalesces into a cyst. The successful treatment of dorsal ganglion cysts by excising only the stalk supports this theory.

Treatment consists of observation if asymptomatic. If symptoms exist or the patient desires removal for cosmetic appearance, aspiration of the cyst may be performed with a
specific cure rate ranging from 15% to 89%. The benefit of injected steroids is inconclusive. Aspiration of a volar wrist ganglion cyst can be dangerous due to the potential of injuring neurovascular structures. Open excision and arthroscopic excision of the cyst stalk are surgical options for cysts that are not amendable to aspiration. A recent meta-analysis reported recurrence rates after either needle aspiration, open excision, and arthroscopic excision as 59%, 21%, and 6%, respectively.

**Mucous Cyst.** A mucous cyst is a ganglion cyst of the DIP joint. They occur most commonly in the fifth to seventh decades, and the underlying cause is associated osteoarthritis of the DIP joint. They are slow growing and usually occur on one side of the terminal extensor tendon between the DIP joint and the eponychium. The earliest clinical sign is often longitudinal grooving of the involved nail plate followed by a small enlarging mass and then attenuation of overlying skin. X-rays will show signs of osteoarthritis within the DIP joint. Heberden nodes (osteophytes within the DIP joint) are often seen on X-ray.

Possible treatment includes observation, aspiration, or excision. If the cyst is not draining and the overlying skin is intact, the patient may be offered reassurance. A draining cyst poses risk of DIP joint infection due to the tract communicating with the DIP joint and should be excised. If the cyst is symptomatic, painful, or the patient desires removal for cosmetic purposes, excision should be performed. Any osteophytes in the DIP joint must be removed to reduce recurrence. Aspiration is an option for treatment, but this poses the risk of DIP joint infection through seeding of bacteria into the joint or by the development of a draining sinus tract. It is generally not performed.

**Giant Cell Tumor of the Tendon Sheath.** Also known as a xanthosarcoma, fibrous xanthoma, localized nodular synovitis, sclerosing hemangioma, or pigmented villonodular tenosynovitis, giant cell tumor of the tendon sheath is the second most common soft tissue mass of the hand and wrist. It is a benign lesion with no clear pathogenesis. The tumor is a growth of polyclonal cells with no risk of malignant transformation. Despite the similarity in name, it is not histopathologically related to giant cell tumor of the bone.

Giant cell tumor of the tendon sheath occurs as a firm slow-growing painless mass over months to years and will often feel bumpy or nodular, which is a distinguishing characteristic helpful for diagnosis. It has a predilection for occurring in close proximity to joints along flexor surfaces of the wrist, hands, and digits (especially the PIP joints of the radial digits) and occurs most commonly between the second and fifth decades (Fig. 44-25A). These tumors do not transilluminate. Direct extension into joints and ligaments can make complete excision difficult. Gross appearance of the tumor will show a well-circumscribed nodular firm mass with a deep brown color due to the large amount of hemosiderin content, which is easily detected on histologic staining (Fig. 44-25B). Multinucleated giant cells and hemosiderin-laden macrophages are characteristic.

This tumor is not visible on radiographs. Approximately 20% will show extrinsic cortical erosion on X-ray. This is a risk factor for recurrence, and removal of the cortical shell should be considered. MRI is useful for delineating involvement with tendons, ligaments, and joints.

The standard treatment is marginal excision. These tumors will often grow next to or around neurovascular bundles, and an Allen’s test should always be performed preoperatively to confirm adequate blood supply by both ulnar and radial arteries as well as dual blood supply to an involved digit via the ulnar and radial proper digital arteries. It is important to completely excise the stalk because this will greatly reduce tumor recurrence even in the setting of residual tumor. If tumor is suspected to have extended into the joint, the joint must be opened and all tumor removed. Despite this being a benign lesion, local recurrence is varies widely from 4% to 44%. Some variants can mimic more aggressive processes, and malignancy must be considered if aggressive features are identified, such as direct bony invasion.

**Lipoma.** Lipomas of the hand and wrist may occur in multiple anatomic locations, including subcutaneous tissues; intramuscularly (especially thenar or hypothenar muscles); deep spaces; carpal tunnel or Guyon’s canal; and rarely bone or nerve. They typically present as a painless, slow-growing, soft, and mobile mass over a period of months to years. Painful findings suggest close approximation to a neurovascular structure or, less commonly, a malignant lesion such as liposarcoma. Lipomas do not transilluminate. They resemble mature fat histologically. X-rays typically reveal no abnormality. MRI is a helpful imaging modality to evaluate a lipoma and will show signal characteristics that are suggestive of adipose tissue.

Asymptomatic lesions with no aggressive findings may be observed. Marginal excision is recommended for symptomatic, painful, or enlarging lipomas or those that cause dysfunction. MRI is recommended for deep lipomas to evaluate proximity or involvement of critical structures, followed by marginal excision if MRI findings are consistent with a lipoma. If MRI findings are not consistent with a lipoma, incisional biopsy is warranted. Recurrence after marginal excision is rare.
Schwannoma. A schwannoma, also known as a neurilemmoma, is a type of benign peripheral nerve sheath tumor. It is the most common benign peripheral nerve sheath tumor of the upper extremity.\(^8^3\) The majority occur as single solitary masses. Patients with neurofibromatosis type 1 (NF1) or 2 (NF2) may develop multiple schwannomas involving large peripheral nerve trunks or bilateral acoustic schwannomas, respectively. These tumors arise from the Schwann cell and occur most often in the middle decades of life. They grow as painless, slow-growing, firm, round, well-encapsulated masses with a predilection toward flexor surfaces of the forearm and palm (given their presence of large nerves). Schwannomas grow from the peripheral nerve sheath and are usually connected by a pedicled stalk. The tumor is well demarcated and can be readily separated from the nerve fascicles (Fig. 44-26). Unlike neurofibromas, they do not grow within the nerve. Paresthesias or other neurologic findings may occur, but they are usually absent, as is the Tinel’s sign. Findings such as pain, paresthesias, or numbness should raise concern for a tumor causing a compressive neuropathy or a tumor that is malignant.\(^8^3\)

Histologic exam reveals Antoni type A palisades of spindle cells with large oval nuclei with interlacing fascicles. Less cellular regions appear as Antoni type B areas. Mutations of the schwannomin gene on chromosome 22 are found in 50% of sporadic cases and 100% of acoustic schwannomas in patients with NF2.\(^8^4\)

Surgical treatment is reserved for symptomatic tumors and those that require biopsy to rule out a malignant process. An MRI should be obtained prior to surgery to confirm that the tumor is not located within the nerve (i.e., a neurofibroma) and that it is consistent with a schwannoma. Operative treatment involves excisional biopsy. If the tumor is adherent to adjacent soft tissue or not encapsulated, incisional biopsy is performed and excision is delayed pending pathology results. Malignant degeneration is exceedingly rare.\(^8^3\)

Malignant Soft Tissue Tumors—Cutaneous

Squamous Cell Carcinoma. Squamous cell carcinoma (SCC) is the most common primary malignant tumor of the hand, accounting for 75% to 90% of all malignancies of the hand. Eleven percent of all cutaneous SCC occurs in the hand.\(^8^5\) It is the most common malignancy of the nail bed. Risk factors include sun exposure, radiation exposure, chronic ulcers, immunosuppression, xeroderma pigmentosa, and actinic keratosis. Marjolin’s ulcers represent malignant degeneration of old burn or traumatic wounds into an SCC and are a more aggressive type. Transplant patients on immunosuppression have a fourfold increased risk, and patients with xeroderma pigmentosa have a 65 to 200-fold increased risk of developing an SCC.\(^8^6\) They often develop as small, firm nodules or plaques with indistinct margins and surface irregularities ranging from smooth to verruciform or ulcerated (Fig. 44-27). They are locally invasive, with 2% to 5% lymph node involvement. Metastasis rates of up to 20% have been reported in radiation or burn wounds. Standard treatment is excision with 0.5- to 1.0-cm margins. Other treatment options include curettage and electrodesiccation, cryotherapy, and radiotherapy.\(^8^5\)

Basal Cell Carcinoma. Basal cell carcinoma (BCC) is the second most common primary malignancy of the hand, accounting for 3% to 12%; 2% to 3% of all BCCs occur on the hand. Risk factors are similar for SCC and include chronic sun exposure, light complexion, immunosuppression, inorganic arsenic exposure, and Gorlin’s syndrome. Presentation includes a small, well-defined nodule with a translucent, pearly border and overlying telangiectasias (Fig. 44-28). Metastasis is very rare. Standard treatment is excision with 5-mm margins. Other treatment options include curettage and electrodesiccation, cryotherapy, and radiotherapy.

Melanoma. Melanoma accounts for approximately 4% of skin cancers and is responsible of 80% of all deaths from skin cancer. Approximately 2% of all cutaneous melanomas occur in the hand.\(^8^7\) Risk factors include sun exposure (especially blistering sunburns as a child), dysplastic nevi, light complexion, family history of melanoma, immunosuppression, and congenital...
nevus. Pigmented lesions with irregular borders, color changes, increase in growth, or change in shape are suggestive of melanoma. Breslow thickness is the most important factor in predicting survival for a primary melanoma. Melanoma in situ lesions should be surgically excised with 0.5 cm margins. For lesions up to 1 mm in thickness, 1-cm margins should be used. Two centimeter margins should be used for lesions over 1 mm in thickness. Sentinel lymph node biopsy is done for lesions over 1 mm in thickness or for any lesion that is over 0.76 mm in thickness and exhibits ulceration or high mitotic rate. Any clinically palpable lymph node requires a formal lymph node dissection of the involved basin, as do sentinel lymph nodes positive for melanoma. Lymph node dissection has not been shown to offer any long-term survival benefit, but the information gained from sentinel lymph node biopsy (or lymph node dissection) does offer valuable staging information that is important for prognosis. For cases of subungual melanomas, DIP amputation is the current standard of care. A recent study reported similar recurrence and survival rates when comparing patients treated with either DIP amputations or wide local excision; however, there was insufficient evidence to conclude if one treatment was superior to another.

Malignant Soft Tissue Tumors—Noncutaneous

Primary soft tissue sarcomas of the upper extremity are very rare. Approximately 12,000 new cases of sarcomas are diagnosed each year and of those, only 15% occur in upper extremity. Statistical inference is limited due to the rare occurrence of these tumors, but mortality rate is very high despite the aggressive treatments. Fewer than 5% of soft tissue sarcomas of the upper extremity will develop lymph node metastasis. Cutaneous malignancies must be considered in the differential diagnosis for any patient with palpable lymph nodes in the setting of any upper extremity mass. Any lesion of the upper extremity that is over 5 cm in diameter, rapidly enlarges, or is painful should be considered malignant until proven otherwise.

Treatment for soft tissue sarcomas can range from palliative debulking to attempted curative resection. Many muscles of the upper extremity and their compartments cross joints (e.g., forearm flexors). Any malignancy within a compartment mandates complete resection of that compartment, and therefore, amputations must often be performed at levels much more proximal than the level of the actual tumor. Many soft tissue sarcomas are not responsive to radiation or chemotherapy, and use of these adjuvant treatments must be decided upon after discussion with medical and radiation oncologists in a multidisciplinary team. Several studies have shown higher mortality rates in patients who undergo initial tumor biopsy of sarcomas at institutions from which they do not ultimately receive treatment. These studies recommend biopsy be performed at the institution at which definitive treatment will be provided. Institutions best suited for such treatment should have pathologists familiar with soft tissue sarcomas, medical and radiation oncologists, surgical oncologists, and a multidisciplinary tumor board.

An in-depth review of each type of soft tissue sarcoma is beyond the scope of this chapter. Epithelioid sarcoma is the most common primary soft tissue sarcoma of the upper extremity and usually presents as a benign-like slow-growing mass during the third or fourth decades. It has a propensity for the forearm, palm, and digits. Spread to lymph nodes has been reported. It typically spreads along fascial planes. Synovial sarcoma is argued by some to be the most common primary soft tissue sarcoma of the hand and wrist, but the paucity of case reports is inconclusive. It is a high-grade malignancy that is painless and slow-growing and usually occurs adjacent to, but not involving, joints. It is most common in the second to fifth decades of life. Tumor size (greater than 5 cm) is positively correlated with mortality. Other sarcomas include malignant fibrous histiocytoma, liposarcoma, fibrosarcoma, dermatofibrosarcoma protuberans, and malignant peripheral nerve sheath tumors, and more information can be found in further selected reading. The majority of metastases to the hand involve secondary bone tumors and are discussed later in the section, “Secondary Metastatic Tumors.”

Benign Bone Tumors

Primary benign bone tumors of the hand and wrist make up a total of 7% of all primary benign bone tumors in the body. Benign tumors of cartilage origin comprise 79% of all primary benign bone tumors of the hand and wrist. This is the most common primary benign bone tumor of the hand and wrist and is of cartilage origin. Up to 90% of all bone tumors in the hand and wrist are enchondromas, with 35% to 54% of all enchondromas occurring in the hand and wrist. They are often found incidentally on X-rays taken for other reasons (e.g., hand trauma). They are usually solitary and favor the diaphysis of small tubular bones and are most common in the second and third decades of life. The most common location is in the proximal phalanges, followed by the metacarpals and then middle phalanges. Enchondroma has never been reported in the trapezoid. Presentation is usually asymptomatic, but pain may occur if there is a pathologic fracture or impending fracture. The etiology is believed to be from a fragment of cartilage from the central physis. Histology shows well-differentiated hyaline cartilage with lamellar bone and calcification.
Two variants of enchondroma include Ollier’s disease (multiple enchondromatosis) and Maffucci’s syndrome (multiple enchondromatosis associated with multiple soft tissue hemangiomas). Malignant transformation is very rare in the solitary form, but there is a 25% incidence by age 40 in Ollier’s patients and a 100% life-time incidence in Maffucci’s patients. When malignant transformation does occur, it is almost uniformly a chondrosarcoma with pain and rapid growth.\(^95\)

Diagnosis is usually made based on history, physical exam, and X-rays. There is a well-defined, multilobulated central lucency in the metaphysis or diaphysis that can expand causing cortical thinning or, sometimes, thickening (Fig. 44-29A). Further imaging is seldom needed, but a CT would be the study of choice.

Observation is indicated for asymptomatic enchondromas with no risk of impending fracture, followed by annual X-rays for 2 years. If a pathologic fracture is found, it is treated with immobilization until fracture union and then surgically treated. If there is any uncertainty as to whether it is an enchondroma, incisional biopsy is indicated, and definitive treatment is postponed pending final pathology. Symptomatic lesions and those with impending fracture are treated surgically. Surgical treatment consists of an open incisional biopsy and confirmation by frozen section that it is well-differentiated hyaline cartilage. Curettage and high-speed burring are used to ablate the tumor. Intraoperative fluoroscopy is used to confirm complete ablation (Fig. 44-29B). The defect is then packed with bone graft or bone substitute. Recurrence ranges from 2% to 15%. X-rays should be obtained serially after surgery.\(^94\)

**Periosteal Chondroma.** Periosteal chondromas are benign bone tumors of cartilage origin that arise most commonly within or adjacent to periosteum at the metaphyseal-diaphyseal junction in phalanges. They occur usually in the second or third decade as solitary lesions with pain, swelling, deformity, and possible pathologic fracture. X-rays reveal a subperiosteal lytic, unilobular lesion with erosion into adjacent cortex. There is often a rim of sclerosis. Histologically, they appear as aggressive cartilage with atypia, and it can be difficult to differentiate these from chondrosarcomas.\(^94\)

Diagnosis involves X-rays with incisional biopsy to confirm the benign diagnosis and avoid unnecessary amputation. Treatment includes en bloc resection of periosteum and corticocancellous bone. Recurrence is less than 4%.

**Osteoid Osteoma.** This is a tumor of bone origin. Approximately 5% to 15% of all osteoid osteomas occur in the hand and wrist and are most often found in the proximal phalanx or carpus. They usually occur in the second or third decade and present with a deep, dull ache that is classically worse at night and relieved by nonsteroidal anti-inflammatory drugs (NSAIDs). X-rays reveal a central lucency that is usually less than 1 cm in diameter surrounded by reactive sclerosis. Bone scan or CT is helpful to secure the diagnosis.\(^96\)

Treatment consists of NSAID therapy only, and resolution occurs at an average of 33 months. If the patient does not wish to undergo prolonged discomfort with conservative therapy, curettage or percutaneous ablation of the nucleus may be performed.\(^96\)

**Giant Cell Tumor of Bone.** Giant cell tumors of bone make up only 4% to 5% of all benign bone tumors in the body, and only 12% of these occur in the hand or wrist. Although its name is similar to that of “giant cell tumor of tendon sheath,” they are two separate tumors and do not share the same clinical or histopathologic characteristics. Approximately 2% occur in the hand and 10% occur in the distal radius; those within the distal radius are more aggressive. They usually occur in the fourth decade with pain and swelling and possibly pathologic fracture.\(^97\)

Giant cell tumor of the bone is unique in that it is benign on histology but does have metastatic potential and can cause death. It should be considered a low-grade malignancy.\(^97\)

Workup includes a CT of the chest and total-body scintigraphy to evaluate for metastases and multifocal lesions and MRI to evaluate the extent of local tissue involvement. The recommended treatment consists of surgical resection of the involved phalanges or metacarpals and wide excision of entire carpal rows. Treatment with curettage and adjuvant treatments only results in a high rate of recurrence. Local and systemic surveillance must be done for at least 10 years because metastasis has been reported to occur as late as 10 years postoperatively.\(^97,98\)

**Malignant Bone Tumors**

Malignant primary and secondary bone tumors of the hand, like soft tissue malignancies, are exceedingly rare. An in-depth
review is beyond the scope of this chapter. The same principles for soft tissue sarcomas of the upper extremity apply here with regard to evaluation, biopsy, and treatment.

Chondrosarcoma comprises 41% of all primary malignant bone tumors of the hand and wrist but only 1.5% of all chondrosarcomas overall. It is most likely to occur from malignant degeneration from a preexisting lesion, with enchondromatosis and osteochondromatosis being the most common. It usually presents as a slow-growing, painless mass in the fourth to sixth decades and can be difficult to differentiate from its benign counterparts. X-ray reveals endosteal erosion, cortical expansion, cortical destruction, and calcification. Metastasis has never been reported for chondrosarcomas of the hand. Chondrosarcomas are not responsive to chemotherapy or radiation.99

Osteosarcoma of the hand is exceedingly rare; only 0.18% of osteosarcomas occur in the hand. It usually presents as a painful swelling with pathologic fracture in the fifth to eighth decades of life. Radiation exposure is believed to be a possible risk factor. X-ray findings vary widely, with 90% of tumors occurring at a metaphyseal location. Findings include an osteoblastic or osteolytic lesion, cortical breakthrough with soft tissue extension, a “sunburst” pattern radially, or periosteal elevation (Codman’s triangle). The presence or absence of metastasis is the most important prognostic factor, with a 5-year survival of 70% in the absence of metastases and a 5-year survival of 10% if present. Preoperative chemotherapy is usually given, but radiation therapy plays no role.100

Secondary Metastatic Tumors
Metastases to the hand or wrist are rare, with only 0.1% of skeletal metastases occurring in the hand. The majority of metastases to the hand are bone lesions, but soft tissue metastases have been reported. The most common primary site is the lung (40%), followed by the kidney (13%) and the breast (11%). Approximately 16% will have no known diagnosis of cancer.101 The most common sites are the distal phalanges, followed by the proximal and middle phalanges, metacarpals, and carpus. Patients will present with pain, swelling, and erythema. Differential diagnosis includes felon, gout, osteomyelitis, trauma, RA, or skin cancer. Treatment of a hand or wrist metastatic lesion must not interfere with treatment of the primary cancer. Treatment is usually palliative (simple excision or amputation). The average life expectancy for these patients is less than 6 months.101

BURNS
The palm of the hand makes up approximately 1% of the total body surface area. A burn involving the entire hand and digits is unlikely to cause life-threatening injury or shock, but seemingly small burns to the hand may cause severe permanent loss of function if not treated appropriately. Burns to the hand can cause serious short- and long-term disability. All burns to the hand are considered severe injuries that warrant transfer to a dedicated burn center for specialized treatment. This management will include a multidisciplinary team consisting of hand surgeons, burn surgeons, burn-specialized nurses, occupational therapists, case managers, and social workers.

Superficial burns involve damage to the epidermis only and present with erythema, no blistering, and full sensation with blanching of skin. These will heal without scarring. Superficial partial-thickness burns involve damage to the papillary dermis; all skin appendages are preserved, and therefore, these readily reepithelialize with minimal to no scarring. Superficial partial-thickness burns are sensate and present with pain, erythema, blistering, and blanching of skin. Topical dressings are the mainstay of treatment. Deep partial-thickness burns involve damage to the reticular dermis with damage to skin appendages, as well as the dermal plexus blood vessels and nerves. These have decreased sensation and no cap refill and appear pale or white. Blistering may be present. Damage to the skin appendages and blood supply in the dermal plexus precludes spontaneous healing without scar. Excision with skin grafting is needed. Third-degree burns involve full-thickness damage through the dermis and are insensate with no blistering. They appear dry, leathery, and even charred.

Acute Management
Advanced trauma life support guidelines should be followed. After primary survey, circulation to the hand should be assessed. Palpation and Doppler ultrasound should be used to evaluate blood flow within the radial and ulnar arteries, the palmar arches, and digital blood flow at the radial and ulnar aspect of each volar digital pad. A sensorimotor exam should be performed. Objective evidence of inadequate perfusion (i.e., deteriorating clinical exam with changes in or loss of pulse or Doppler signal) indicates the need for escharotomy, especially in the setting of circumferential burns. Escharotomy may be performed at bedside with scalpel or electrocautery under local anesthesia or intravenous sedation. In the forearm, axially oriented midradial and midulnar incisions are made for the entire extent of the burn. Escharotomy should proceed as distally as necessary into the wrist and hand to restore perfusion. Digital escharotomies are made via a midaxial (the middle of the longitudinal axis on sagittal view) incision over the radial aspects of the thumb and small finger and the ulnar aspects of the index, middle, and ring fingers.102 These locations for digital escharotomies avoid painful scars on the heavy-contact surfaces of each respective digit. After primary survey, vascular, and sensorimotor exams are complete, careful documentation should be made of all burns. This is best done with a Lund and Browder chart and includes location, surface area, and initial depth of burn.

The burns should be dressed as soon as examination is complete. Gauze moistened with normal saline is a good initial dressing because it is easy, readily available, and will not leave ointment or cream on the wounds, which can hinder frequent examinations in the initial period. It is critical that no dressing is wrapped in a circumferential manner around any body part. Edema and swelling can lead to extremity ischemia if a circumferential dressing is in place. It is important to maintain body temperature above 37°C, especially in burn patients who have lost thermoregulatory function of the skin and now have moist dressings in place. The hands should be elevated above heart level to decrease edema formation, which can hinder motion and lead to late scar contracture. The hand should be splinted in the intrinsic plus position with the MP s flexed to 90° (placing MP collateral ligaments under tension), the IPs in straight extension (prevents volar plate adhesion), and the wrist in approximately 15° of extension.103 In rare cases, Kirschner wires or heavy steel wires/pins are needed to keep a joint in proper position. These are placed percutaneously through the involved joint and serve as a temporary joint stabilizer.

After the primary and secondary surveys are complete, the wound should be evaluated again. Devitalized tissue should be
debrided. Wounds should be cleansed twice daily, typically with normal saline. Second-degree superficial burns may be dressed with Xeroform gauze and bacitracin. Silver sulfadiazine cream is another option for any second- or third-degree wound. It covers gram-positive and gram-negative microbes, but it does not penetrate eschar. It should be applied at least one-sixteenth of an inch thick. Sulfamylon can be used in conjunction with silver sulfadiazine or alone. It deeply penetrates eschar and tissues and has good gram-positive coverage.

Surgical Management
Any burn wound will eventually heal with proper wound care. However, this may involve unacceptable scarring, deformity, contractures, pain, and unstable wounds that are prone to breakdown. The goal is to restore preinjury function as much as possible with a wound that is durable, supple, nonpainful, and allows the patient to return to society as an active member. Local wound care is the ideal treatment for wounds that can heal completely within 14 days while not sacrificing function. For deep partial-thickness or full-thickness burns, early surgical excision and skin grafting is necessary.103

Considerable controversy surrounds the need, timing, and method of grafting burns. Careful consideration must be given to the patient’s overall status, their preinjury state, and the type of work and recreational activities they enjoyed in order to have a better understanding of which issues should be addressed. Tangential excision of the wounds should be performed under tourniquet to minimize blood loss and is carried down to viable tissue. Avoid excising through fascia (epimysium) overlying muscles or exposing tendons, bone, joint capsules, or neurovascular structures. Tissues capable of receiving a skin graft include well-vascularized fat, muscle, perineurium, paratenon, perichondrium, and periosteum. Exposure of deep structures without an adequately graftable bed mandates further coverage before skin grafting can occur (discussed later in “Reconstruction”).

Once there is an adequate bed, grafting is the next step. If there is any doubt as to whether the wound bed can support a skin graft, a temporary dressing such as Allograft (human cadaver skin) should be placed and the patient reexamined frequently for signs of granulation tissue and wound bed viability. It can remain in place for up to 14 days before rejection and can serve as a way of “testing” if a wound is ready to receive a skin graft. Skin grafts to the dorsum of the hand are typically split-thickness sheet grafts (not meshed), as sheet grafts have a superior aesthetic appearance. Skin grafts to the palmar aspects of the hand should be full-thickness in order to provide the dermal durability needed for daily functions. Skin grafts are secured with staples, sutures, fibrin glue, or even skin glue. It is important to bolster every skin graft. This prevents shearing loss and also keeps the skin graft in contact with the wound bed, preventing fluid collections that can lead to graft loss. A bolster may consist of a tie-over bolster and a splint or a negative-pressure dressing. The hand should be splinted in intrinsic plus for 7 days after skin grafting. Once the graft is adherent, hand therapy should begin, consisting of active and passive range-of-motion exercises and modalities.103

Reconstruction
Reconstruction of burn wounds can begin as early as the acute setting and continue into the subacute and late stages. Burns may initially be superficial but later convert to deep burns (especially with grease, oil, and alkali burns) due to infection, tissue desiccation, or continued trauma, or they may be deep from the outset of injury. Debridement or excision of burns may result in exposure of viable muscle, bone, tendon, cartilage, joints, and neurovascular structures, as well as loss of fascial layers that are required for overlying soft tissue to glide during movement. Simply skin grafting these exposed structures will result in unstable wounds that are prone to chronic breakdown. Soft tissue contractures will develop as the skin grafts adhere to the structures, effectively anchoring them in static position. This is especially true for tendons, where gliding capability is paramount for function. Flap coverage is required in these situations. The reversed radial forearm flap is a local flap and is often the first choice for flap coverage of the hand. If the zone of injury or size of defect precludes its use, other skin and fat flaps, including the free lateral arm, free anterolateral thigh, or even free parascapular flaps, may be useful, provided the patient can tolerate a free tissue transfer (see Chapter 45) operation (Fig. 44-30). The digits may also be buried subcutaneously in the lower abdominal skin or groin crease. Vascular ingrowth from the digits into the abdominal or groin skin occurs over 2 to 3 weeks, allowing division of the flap(s) and achieving full-thickness coverage of the wounds.104

An acellular dermal regenerative substitute (e.g., Integra) may be used for wounds that have exposed structures and require more durability than is offered by a skin graft such as full-thickness loss overlying the extensor tendons of the wrist and hand.105 Dermal substitute is a good option for wounds that are not extensive enough to warrant a flap and for patients who are poor candidates for an extensive surgery. Integra is composed of acellular cross-linked bovine tendon collagen and glycosaminoglycan with an overlying silicone sheet. It is applied much like a skin graft. After incorporation in 14 to 21 days, it is capable of accepting a skin graft (after removing the silicone sheet). Conceptually, it works by replacing the lost dermis and adds durability to a wound bed. It may be reapplied multiple times to the same area if thicker neodermis is desired. Although cultured autologous keratinocytes have been used, they are expensive, time-consuming, and do not provide prompt or durable coverage.

Web space contractures are the most common deformity resulting after hand burns. They may occur late despite the best efforts. In the normal web space, the leading edge of the volar

Figure 44-30. Free anterolateral thigh flap reconstruction of a large dorsal hand wound. Once wound coverage is stable, this flap will need to be surgically revised to achieve proper contour.
lavage, phosphate buffer soaks and immediate surgical excision. Cement can result in chemical burns and should be treated with immediate irrigation and topical antibacterial ointments. Alkaline and acid burns require copious irrigation with water, with alkali burns often requiring hours of irrigation. Phenol burns should be irrigated with dilute polyethylene glycol wash followed by high-flow water lavage.\textsuperscript{106}

**VASCULAR DISEASE**

Vascular disease encompasses a broad spectrum of disorders leading to compromised perfusion to the hand and digits and may potentially cause ischemia and necrosis. Chronic vascular disorders tend to develop slowly and are typically seen in older patients. This includes progressive thrombosis, aneurysms, systemic vasculopathy, and vasospastic disorders. Disorders unique or common to the hand are discussed in the following sections.

**Progressive Thrombotic Disease**

Hypothenar hammer syndrome involves occlusion of the ulnar artery at the wrist and is the most common occlusive vascular disorder of the upper extremity. The etiology is believed to be chronic trauma to the ulnar artery as it exits Guyon’s canal. The classic example is a construction worker who frequently uses heavy equipment, such as jackhammers, that cause prolonged vibration and repetitive impact on the ulnar aspect of the palm. This causes periadventitial arterial damage that results in scarring and eventual compression, as well as medial and intimal damage.\textsuperscript{107} The artery then becomes weakened and prone to aneurysm and/or thrombosis. If a thrombus forms, it may embolize, producing digital ischemia. Symptoms may be chronic or acute and include pain, numbness and tingling, weakness of grip, discoloration of the fingers, and even gangrene or ulcers of the fingertips.

If acute in onset, proximal occlusions may be extracted with a balloon catheter or, sometimes, under direct vision via an arteriotomy. Very distal embolism may require infusion of thrombolitics to dissolve clots and allow reperfusion. Large-vessel acute embolism and reperfusion may result in edema and compartment syndrome, requiring fasciotomy. A high index of suspicion must be maintained.

For the more common scenario of chronic, progressive occlusion, the involved segment of ulnar artery should be resected. There is disagreement in the literature regarding whether simple ligation and excision is sufficient for patients with sufficient distal flow or if all patients should undergo vascular reconstruction.\textsuperscript{108} The authors’ personal preference is to reconstruct all patients.

**Systemic Vasculopathy**

Buerger’s disease (thromboangiitis obliterans) is an inflammatory occlusive disease affecting small and medium-sized arteries and veins. It is strongly influenced by smoking and will often resolve upon smoking cessation. The disease is classified into acute, intermediate, and chronic, depending on histologic progression of the disease. Migratory phlebitis occurs distal to the elbow, resulting in ischemia, rest pain, and ulceration and necrosis of the digits. It can continue to cause more proximal ischemia and ultimately lead to loss of the hands. Treatment must start with smoking cessation. Failure to stop smoking will make any surgical intervention unsuccessful. Arteriography is useful to determine arterial flow and whether bypass is possible.
If direct bypass is not possible, alternatives include arterialization of the venous system by connecting the dorsal venous network to the brachial artery or possible free microvascular omental transfer beneath the dorsal forearm or hand for indirect revascularization.109

**Vasospastic Disorders**

Raynaud’s syndrome results from excessive sympathetic nervous system stimulation. Perfusion is diminished and fingers often become cyanotic. Although the onset of the symptoms is benign, chronic episodes can result in atrophic changes and painful ulceration or gangrene of the digits. Raynaud’s disease occurs without another associated disease. This disease predominately affects young women and is often bilateral. The vascular system is structurally intact without any obstructions. There is no ulceration, gangrene, or digit loss. In contrast, Raynaud’s phenomenon is associated with an underlying connective tissue disorder, such as scleroderma. Arterial stenosis is present due to disease changes in blood vessels as a result of the specific medical disorder.110

Scleroderma is an autoimmune connective tissue disorder resulting in fibrosis and abnormal collagen deposition in tissue. Many organs can be affected, with the skin most commonly and noticeably involved. In this disease, blood vessels are injured by intimal fibrosis leading to microvascular disease. The vessels become subject to Raynaud’s phenomenon, and patients develop painful, ulcerated, and sometimes necrotic digits.109,110

Sympathectomy can provide pain relief and healing of ulcers for patients with scleroderma and Raynaud’s phenomenon. In this procedure, adventitia is stripped from the radial artery, ulnar artery, superficial palmar arch, and digital arteries in various combinations based on the affected digits being treated. The decrease in sympathetic tone allows for vasodilation and increased blood flow. If the patient notes significant distal pain relief and/or previously ischemic tissue improves in color after a test administration of local anesthetic, sympathectomy may provide the same results in a long-term fashion.111

Recently, several studies have investigated the use of botulinum toxin on improving digital perfusion in patients with Raynaud’s. Reports have shown improved objective measurements of hand function 8-12 weeks after injection.112

**CONGENITAL DIFFERENCES**

Congenital differences in a newborn can be particularly disabling as the child learns to interact with the environment by using the hands. The degree of anomaly can range from minor, such as a digital disproportion, to severe, such as total absence of a forearm bone. In recent years, increasing knowledge of the molecular basis of embryonic limb development has significantly enhanced the understanding of congenital differences. Congenital hand differences have an incidence of 1:1500 births. The two most common differences encountered are syndactyly and polydactyly.113

There are numerous classification systems for hand differences. The Swanson classification, adopted by the American Society for Surgery of the Hand, delineates seven groups organized based on anatomic parts affected by types of embryonic failures.114,115

**Failure of Formation**

The failure of the formation of parts is a group of congenital differences that forms as a result of a transverse or longitudinal arrest of development. Conditions in this group include radial club hand, a deformity that involves some or all of the tissues on the radial side of the forearm and hand, and ulnar club hand, which involves underdevelopment or absence of the ulnar-sided bones.

**Failure of Differentiation**

The failure of the differentiation of parts comprises conditions where the tissues of the hand fail to separate during embryogenesis. Syndactyly, in which two or more fingers are fused together, is the most common congenital hand deformity and occurs in 7 out of every 10,000 live births. There is a familial tendency to develop this deformity. This deformity often involves both hands, and males are more often affected than females. Syndactyly is classified as either simple (soft tissue only) or complex (bone and/or cartilage also involved), and complete (full length of the digits) or incomplete (less than the full length).

Surgical release of syndactyly requires the use of local flaps to create a floor for the interdigital web space and to partially surface the adjacent sides of the separated digits (Fig. 44-32). Residual defects along the sides of the separated fingers are covered with full-thickness skin grafts. Surgery usually is performed at 6 to 12 months of age.

**Duplication**

Duplication of digits is also known as polydactyly. Radial polydactyly usually manifests as thumb duplication. Wessel described a classification system for thumb duplications based on the level of bifurcation.116 When two thumbs are present in the same hand, they are rarely both normal in size, alignment, and mobility. In the most common form of thumb duplication, a single broad metacarpal supports two proximal phalanges, each of which supports a distal phalanx. Optimal reconstruction requires merging of elements of both component digits. Usually the ulnar thumb is maintained. If the duplication occurs at the MP joint, the radial collateral ligament is preserved with the metacarpal and attached to the proximal phalanx of the retained ulnar thumb. Surgery is usually performed at 6 to 12 months of age. Ulnar-sided polydactyly may often be treated by simple excision of the extra digit.

**Overgrowth**

Overgrowth of digits is also known as macrodactyly, which causes an abnormally large digit. In this situation, the hand and the forearm also may be involved. In this rare condition, all parts of a digit are affected; however, in most cases, only one digit is involved, and it is usually the index finger. This condition is more commonly seen in males. Surgical treatment of this condition is complex, and the outcomes may be less than desirable. Sometimes, amputation of the enlarged digit provides the best functional result.

**Constriction Band Syndrome**

Underdeveloped fingers or thumbs are associated with many congenital hand deformities. Surgical treatment is not always required to correct these deformities. Underdeveloped fingers may include the following: small digits (brachydactyly), missing muscles, underdeveloped or missing bones, or absence of a digit.

**Generalized Skeletal Anomalies and Syndromes**

This is a rare and complex group of unclassified problems.
Hand transplantation was first performed in humans in the late 1990s both in Louisville, Kentucky, and Lyon, France. The treating surgeons were able to successfully remove an upper extremity from a brain-dead donor, attach it to an upper extremity amputee, and have the tissue survive. In the subsequent 15 years, many additional centers have achieved technical success with upper extremity transplantation as well.

The technical considerations of hand transplantation have proven to be only the beginning of challenges in bringing this treatment option to the general public. Replantation of an amputated limb was first reported by Malt in 1962. In a limb replantation, there is a zone of injury, and cold preservation of the amputated part does not begin immediately. In a limb transplant, the harvest can be done as proximally as necessary to ensure that only healthy tissue is present on both sides of the repair and to obviate the need for limb shortening, and cold preservation of the amputated part can begin immediately after harvest.

A major concern regarding the use of limb transplantation is the immunosuppression medications required to prevent rejection of the transplanted limb. Unlike organ transplantation, which provides a critical organ without which the recipient could not survive or would require chronic mechanical support (e.g., hemodialysis), the absence of one or even multiple limbs does not represent an immediate threat to a patient’s survival. Multiple studies have documented the nephrotoxic and other side effects of tacrolimus (FK 506), the principle antirejection agent used in transplant immunomodulation protocols. Due to these concerns, much research has been directed at minimizing the amount of antirejection medication as well as promoting tolerance or even chimerism. Donor bone marrow transplantation to the limb transplant recipient has been shown to be beneficial toward this purpose and is part of the limb transplant protocol in some centers. Recent research with donor bone marrow infusions has shown that lower levels of immunosuppressive drugs may be possible, as well as fewer immunosuppressive agents. Further research is needed in order to determine the efficacy and utility of donor bone marrow transusions and how they impact transplant recipients in the short and long term.

The final challenge in consideration of a patient for limb transplantation is selection of an appropriate candidate. There are multiple patient factors that need to be considered to determine if a patient is an appropriate candidate for hand transplantation. These include medical concerns, such as immunologic issues (both antibodies and the presence of occult neoplasms or indolent viruses such as cytomegalovirus), hematologic issues including coagulopathies, and anatomic issues such as quality of skin envelope and amputation level of the bone and neuro-muscular structures. Psychological and social factors must also be considered related to the recipient’s ability to comply with postoperative medication and therapy protocols as well as to cope with a continuous visible presence of a limb originating from another person.

The promise of upper limb transplantation as a reconstructive technique remains high. Both civilian and military amputees stand to receive a marked functional benefit from this treatment. With the number of transplants performed worldwide...
approaching 100 as well as decades of animal research, understanding of how best to use this technique from functional, patient safety, and cost-effectiveness standpoints continues to grow.

REFERENCES

Entries highlighted in bright blue are key references.

38. Al-Qattan MM, Abou Al-Shaar H, Al Mugaren FM. Non-union without avascular necrosis of finger phalangeal neck
50. Trumble TE, Diao E, Abrams RA, Gilbert-Anderson MM. Single-portal endoscopic carpal tunnel release compared with open release: a prospective, randomized trial. J Bone Joint Surg Am. 2002;84-a(7):1107-1115. Carpal tunnel release is one of the most common procedures performed by hand surgeons. This study by Trumble highlights that although patients undergoing endoscopic carpal tunnel release have less pain in the immediate postoperative period, clinical outcomes after 3 months show no difference compared to traditional open approaches.
63. Chung KC, Pushman AG. Current concepts in the management of the rheumatoid hand. J Hand Surg Am. 2011;36(4):736-747; quiz 747. Surgical treatment for rheumatoid arthritis of the hand has decreased due to the advances in medical management. This article serves as thorough review for hand surgeons on the treatment of rheumatoid hand.
69. van Rijssen AL, ter Linden H, Werker PM. Five-year results of a randomized clinical trial on treatment in Dupuytren’s disease: percutaneous needle fasciotomy versus limited fasciectomy. Plast Reconstr Surg. 2012;129:469-477. Although percutaneous needle fasciotomy is less invasive than limited fasciectomy, this study showed that fasciectomy provided more durable and lasting results.
93. Murray PM. Soft tissue sarcoma of the upper extremity. *Hand Clin.* 2004;20(3):325-333, vii. The subject of soft tissue sarcomas is very broad and specific. This article by Murray provides a concise and accurate summary of soft tissue sarcomas of the upper extremity.


Plastic and Reconstructive Surgery

Rajiv Y. Chandawarkar, Michael J. Miller, Brian C. Kellogg, Steven A. Schulz, Ian L. Valerio, and Richard E. Kirschner

INTRODUCTION

Plastic and reconstructive surgery is a unique subspecialty of surgery that consists of a set of techniques intended to modify the amount, position, quality, or organization of tissues in order to restore function and appearance. The name of the field is derived from the Greek word plastikos, which means “to mold.” An object is considered plastic if its shape can be modified without destruction. In this sense, all human tissues have some degree of plasticity. They can be nondestructively modified if the surgeon adheres to certain principles. Understanding and applying these principles to solve clinical problems is the essence of plastic and reconstructive surgery. Although informal references to this type of surgery can be found in the modern literature as early as the 17th century, American surgeon John Staige Davis published the first textbook dedicated to the field in 1919, entitled Plastic Surgery—Its Principles and Practice. He coined the term that we have used to refer to the specialty ever since. Science has always evolved in a nonlinear fashion: seminal discoveries in different parts of the world have all collectively fueled progress and addressed an unmet need. The evolution of plastic and reconstructive surgery has followed the same path: the Edwin Smith Papyrus¹ (Egypt, 1600 B.C.) (Fig. 45-1) described facial reconstruction; the Shushruta Samhita (India, 1500 B.C.) (Fig. 45-2) described nasal reconstruction; and Aulus Cornelius Celsus (Rome, 1 A.D.) described operations for facial reconstruction. The underlying impetus for this evolution is the common unmet need for restoring defects, be they congenital, traumatic, or functional.

This strong thread of advances in reconstructive surgery continues even today. What does seem under-recognized is that the clinical practice of plastic and reconstructive surgery touches on every other area of surgery. Enhanced reconstructive capabilities strengthen all other specialties significantly, such as the ability to safely perform radical cancer operations, salvage traumatic limbs, or extend the reach of neonatal medicine by congenital reconstruction. Each surgical specialty encounters problems that might be addressed by some form of tissue repair, modification, rearrangement, transfer, or replacement. Since its inception, plastic surgeons have routinely responded to the medical needs of the society and helped restore form and function. One of the most powerful examples of this response is the advances that occurred as a result of World Wars I and II. Walter Yeo, a sailor injured at the Battle of Jutland, is assumed to have received plastic surgery in 1917. The photograph shows him before (Fig. 45-3, left) and after (right) receiving a flap surgery performed by Gillies.

The Gulf war and the conflicts in the Middle East have prompted several revolutionary reconstructive surgical advances in limb salvage, microsurgery, supermicrosurgery, hand, face, and abdominal wall transplantation. Plastic surgeons have also targeted muscle reinnervation, tissue engineering, and regenerative medicine.

When society calls, plastic surgeons rise to the challenge and create novel methods to address its needs. For example, neurosurgeons at times must replace or stabilize bone in the cranium or spine, and healthy soft tissue coverage is essential for optimal healing. Head and neck surgeons face tissue replacement problems in order to restore normal function and appearance after major tumor ablation. Thoracic surgeons must manage bronchopleural fistulae, esophageal defects, or loss of chest wall integrity after trauma or tumor resection. Cardiologists and cardiac surgeons at times face complicated wound
Key Points

1. It is critical to understand the physiologic basis and rationale of wound healing in order to further assimilate surgical and nonsurgical care of wounds and methods of wound care.

2. Understanding the reconstructive choices in tissue repair cases is critical for any surgeon. The principles of soft tissue and skin repair are important for the reconstruction of defects, whether in a trauma situation of after excision of lesions.

3. Children with cleft and craniofacial differences have complex medical, surgical, and social needs. Coordinated, interdisciplinary team care is crucial to success.

4. Robin sequence, characterized by micrognathia, glossophtosis, and airway obstruction, can be managed with prone positioning, tongue-lip adhesion, mandibular distraction osteogenesis, or tracheostomy.

5. The first-line treatment for high-risk hemangiomas is oral propranolol, which can induce rapid involution and has a more favorable side effect profile than systemic steroids.

6. The coordination of care for patients in a trauma department is an important part of a surgeon’s role, whether that role be as a trauma emergency department surgeon or a surgeon in practice.

7. The careful evaluation of a patient in a polytrauma involves a thorough assessment of internal and soft tissue injuries, planning of care, and the appropriate triage of reconstructive procedures. As a leader in a trauma bay of the trauma service, the surgeon typically assumes a captain’s role in decision-making.

8. Principles of oncologic reconstruction have evolved significantly, and a deeper understanding of these reconstructive choices is essential for a surgeon who is often the first point of contact for cancer patients and responsible for making critical referrals.

9. The combined work of general surgeons and reconstructive plastic surgeons has revolutionized the care of abdominal wall defects, including ventral hernias, repair after tumor ablation, and bariatric surgery.

10. Any critical care unit or a medical surgical team that takes care of debilitated patients needs a detailed understanding of pressure sores, including their etiology and the reconstructive options that are available to these patients.

Infections, sternal osteomyelitis, or failure of soft tissue coverage that leads to exposure and contamination of implanted devices such as left ventricular assist devices or cardiac pacemakers. Orthopedic surgeons managing segmental bone defects in the extremities at times require replacement by surgical transfer of vascularized bone segments rather than conventional bone grafts or alloplastic substitutes. Urologists, colorectal surgeons, and gynecologists who commonly perform surgery in the perineum encounter nonhealing wounds or fistulae. All of these problems may be managed or potentially prevented by judicious application of tissue methods developed and practiced by plastic and reconstructive surgeons.

Plastic and reconstructive surgery is a field characterized by innovation, and it has yielded important contributions to other surgical specialties. These include notable advances in hand and upper extremity surgery, craniofacial surgery, peripheral nerve surgery, and reconstructive microsurgery. Entirely new fields of have emerged from plastic surgery research. Joseph E. Murray, a Boston plastic surgeon, and his team performed the first renal transplantation procedures and laid the foundation for modern organ transplantation, an achievement for which he was awarded the Nobel Prize in Medicine in 1990 (Fig. 45-4). This spirit of innovation continues with ongoing active research by plastic surgeons in composite tissue allotransplantation, tissue engineering, biomaterials, cell transplantation, regenerative medicine, computer-assisted surgical planning, medical application of three-dimensional manufacturing methods, infection control, and outcomes research. Plastic and reconstructive surgery is a vibrant field that brings tremendous value to people’s health and quality of life through life-changing reconstructive, restorative, and transformative surgeries.
**PURPOSE**

The purpose of this chapter is to inform about the general principles of plastic and reconstructive surgery, which apply to all areas of surgery, and to provide current examples of practice. Studying this chapter will help the reader to understand (a) the principles of plastic surgery that translate into other surgical specialties; (b) the kind of clinical problems that may be addressed using plastic surgery techniques; and (c) the types of research found in plastic and reconstructive surgery. It will make clearer the nature of the field and its role in the multidisciplinary care environment of modern healthcare.

**GENERAL PRINCIPLES**

General principles of plastic surgery relate to technical aspects of incision planning and wound repair. These principles apply to all surgical disciplines. As such, every surgeon can benefit from learning and applying them. Previously, tremendous emphasis was placed on simply understanding the nature of skin, which is completely justified; however, over the past few years plastic surgical focus has expanded to include the entire integument. Muscles, fascia, fat, skeletal framework, nerves, vascular networks, and their dynamic interactions have become far more important factors that are choreographed in most reconstructive processes.

**Skin Incisions**

From a surgical viewpoint, the skin is a multilayered tissue formed by dermis and epidermis. It is the largest organ in the human body and exists in a state of dynamic equilibrium from the balance of tension created by external and internal factors. Externally, skin and underlying subcutaneous tissue are acted on by gravity and clothing. Internal factors include skin elasticity, which is simply the ability to stretch and return to prestretch architecture upon removal of the stretch. The dermis is composed of different types of collagen and elastic protein fibers (elastin), and epidermis, composed primarily of cells anchored together in various stages of maturation. The skin serves important functions of thermoregulation, affording tactile sensation, and protection from foreign materials and microorganisms. Areas of skin exposed to view in normal clothing play a significant role in personal appearance and social interaction. As a result, even favorable scars from surgical incisions can have an undesirable effect on personal appearance. Thoughtful placement and performance of a surgical incision will minimize the risk of adverse consequences that can result in short- and long-term morbidity.

Human skin exists in a resting state of tension caused by gravity and its conformation over underlying structures between sites that are tethered by subcutaneous fibrous tissue, which secure the deep surface of the dermis to underlying points of fixation. When the skin is incised linearly, the wound edges separate in a predictable fashion forming an ellipse with the long axis perpendicular to the lines of greatest tension. These tension lines are often called “Langer’s lines,” after Carl Langer, a 19th century anatomist from Vienna who first described them based on studies in fresh cadavers (Fig. 45-5). Later, Borges described relaxed skin tension lines, which follow furrows formed when the skin is relaxed and are produced by pinching the skin. Incisions placed parallel to these lines often heal with less conspicuous scar because the skin often has natural wrinkles following these lines and there is less tension perpendicular to the orientation of the wound1 (Fig. 45-6). Based on these principles,2 a recommended pattern for incisions can be made (Fig. 45-7).

Using the proper technique for creating and repairing skin incisions ensures uncomplicated wound healing with few distorting surface scars. The epidermis and superficial dermis should be incised sharply with a scalpel. The incision is then continued through the deep dermis and subdermal plexus of blood vessels with electrocautery. This technique helps to minimize collateral tissue injury along the wound margins to facilitate prompt and reliable healing. It is essential to maintain the orientation of the scalpel or electrocautery blade perpendicular to the surface of the skin in order to facilitate accurate reapproximation during wound closure. As the incision is deepened through the subcutaneous tissue to expose underlying structures, it is important to avoid creating multiple pathways through the tissue, which can create focal areas of devitalized tissue that form a nidus of infection or lead to delayed wound healing.
surgeon should extend the incision through the subcutaneous fat by tracing the same line each time with the scalpel or electrocautery in order to reach the deeper structures.

Traumatic wounds do not permit the same careful planning that is possible with incisions made in undamaged skin. Nevertheless, optimum repair of traumatic lacerations involves similar principles applicable in nontraumatic circumstances. The surgeon must remove as much traumatized tissue as possible from the wound edges, converting the uncontrolled traumatic wound into a controlled surgical wound. All devitalized tissue is excised. The same principles of making incisions perpendicular to the skin surface and avoiding creating multiple pathways through the subcutaneous tissues apply. In this process, an attempt can be made to reorient the wound into a more favorable direction. A variety of methods are available to perform this reorientation, and they often involve creating small local flaps of undamaged tissue using geometric tissue rearrangements. These techniques will be considered later in this chapter. Following these principles increases the likelihood of uncomplicated wound healing and reduces the need for later treatment of unfavorable scars. However, there are situations in which the direction of the incision has been preestablished, as in acute lacerations, burns, or old contracted and distorting scars. In these circumstances, the principles of proper incision placement can be combined with simple surgical techniques to reorient the scar and lessen the deformity.

When making an incision in an area of previous scarring, such as in a scar revision or a reoperation, it is preferable to completely excise the scar when making the skin incision and not simply make the incision through the old scar. Closing scarred wound edges increases the likelihood of delayed wound healing, infections, and unfavorable new scars. It only takes a few moments to make the skin incision outside of the area of scarring through unscarred skin. Once the skin incisions on each side of the previous scar reach into the subcutaneous tissue, then the surface scar can be removed completely at the subdermal level. This approach ensures that the final repair relies on undamaged tissues, thus facilitating uncomplicated healing and lowering the risk of an unfavorable scar.

**Incision Repair**

A well-performed skin incision sets the stage for an accurate repair that minimizes the risk of unfavorable scarring. An unfavorable scar is characterized by excessive amount of collagen

---

**Figure 45-5.** “Langer’s lines,” named after Carl Langer, a 19th century anatomist from Vienna.

**Figure 45-6.** Lines of relaxed skin tension.

**Figure 45-7.** Planning of incisions based on lines of skin tension.
deposition, leading to hypertrophic scarring or keloid formation (Fig. 45-8). The difference between them is that a hypertrophic scar stops growing 6 months after the injury, whereas a keloid continues to grow, even growing well beyond its borders. Accurate approximation and stabilization of the skin edges helps to minimize the amount of collagen deposition required for skin healing. The most important layer to approximate is the dermis because this layer contains the healing elements such as blood supply and cellular elements that create the extracellular matrix necessary for healing. Optimal wound closure involves placing deep dermal sutures followed by superficial sutures that incorporated the upper layers of the dermis and epidermis. Absorbable deep dermal sutures have the advantage of disappearing over time; however, they can promote prolonged inflammation during this process. Nonabsorbable sutures minimize inflammation and might be indicated in individuals who are particularly prone to scar formation. A step-off between each side of the wound should be avoided because an uneven surface on each side of the wound can cause a shadow that accentuates the presence of the scar. Stability between the two wound edges is important because motion between the two sides of the wound prolongs the inflammatory phase of healing and requires additional collagen to be deposited. The timing of suture removal depends on the type of suture placed in the superficial closure. Sutures placed at the surface that go deep into the dermis can leave additional scarring at the entry and exit points of the suture material in addition to the incisional scar. Sutures like this should be removed within the first week. If the superficial sutures are placed more shallowly in the dermis, there is a reduced tendency to form additional scarring. A subcuticular suture may be used instead of simple sutures. This type of technique avoids the risk of additional scarring along the wound edge; however, it can be more difficult to accurately reapproximate the skin edges without a step-off between the two sides.

**Wound Healing**

In the United States, nonhealing wounds affect about 3 to 6 million people, with persons 65 years and older accounting for 85% of these events. The annual cost of this problem is estimated to be as high as $25 billion for hospital admissions, antibiotics, and local wound care.

Normal wound healing is achieved through four highly choreographed, overlapping biophysiologic phases: hemostasis, inflammation, proliferation, and tissue remodeling or resolution. Each phase initiates a cascading set of processes critical to the desired result of a healed wound.

**Phases of Wound Healing**

There are different processes that characterize healing in several types of tissue, such as skin, muscle, or bone, and there is a strong underlying mechanism that is best understood in terms of a simple skin injury. The process of wound healing is comprised of four integrated processes that overlap: (a) bleeding and hemostasis, (b) inflammation, (c) proliferation, and (d) tissue modeling or resolution (Fig. 45-9).

These processes occur in sequence over a 1-year duration, but they also significantly overlap and work in terms of a “continuum of processes” rather than discrete “stop-and-go” phases. As shown in Fig. 45-9, each phase is characterized by several
well-defined processes that are dominated by cellular as well as noncellular elements, such as platelets, macrophages, and cytokines, that act in concert.

**Hemostasis.** This phase of healing occurs immediately after tissue injury. The most important cells that play a role in the hemostatic process are platelets that degranulate and result in the formation of a clot. The extracellular matrix that supports the tissue framework and otherwise acts as a barrier is now open to the vascular compartment, resulting in the release of several factors into the wound. In addition, the release of proteins—otherwise stored within the extracellular matrix—and the resident cells act as further stimulants that start the hemostatic process. Inflammatory plasma proteins and leukocytes also migrate into the wound. On the cellular level, the plasma membrane of each platelet contains several receptors for collagen (glycoprotein 1A and 2A). Once these receptors are activated, glycolated granules holding multiple factors that activate hemostasis and inflammation are disrupted, releasing bioactive factors that stimulate platelet aggregation, vasoconstriction, and the subsequent activation of the clotting cascade. As these initial platelet activation factors are released, there is a subsequent push that influences angiogenesis inflammation. These systemic immune response platelet-derived factors include biologically active proteins, such as PDGF, TGF-β, and VEGF, as well as other cytokines, such as PF4 and CD40L.

In addition to the release of these factors, the binding of selected proteins within the already developed fibroblasts and the combination of two elements within the extracellular matrix create a chemotactic gradient that activates cell recruitment, cell migration, and cell differentiation and promotes tissue repair. This has been demonstrated clinically in several instances, including orthopedic surgery, cardiac surgery, and certain types of skin repair, where autologous platelet transfusions have shown to accelerate the healing process.

The subsequent fate of the platelet plug is determined by the amount of circulating fibrinogen. The vascular system interacts with the sympathetic nervous system by eliciting vasoconstriction from the actions of cytokines, prostaglandins, and catecholamines. There is also an alteration of capillary permeability caused by histaminic responses and the mediation of VEGF, which is released from micelles and the damaged endothelium. This highly interactive process results in decreasing blood loss while simultaneously delivering bioactive proteins and cells into the wound environment that kick start the inflammatory process.

**Inflammation.** This is the second phase of wound healing and arguably overlaps the hemostatic face. Polymorphonuclear leukocytes (PMNs) and macrophages appear in the wound right after platelets, and their primary role is mainly to act as scavengers. They clear the wound environment of debris, foreign material, bacteria, dead tissue cells and any other devitalized issues that would otherwise impede the healing process. Both macrophages and PMNs aid in phagocytosis and the secretion of free articles that kill bacteria and reduce the bioburden. Cellular migration into the wound is highly controlled by bioactive agents within the wound and within the vascular compartment. These include cytokines, integrins, selection, and other collagen-derived substances that act in concert. Through antibody activation, polymorphonuclear cells also interact with the humoral system to facilitate the key functions of cell activation, recruitment, and proliferation, as well as migration from the intravascular compartment to the extracellular matrix. Within 48 hours of tissue injury, PMNs and macrophages are recruited to the wound in very large numbers, heralding the inflammatory response. As described in other chapters in this text, macrophages possess a very large repertoire of functions, all of which are geared towards removing the nonviable elements in the wound and recruiting other cell types into the wound that facilitate angiogenesis, fibroblast function, and subsequent repair. A summary of various macrophage-related functions is broadly classified into 7 major categories:

1. Phagocytosis
2. Release of reactive oxygen species that result in cellular killing specifically related towards bacterial lysis
3. Release of nitric oxide that is deadly to several otherwise antibody-resistant bacteria
4. Cytokine release of interleukins (IL1, IL2, IL4, and IL12)
5. Angiogenesis via VEGF that promotes capillary budding
6. Recruitment of other cells into the wound that continue the healing process
7. Different homeostatic roles of macrophages and Langerhans cells, including wound repair, follicle regeneration, salt balance, and cancer regression and progression in the skin

Interestingly, the inflammatory phase determines the difference between chronic and acute wounds. Uncomplicated wounds heal within 4 to 6 weeks. If they continue to remain nonhealing beyond this time, they are termed chronic. Several local and systemic factors affect the inflammatory phase of wound healing directly. These include pressure, tissue hypoxia, infection, tissue contamination, desiccation, and maceration. Systemic factors include age, stress, and comorbid conditions such as diabetes, vascular insufficiency, immunocompromise, malnourishment, obesity, and smoking. The common thread, however, in all nonhealing chronic wounds is the persistence of proinflammatory conditions. These specific tissue deficits result in a chronic cycle of chronically migrating inflammatory cells (PMNs, macrophages) that scavenge early healing tissue, degrade the newly formed matrix proteins, and then cyclically recover only to restart the inflammatory phase. This cycle leads to a chronically unstable wound that is unable to progress to the next phases of healing: cell proliferation, tissue remodeling, and resolution.

**Biofilm** One of the recent discoveries in the area of biofilm is an important microbial factor that impedes healing by affecting inflammatory processes in the wound-healing continuum. Biofilm comprises a colony of microorganisms enveloped with a matrix of extracellular polymers also known as extracellular polymeric substance (EPS) (Fig. 45-10). EPS affects chronic and acute dermal wounds. Its life cycle and effects on the bacterial colonies it protects are shown in Figs. 45-11 and 45-12. These include antibiotic resistance; latency (the ability to enter into latent states during inhospitable conditions); increasing species diversity; and quorum sensing (bacteria in the biofilm engage in a type of decision-making process in which behavior is coordinated through a “chemical” vocabulary).

**Proliferation.** This phase is arguably the first step towards restoration of tissue continuity. It is characterized by the production of extracellular matrix by the fibroblast, the most prominent cell type in the proliferative phase. Fibroblasts are
the architects of wound healing and appear in the wound right at the end of the inflammatory phase. Collectively, fibroblasts support several major functions that lead to tissue repair, including the formation of collagen and the structural creation of the extracellular matrix. The formation of fibrin and fibronectin that is precipitated from the blood clot results in the formation of a provisional extracellular matrix that serves as a scaffold. Typically, this matrix can be compared to the framework of a building without any walls or windows. The protein scaffold serves as a solid framework that subsequently hosts cells including human macrophages and fibroblasts. Simultaneous VEGF-derived angiogenesis promotes the formation of small vascular loops, known as capillary buds, that proliferate within the fibroblast matrix. Paradoxically, the major activating factor responsible for the formation of capillary buds is low oxygen tension. Poor oxygenation of the tissues increases...
the expression of hypoxia inducible factor (HIF) by endothelial cells. Specific DNA sequences of cells that regulate angiogenesis are turned on by HIF. This paradoxical negative loop is directly related to a low oxygen tension within the tissues. Subsequent release of the epidermal growth factor TGFα and the transforming growth factor TGF-β by several cell types, including macrophages, platelets, and keratinocytes, strengthen the newly formed extracellular matrix. Once a robust scaffold is built, the epidermal cells from the edges of the wound on all sides migrate towards the center of the wound. This process is facilitated by several factors, including angiogenesis, neovascularization, and the release of fibroblast growth factor TGF-β and epidermal growth factor. The formation of the extracellular matrix is the key process that leads to subsequent reepithelialization. The extracellular matrix is primarily made of collagen. The different types of collagen that occur more predominantly in different types of tissues characterize the type of healing that occurs. Specifically, type I is present in scar tissues. After the formation of collagen, the fibers are now attached to form a provisional fibrin matrix. After a variety of complicated signaling that includes the transcription and processing of collagen messenger RNA, the collagen gets attached to hydroxylation of protein and lysine. The hydroxyproline in the collagen is responsible for the stable helical confirmation that is critical for the formation of a robust strong scar. It then transforms itself into a classical triple helical structure that is subsequently modified through glycosylation. It is important to realize that increased collagen stability is directly related to the degree of hydroxylation of the collagen and that fragile forms of collagen—which result in a fragile scar—are largely due to increases in nonhydroxylated collagen forms. Certain diseases including scurvy (vitamin C deficiency) or other diseases that are predominantly anaerobic in their nature can cause the formation of week nonhydroxylated collagen, which is fragile and can easily undergo denaturation and lysis.

The next step is the cleavage of the procollagen N and C terminal peptides. A very important extracellular enzyme called lysyl oxidase is responsible for the strengthening of collagen by the formation of strong, stable cross-linkages. Microscopic examination of stable mature scars reveals that strong cross-linkages present in the intramolecular and the intermolecular compartments directly correlate with strength and stability. Epidermal cells migrate over the scaffold, and after the epithelial bridge is completed, enzymes are released to dissolve the attachment at the base of the overlying scab that falls off.

Contraction is one of the key end phases of proliferation. Typically, contraction starts approximately 7 days from tissue injury, when the fibroblasts differentiate into myofibroblasts. Myofibroblasts are similar to smooth muscle cells, have the same amount of actin (responsible for mobility), and are responsible for contraction it peaks at around 10 days post injury but can continue for several weeks. Myofibroblasts attach to the extra cellular matrix (ECM) at the wound edges and to each other as well as to the wound edges via desmosomes and the fibronexus, through which actin in the myofibroblast is linked across the cell membrane to molecules in the extracellular matrix like fibronectin and collagen. This in turn facilitates the myofibroblasts to pull the ECM when they contract, thus reducing the wound size. Wounds contract at the rate of 0.75 mm to 1 mm daily. The formation of a strong, contracted, cross-linked collagen scar with reepithelialization heralds the end of the proliferative phase. Contraction usually does not occur symmetrically; instead, most wounds have an “axis of contraction” that allows for greater organization and alignment of cells with collagen.

Remodeling/Maturation. The remodeling phase is also termed the maturation phase. It is primarily characterized by the remodeling of collagen through a balance between collagen formation and collagen lysis that results in the formation of a strong scar. Biochemically, the collagen is remodeled from type III to type I and is also accompanied by complete reepithelialization of the wound. The lysis of collagen is mediated by collagenases that are secreted by various cells—fibroblasts, neutrophils, and macrophages—of which can cleave the collagen molecule at different but specific locations on all three chains and break it down to characteristic three-quarter and one-quarter pieces. These collagen fragments undergo further denaturation and digestion by other proteases. There is significant remodeling of the collagen during this process. It is aligned along tension lines, and significant reabsorption of water from the collagen fibers result in a denser alignment and stronger cross-linking. The remodeling phase establishes a new equilibrium with the formation of an organized scar. Several molecules, including TGF-β, which induces intracellular signaling of SMAD proteins, play an important role in the remodeling phase. Using SM 80 knockout mice and transgenic animals, a critical role of the SMAD pathway in the formation of scar has been delineated. This process is also facilitated by apoptosis and programmatic cell death, which helps to former a thinner scar that is stronger and more cosmetically appealing. This phase begins 3 weeks after the injury and continues for over 1 year. One must realize that despite the best cross-linking, scar tissue is weaker than injured skin and regains only 80% of its uninjured tensile strength. As it matures further, it becomes less red and less vascular because the reduced biologic activity within the scar renders the vascular capillaries redundant and they apoprose.

RECONSTRUCTIVE SURGERY

Reconstructive surgery restores normal anatomy and function using plastic surgery methods of tissue repair, rearrangement, and replacement. Tissues can be missing or damaged as a consequence of trauma, cancer, degeneration, congenital abnormalities, and aging. The primary adverse consequence of lost or impaired tissue is functional disability, which includes physical, psychologic, or social dysfunction. The clinical objective is to reestablish normal anatomy, function, and appearance in order to restore the patient as closely as possible to normal health. The most useful techniques transfer and modify tissues in the form of tissue grafts and surgical flaps.

RECONSTRUCTIVE STRATEGIES AND METHODS

The main aim of wound healing is to achieve a closed wound. Ordinarily, wounds heal via three main mechanisms:

1. **Primary intention:** This type of healing occurs in a clean wound without any apparent tissue loss. Mostly seen in surgical incisions that have been approximated (primary closure), healing by primary intention can only be implemented when the closure of the wound is precise and there is minimal disruption to the local tissue or the epithelial basement membrane. Typically, this wound seals off within 24 hours. Healing is faster than healing by secondary intention, and there is the least amount of scarring.
2. **Secondary intention**: Tissue loss following major trauma results in the formation of granulation tissue, which results in a broader scar (see earlier section, “Phases of Wound Healing”).

3. **Tertiary intention (delayed primary closure or secondary suture)**: The wound is initially cleaned, debrided, and observed, typically 4 or 5 days before closure. Examples of this type of healing include healing through the use of tissue grafts, including skin grafts and substitutes.

**Skin Grafts and Skin Substitutes**

Skin grafting methods date back millennia to ancient India, where they were used to resurface nasal defects. They were introduced in the modern era by Giuseppe Baronio, an Italian physician who studied skin grafting techniques in sheep and published his work entitled *Degli Innesti Animali (On Grafting in Animals)* in 1804.4

It is important to know the basic anatomic structure of skin in order to understand the principles of skin grafting. Skin is comprised of the epidermis, the dermis, specialized sensory nerve endings, and various skin appendages that lubricate and protect the skin as well as contribute to functions such as thermoregulation. The epidermis is a layer of cells that affords primary barrier function. It begins with a layer of cells called the basal layer. These are cuboidal-shaped cells that multiply and differentiate into flattened, keratinized squamous cells, which progressively migrate from the basal layers until they are finally released from the surface in a process known as desquamation. The junction between the dermis and the epidermis is composed of projections from the dermis into the epidermis, which are called *dermal papillae*. This feature secures the epidermis to the dermis by resisting sheer forces transmitted from the skin surface, helping to prevent separation of the epidermis from the dermis. The dermis contains sebaceous glands, whereas sweat glands and hair follicles are actually located below the dermis in the subcutaneous tissue and traverse the dermis and epithelium to reach the body surface. The dermal thickness and concentration of skin appendages vary widely from one location to another on the body. The blood supply to the skin occurs in a variety of patterns that form the basis for transferring tissue-containing skin, which will be discussed later in this chapter. Regardless of the pattern, there is a network of vessels just below the dermis called the *subdermal plexus* that supplies the skin immediately above and is important in thermoregulation. Finally, terminal vessels and capillaries fill the dermis and penetrate the dermal papillae to perfuse the cellular elements of the dermis and epidermis.

Skin grafting methods include split-thickness skin grafts (STSG), full-thickness skin grafts (FTSG), and composite tissue grafts. Each has its advantages and disadvantages, and selecting the best technique for a given circumstance depends on the reconstructive requirements, the quality of the recipient wound bed, and the availability of donor site tissue.

**Split-Thickness Grafts.** An STSG is the simplest method of tissue transfer. The name is derived from how these grafts are harvested by cutting through (i.e., splitting) the dermis at various levels. Thin STSGs are harvested through the superficial levels of the dermis. Thick grafts are harvested through deeper layers and include a larger amount of dermal tissue. The important characteristics of STSGs are determined by the thickness of dermis present in the graft. Thin grafts undergo less primary contraction after harvest because they contain fewer elements of the dermal extracellular matrix such as elastic fibers. Thick grafts undergo greater amounts of primary contraction. This is important to remember when harvesting the graft because it is necessary to obtain sufficient tissue in order to restore the defect. On the other hand, thin grafts allow the wound to undergo a greater amount of contraction in a process traditionally referred to as secondary contraction of the graft. This becomes important if the wound is adjacent to a mobile structure such as the oral commissure, which might be distorted as healing progresses. Thin grafts also have improved chances of complete engraftment, or “taking,” as they contain mostly epidermis, which has low metabolic demands, in contrast to thicker grafts that contain more dermis with greater metabolic needs.

A variety of techniques have been described to maximize the surface area that can be covered by harvested skin amount while minimizing the size of the donor site. One approach is to process the harvested skin into micrografts using devices specially designed for this purpose in the operating room. Another method is fractional skin harvesting, which involves harvesting a large number of full-thickness skin tissue columns that are then seeded onto the wound surface. The traditional method, however, is to mesh the graft. Meshed grafts usually also have enhanced reliability of engraftment because the fenestrations allow for egress of wound fluid and excellent contour matching of the wound bed by the graft. The fenestrations in meshed grafts must epithelialize by secondary intention from the surrounding graft skin. The major drawbacks of meshed grafts are poor cosmetic appearance and high rates of secondary contraction. Meshing ratios used usually range from 1:1.5 to 1:6, with higher ratios associated with magnified drawbacks related to meshing. For any case, a decision to mesh the graft must be balanced against the disadvantages. Other differences between thin and thick STSGs include final durability, pigmentation, and tendency to desiccation of the final result. The distinguishing characteristics of skin grafts types based on thickness are summarized in Fig. 45-13.

STSG donor sites heal by regeneration from dermal and epidermal elements remaining in the harvest site. Recesses between dermal papillae projecting into the dermis are lined by basal cells. These cells migrate across the wound surface and

![Figure 45-13A. Skin grafts categorized based on thickness.](image-url)
Full-Thickness Grafts. By definition, full-thickness skin grafts include the epidermis and the complete dermis. When harvesting and preparing this type of skin graft, the surgeon must carefully remove any retained subcutaneous tissue from the deep surface of the dermis in order to maximize the potential for engraftment. Full-thickness grafts are associated with the greatest amount of primary contraction, the least amount of secondary contraction, the highest durability, and ultimately the best cosmetic appearance. As a result, they are frequently used in reconstructing superficial wounds of the face and the hands. These grafts require clean, well-vascularized recipient beds free of bacterial colonization, previous irradiation, or fibrous wound tissue. They also work poorly in wounds associated with previous radiation treatments in cancer patients. The harvest site for an FTSG must be closed primarily because no skin elements remain in the area of harvest.

Skin Substitutes. Skin substitutes are typically types of extracellular matrices that are often acellular in nature and are either human-derived (allografts), animal-derived (xenografts), tissue engineered, or a combination of the three. These substitutes most often are employed to replace lost dermal and/or epidermal skin layers resulting from burns, trauma, and other superficial injuries to the outer skin layers. While a complete review of all of these commercially available materials is beyond the scope of this chapter, the benefits and applications of these useful adjuncts is growing, and they have been shown to play an important role in current as well as future reconstructive, regenerative, and restorative measures for tissue and skin replacement. Essentially, they act similarly to grafts as they rely on revascularization and autologous cell repopulation of the construct in order to “take” and become part of the lost anatomic structure they are acting to restore.

Graft Take. Skin graft healing, or “take,” occurs in three phases: imbibition, inosculation, and revascularization. Plasmatic imbibition takes place during the first 24 to 48 hours after placement of the graft onto the defect. During this time, the graft is held in place by a thin film of fibrin, and the cellular elements survive by diffusion of oxygen and substrate from plasma present in the open wound. After 48 hours, a fine vascular network forms from capillaries and small blood vessels in the wound bed and advances through the fibrin layer toward the graft. These new vascular buds encounter open, cut end vessels on the deep surface of the dermis of the graft and line up, forming loose anastomoses that begin to allow blood flow and the transfer of some nutrients and oxygen. This phase is called inosculation and is the period during which the graft is most at risk for failure. If the tenuous alignment of vessels between the wound bed and the graft are disrupted, then the final phase of healing will not occur. Events that can cause graft failure at this time include mechanical shear, formation of a seroma or hematoma, or bacterial contamination. The final phase of engraftment is called revascularization. During this phase, firmer vascular anastomoses are formed as the vessels heal, and the graft becomes perfused from the wound bed. Signs of perfusion, such as improved coloration and evidence of capillary refill, confirm engraftment and graft take. In most circumstances, these phases are complete by 4 to 5 days after graft placement. The dressing used after placing the skin graft is a critical part of success. It must prevent desiccation and shear stress from disrupting the graft, especially during the critical period of inosculation. Tie-over bolster dressings are a traditional method. Topical negative pressure wound dressings have been demonstrated to increase quantity and quality of split-thickness skin graft take compared to traditional bolster dressings. The benefits are particularly evident in wounds with irregular surface contours in areas that might be difficult to avoid motion.

After skin graft take, the graft remains subject to late failure due to mechanical shear, desiccation, or bacterial infection. Depending on the location and clinical setting, the graft should continue to be protected using dressings, topical moisturizing creams, or antibacterial medications as indicated until stable healing obtains in up to 2 weeks.

Composite Grafts. Composite grafts contain other types of tissue besides skin. Additional elements must have low metabolic requirements in order to survive the time required for revascularization. Composite grafts might include subcutaneous fat, cartilage, perichondrium, and small amounts of muscle. Indications for composite grafts are limited to small areas with specialized tissue requirements such as nasal reconstruction. For example, excision of a skin cancer involving the nasal lobule may create a composite defect that involves internal nasal lining, supporting nasal cartilage, and external skin. The ear is a good donor site for a composite graft of tissue with a good color match for the face and small amounts of tissue configured naturally to simulate the contours of the nose. For example, harvest of tissue from the root of the helix of the ear causes a relatively inconspicuous donor site. The donor site for composite tissue grafts must be repaired with primary closure.

Surgical Flaps. A surgical flap is a unit of tissue harvested from a donor site and transferred to another location for...
reconstructive purposes. The term “flap” is derived from techniques of adjacent skin tissue transfers fashioned as flaps of skin that were elevated and folded into the defect. The distinguishing feature of a surgical flap is having a blood supply independent of the injured area. A graft must go through the phases of healing described previously as it derives a new blood supply from the wound bed. A flap is brought to the wound with its own blood supply. This allows restoring tissue in areas of poor blood supply or with tissue requirements greater than what can be supported through a period of diffusion only.

There are a tremendous variety of surgical flaps that can be created depending on the individual patient’s reconstructive needs and available tissues. The challenge of reconstructive surgery is to design an appropriate flap to restore the defect with a minimal amount of morbidity related to the flap donor site. The different kinds of flaps can be broadly classified by three distinct characteristics: (a) the types of tissue contained, (b) the proximity to the defect, and (c) the pattern of blood supply.

The first way to classify different types of surgical flaps is by what tissue they contain. Nearly any type of vascularized tissue can be transferred as a surgical flap. One of the most common is a cutaneous flap, which contains skin and subcutaneous tissue. Another versatile type is a muscle flap, which contains only muscle. Musculocutaneous flaps contain a portion of muscle along with the overlying skin and all the intervening tissues. An osseous flap contains a segment of bone, and an osteocutaneous flap includes skin as well as the bone. Flaps can also be designed to include fascia and peripheral nerves. Visceral flaps contain segments of jejunum, stomach, colon, or the greater omentum. The choice of flap depends upon the reconstructive needs and availability of tissue.

The second way to classify surgical flaps is by their proximity to the defect. The location and distance between the flap donor site and the defect usually dictate the method required to transfer the tissue with preservation of the blood supply. Local flaps have a donor site located immediately adjacent to the defect. Regional flaps are harvested from the same anatomic region as the defect. Distant flaps are harvested and transferred from outside the anatomic region of the defect. During the transfer of all of these flaps, the blood supply remains attached to the source anatomic region. The tissue transmitting the blood supply is called the flap pedicle. When the blood supply is not divided during the transfer, it is referred to as a pedicled flap. If the distance between the donor site and the defect exceeds the length of the pedicle, the vessels can be divided and then reattached to uninjured vessels within or adjacent to the defect after the tissue is placed there. This technique is called a free tissue transfer, and flaps transferred in this fashion are called free flaps because for some period of time during the procedure the tissue of the flap is completely separated, or free, of the patient. The diameter of the blood vessels that supply common surgical flaps is usually less than 5 mm. Repairing blood vessels of this caliber is considered microvascular surgery, and techniques for doing this are part of reconstructive microsurgery.

The third and perhaps most important way to classify different surgical flaps is by the pattern of their blood supply. Using this criterion, flaps are traditionally divided into random pattern flaps, axial pattern flaps, musculocutaneous flaps, fasciocutaneous flaps, direct cutaneous flaps, perforator flaps, and free flaps. These designations are based on how vessels reach from the deeper, usually named, arteries and veins to the superficial tissues and skin. These are described in greater detail in the following section.

Random Pattern Flaps. The simplest flap designs are random pattern flaps, so named because the blood supply is based on unnamed vessels in the attached base of the flap that perfuse through the subdermal plexus. Random flaps are typically used to reconstruct relatively small, full-thickness defects, and they are designed following geometric principles of skin rearrangement with a traditional length-to-width ratio of 3:1. Exceptions to this principle regarding reliable dimensions abound, however, because of the variability in the patterns of perfusion and the density of the subdermal plexus in different regions of the body.

Random pattern flaps can be further subdivided based on the geometry of the transfer. Examples of this are transposition flaps, advancement flaps, and interpolated flaps. A transposition flap is fashioned adjacent to an area needing reconstruction and rotated into the defect. Large transposition flaps can require a skin graft to close the donor site. To avoid this problem, specialized types of transposition flaps have been devised. One that is particularly useful is called a Z-plasty. In this technique, two flaps are rotated, each into the donor site of the other, to rearrange the tissues in a way that redirects the lines of tension and lengthens the central limb. Another is the rhomboid (Limberg) flap (Fig. 45-14). In this technique, a skin flap is precisely designed with opposing 60° and 120° angles at the corners of a rhomboid designed immediately adjacent to the defect. This design can be modified to allow the flap to rotate into the defect.

Figure 45-14. Limberg flap.
with primary closure of the donor site with minimal distortion of the surrounding tissues as shown in the case of a gluteal repair (Fig. 45-15A–B, by complex closure; Fig. 45-15C–E, by modified Limberg flap). Modifications on the angle, including the Dufourmental modification, cause the parametric configuration to be optimized based on the defect (Fig. 45-16). Rotational flaps are a type of transposition that is semicircular in design, allowing the tissue to be rotated and permitting primary closure. Advancement flaps differ from transposition flaps because the tissue is moved forward from the donor site along the flap’s long axis rather than being rotated about a point. Two common variants include the rectangular advancement flap (Fig. 45-17) and the V-Y advancement flap (Fig. 45-18). Finally, interpolation flaps rotate about a pivot point but are used to transfer tissue

Figure 45-15. Reconstruction of a gluteal defect using complex closure and reconstruction of a gluteal defect using a modified Limberg flap.
into a nonadjacent area with an intervening portion of undamaged tissue between the donor site and the defect (Fig. 45-19).

**Axial Pattern Flaps.** Historically, surgeons made an increasing variety of surgical flaps to address a greater assortment of reconstructive problems. In the process, they noticed that some of these flaps routinely violated the strict limitations of accepted length-to-width ratio. Further investigation demonstrated that these flaps had significant arteries running parallel to the long axis of the flap. These flaps became known as axial pattern flaps. The earliest example of this type of flap is the deltopectoral flap, originally described in 1971 by Bakamjian (Fig. 45-20A,B). This flap is based on cutaneous vessels perforating from inside the chest from the internal mammary artery and vein. After entering the subcutaneous tissues, they travel obliquely from the sternal border toward the deltoid area of the arm. Long flaps can be designed based on these vessels, which can reach into the head and neck to provide thin tissue from the arm. Long flaps can be designed based on these vessels, which can reach into the head and neck to provide thin tissue from the upper chest to restore defects, especially after tumor ablation. Other important and useful axial pattern flaps are the groin flap and the posterior thigh flap.

**Musculocutaneous Flaps.** The vascular pattern of musculocutaneous flaps arises from major vessels that primarily supply a muscle and then secondarily supply the skin through multiple small vessels traversing between the superficial surface of the muscle and the subdermal plexus. The discovery of this pattern of cutaneous blood supply was a major breakthrough in reconstructive surgery because it made it possible to transfer units of tissue much larger than was possible with random or axial pattern flaps, enabling plastic surgeons to restore a greater range of deformities. Mathes and Nahai classified individual muscles into five types (I–V) according to the number and dominance of the vascular pedicles supplying each (Table 45-1).

There may be advantages to including muscle in a surgical flap besides ensuring aadequate blood supply to the overlying skin. The classic example is breast reconstruction using a latissimus dorsi myocutaneous flap (Fig. 45-21A–C). Here, the latissimus muscle is harvested pedicled on the thoracodorsal vessels and transposed anteriorly onto the chest wall. Muscle is a highly vascularized tissue that is bulky and deformable. It can help to repair visible surface contour deformities by increasing the projection of tissue in the defect to reach the level of the surrounding undamaged tissues. It can also easily contour to fill spaces in a complicated wound surface, thus helping to prevent small fluid collections in recesses, which can be a harbor bacteria and become a nidus of infection. It is also possible to provide functional restoration using musculocutaneous flaps by coaping the motor nerve of the muscle in the flap to a corresponding motor nerve in the defect. This method can be used to restore motor function in patients with motor loss in the extremities or face.

**Fasciocutaneous Flaps.** Rather than having a blood supply primarily from underlying muscle, the skin and subcutaneous tissues of some anatomic regions are supplied from vessels communicating with the underlying superficial or deep fascia. Such flaps are referred to as fasciocutaneous flaps. The artery and vein of the flap pedicle passes between rather than through muscles, form a plexus of vessels within the fascia, and then send multiple small vessels to the subdermal plexus to perfuse the skin. There are clinical circumstances when a fasciocutaneous flap might have advantages over a musculocutaneous flap. Fasciocutaneous flaps are usually thinner compared to musculocutaneous flaps. They also do not create a functional loss of muscle in the donor site. Mathes and Nahai classified fasciocutaneous flaps into types A, B, and C (Table 45-2) based on how the vascular pedicle reaches the fascia from the major vessels deep to the fascia and muscles. Sural perforator fasciocutaneous flaps (Fig. 45-22A–D) are a modern example of reconstructing lower extremity defects that would be difficult to reconstruct without microvascular surgery.

**Direct Cutaneous Flaps.** Some surgical flaps have a vascular pedicle that reaches directly to the superficial tissues and subdermal plexus without passing through a muscle or fascia plexus. These are called direct cutaneous flaps.

**Perforator Flaps.** The final kind of surgical flap classified by the pattern of blood supply is the perforator propeller flap. The geometric measurements that are critical to its success are summarized in Fig. 45-23. Reconstructive procedures based
on these flaps are the result of complementary advances in our understanding of cutaneous blood supply and improved surgical techniques.

Ian Taylor and a team of investigators from Melbourne, Australia, discovered that the blood supply to all portions of the skin was organized into discreet units, which they called angiosomes. Analogous to dermatomes that describe the patterns of cutaneous sensation supplied by single sensory nerves, the cutaneous perfusion is organized into angiosomes supplied by a single artery. These arteries arise from source blood vessels located deep to other structures like muscle and fascia and penetrate through as perforating vessels. Often the artery is accompanied by two venae committantes, but in many regions an additional venous drainage system is present in the superficial planes. The territories of adjacent angiosomes overlap similarly to how dermatomes overlap. An angiosome is defined by the limits of an artery’s terminal branching. At the borders, these arterioles form anastomoses with the neighboring angiosome. The vessels that pass between these anatomic angiosomes are called choke vessels. In life, these may open or close in response to physiologic changes in order to increase or decrease, respectively, an artery’s dynamic angiosome momentarily. Accordingly, at any given time point, the dynamic angiosome of an artery may be approximated by the volume of tissue stained by an intravascular administration of fluorescein into that artery (indicating the reach of blood flow from that artery into tissues). The potential angiosome of an artery is the volume of tissue that can be included in a flap that has undergone conditioning (see the following section). Both the dynamic and potential angiosomes extend beyond the anatomic angiosome of an artery. Although the angiosome concept provides some guidance to the size and volume limits of a flap harvest, there remains no quantifiable method to predict safe flap harvest limits with precision.
A

Figure 45-20A, B. Deltopectoral flap for cheek reconstruction.

B

<table>
<thead>
<tr>
<th>Table 45-1</th>
<th>Mathes-Nahai classification of muscular flaps</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLASSIFICATION</td>
<td>VASCULAR SUPPLY</td>
</tr>
<tr>
<td>Type I</td>
<td>One vascular pedicle</td>
</tr>
<tr>
<td>Type II</td>
<td>Dominant and minor pedicles (the flap cannot survive based only on the minor pedicles)</td>
</tr>
<tr>
<td>Type III</td>
<td>Two dominant pedicles</td>
</tr>
<tr>
<td>Type IV</td>
<td>Segmental pedicles</td>
</tr>
<tr>
<td>Type V</td>
<td>One dominant pedicle with secondary segmental pedicles (the flap can survive based only on the secondary pedicles)</td>
</tr>
</tbody>
</table>

Tissue Expansion. Tissue expansion is a technique that increases the amount of tissue in a surgical flap by first placing an inflatable device into the tissue beneath the planned flap and gradually expanding the tissue by regular inflation. Staged reconstruction using tissue expansion can significantly increase the amount of local, well-matched tissue for transfer while decreasing donor site morbidity. The most common method of skin expansion involves the placement of an inflatable silicon elastomer similar to a balloon with a filling port that is generally positioned in an easily accessible location beneath the skin. After wound healing, the device is gradually inflated by serial injections of sterile saline solution into the filling port. The process can require several weeks, depending on the amount of expansion and compliance of the tissues. When expansion is complete, the expander is removed, and the resulting expanded tissue is transferred into the defect.

The process of expanding flaps confers physiologic benefits that increase the reliability of the flap tissue. Histologically, expanded skin demonstrates thickened dermis with enhanced vasculature and diminished subcutaneous fat. Studies have shown that the increased amount of skin is the result of actual generation of new tissue. Also, the blood supply to an expanded flap is improved because of the period of delay associated with expansion process and the capsule formed around the device is highly vascular and contributes to the quality of blood supply.16

The disadvantages of tissue expansion have to do with possible complications, which include infection, hematoma, seroma, expander extrusion, implant failure, skin necrosis, pain, and peripheral nerve injury. Furthermore, an inflated expander is visible, and the temporary deformity may cause patients distress.

Tissue expansion has found particular usefulness in managing giant congenital nevi, secondary reconstruction of extensive burn scars, scalp reconstruction, and breast reconstruction. Expanders are available in a multitude of shapes and sizes, depending on the reconstructive needs. The technique permits reconstruction with tissue of similar color, texture, and thickness, with minimal donor site morbidity.

PEDIATRIC PLASTIC SURGERY

Congenital Craniofacial Anomalies

In 1981, Whitaker et al introduced a simple classification system to help conceptualize the vast array of congenital pathology involving the craniofacial region.17 Based on anatomy, etiology, and current treatment principles, most craniofacial anomalies can be classified into one of four categories: clefts, synostoses, atrophy-hypoplasia, or hypertrophy-hyperplasia-neoplasia (Table 45-3).

Clefts. Arguably, no operation in plastic surgery is more demanding of reconstructive principle and aesthetic intuition
than a cleft lip repair. Orofacial clefting is the most common birth defect in the world. Cleft lip, with or without cleft palate (CL/P), occurs spontaneously among Caucasian populations in approximately 1 out of every 1000 births. It is over twice as common (1 in 450) among Asians and Native Americans and half as common (1 in 2000) in African Americans. There is a predilection among males, who are twice as likely to be affected as females. Left-sided cleft lip is twice as common as right and nine times as common as bilateral. Of patients born with CL/P, 29% have associated anomalies, which can range from minor physical differences to major organ involvement. While a family history of CL/P remains the strongest known predictive factor, other extrinsic risk factors include maternal smoking or early exposure to the anticonvulsant drug phenytoin.18

Epidemiologically, isolated cleft palate (CP) appears to be distinctly different from CL/P. CP occurs in 1 of every 2000 live births. It is twice as common in females, and it demonstrates no racial or ethnic preponderance. Nearly half of patients with isolated CP have a diagnosable syndrome and additional congenital anomalies. Evaluation by a geneticist is therefore indicated in all babies born with isolated CP. Like CL/P, isolated CP is multifactorial. Known environmental risk factors include maternal smoking or alcohol consumption, folate deficiency, use of steroids or anticonvulsant medications, or retinoid (vitamin A) excess.

Some familial patterns of orofacial clefting have been linked to specific genetic mutations. Van der Woude syndrome, an autosomal dominant form of CL/P associated with lower lip pits, is caused by an IRF6 gene mutation (Fig. 45-24).23 Stickler syndrome should be suspected in patients with isolated CP.
Figure 45-23. Reconstruction of a lateral malleolar defect using a reverse sural perforator flap.

Figure 45-22. Geometric considerations for a propeller flap.

with associated eye defects, sensorineural hearing loss, and joint abnormalities. This constellation of findings is due to an autosomal dominant mutation in a procollagen gene. Stickler is also the most common syndrome associated with Pierre Robin sequence (micrognathia, glossoptosis, and respiratory distress). These examples help emphasize the importance of early genetic workup for patients in whom a syndrome is suspected.

Embryology of the Lip and Palate The “primary palate,” which includes the nostril sill, upper lip, alveolus, and hard palate anterior to the incisive foramen, forms from fusion between the medial nasal and maxillary prominences during weeks 4 through 7 of gestation. Development of the hard palate posterior to the incisive foramen and the soft palate, which are collectively known as the “secondary palate,” occurs during weeks
6 through 12 of gestation. The lateral palatine processes initially hang vertically on either side of the developing tongue. Around week 8, these palatal shelves rotate into a horizontal orientation, bringing their free edges into close proximity with the nasal septum. Midline fusion then commences, proceeding posteriorly from the incisive foramen (Fig. 45-25).23

**Normal and Cleft Anatomy** There are several key defining characteristics of the lip that make its surgical repair so challenging. On the surface, the philtrum of the upper lip is comprised of paired philtral columns and a central philtral dimple. The white roll passes along the vermilion-cutaneous junction, peaking at the base of the philtral columns and dipping centrally to form Cupid’s bow. Deep to the surface, the paired orbicularis oris muscles originate lateral to the oral commissures and encircle the mouth, decussating in the midline and sending off dermal insertions to the philtrum. This intrinsic muscle of the lip provides oral competence and assists with speech production and facial expression. Continuity of the orbicularis oris muscle is disrupted in babies born with a cleft lip. Aberrant muscle insertion into the piriform aperture laterally and the anterior nasal spine medially contributes to the hallmark appearance of cleft lip and nasal deformity (Fig. 45-26).20,25

Clefts of the lip can be described as unilateral or bilateral and microform, incomplete, or complete. Microform cleft lip is the most minor variant and may manifest as subtly as a small notch in the vermillion. An incomplete cleft lip, by definition, requires an intact nasal sill. The term can otherwise be applied to a wide spectrum of anomaly, ranging from a partial cleft of the lip alone (Fig. 45-27A) to a near-complete cleft of the entire primary palate. A complete cleft lip involves all structures of the primary palate in their entirety, extending through the nasal sill and opening into the anterior nasal floor (Fig. 45-27B).20,26

The normal palate functions primarily as a speech organ, but it is also intimately involved in feeding, swallowing, and breathing. The soft palate, or velum, together with lateral and posterior pharyngeal walls, can be conceptualized as a valve that regulates the passage of air through the nasopharynx. The paired levator veli palatini muscles descend from the cranial base and decussate in the midline to form a sling within the soft palate. This sling acts to elevate the velum against the posterior pharyngeal wall, effectively closing the velopharyngeal port. In patients with cleft palate, the levator muscles are unable to cross the midline. Instead, they run parallel to the cleft margin and insert aberrantly into the posterior edge of the hard palate (Fig. 45-28A,B). Air is allowed to leak through the nose during attempts to suck or speak. This inability to build negative or positive intraoral pressure makes either task difficult, if not impossible. The tensor veli palatini muscles, which normally function to vent and drain the Eustachian tubes, are also disrupted in cleft anatomy. Eustachian tube dysfunction predisposes patients to frequent bouts of otitis media, which can lead to permanent hearing loss if left untreated.20

The most clinically useful system to describe cleft palate morphology is the Veau classification. A Veau I cleft is midline and limited to the soft palate alone, whereas a Veau II cleft may extend further anteriorly to involve the midline of the posterior hard palate (the “secondary palate”). A Veau III cleft is a complete unilateral cleft of primary and secondary palates, in which the cleft extends through the lip, the alveolus, the entire length of the nasal floor on the cleft side, and the midline of the soft palate. Veau IV clefts are bilateral complete clefts of the primary palate that converge at the incisive foramen and continue posteriorly through the entire secondary palate (Fig. 45-29A,B). Not included in the Veau classification is the submucous cleft palate, which occurs when there is clefting of the soft palate musculature beneath intact mucosa. Submucous cleft palate classically presents as the triad of a bifid uvula, a midline translucency called the “zona Pellucida” and a palpable notch of the posterior hard palate.21

**Presurgical Infant Orthopedics** Current literature suggests aesthetic outcomes in patients with complete unilateral or bilateral clefts may be improved by reestablishing more normal skeletal, cartilaginous, and soft tissue relationships prior to definitive lip repair. Presurgical infant orthopedics (PSIO) can help to narrow wide clefts and align dental arches in preparation for surgery. Some methods of PSIO, such as nasoalveolar molding (NAM), provide the added benefits of elongating the columella and improving nasal tip asymmetry.22 The most common barrier to PSIO implementation is its imposition on families, who must be willing and able to keep frequent follow-up appointments for appliance adjustment. An excellent alternative to PSIO is a lip adhesion procedure, in which a complete cleft is surgically converted to an incomplete cleft. This preliminary stage of lip repair restores soft tissue continuity at the nasal sill, which helps to realign the underlying dental arches and approximate the soft tissues. In addition, the nasal deformity can be improved, both by repositioning of the cleft side alar base and placement of nasal conformers.23

**Cleft Lip Repair** Although cleft lip surgery can be traced to antiquity, it was not until the first half of the 20th century that surgeons began to realize the inadequacy of a straight-line repair. In 1955, Ralph Millard pioneered his “rotation-advancement” technique, which was the first to address upper lip length deficiency while preserving intricate philtral anatomy (Fig. 45-29C).24 The back-cut is designed high on the medial lip element just beneath the columella, enabling a downward rotation and leveling of Cupid’s bow, while the lateral lip element is advanced into the

---

Table 45-3

**Classification of craniofacial anomalies**

1. Clefts
2. Synostoses
3. Atrophy–hypoplasia
4. Hypertrophy–hyperplasia–neoplasia

---

**Figure 45-24.** Van der Woude syndrome.
Figure 45-25. Facial prominences and their contributions to facial development. Cleft lip results from failure of fusion between maxillary and medial nasal (a component of frontonasal) prominences.

rotation defect. Although other techniques exist, most lip repairs performed today are minor modifications of Millard’s original rotation-advancement principle.20

Bilateral cleft lip presents an even greater set of challenges to the reconstructive surgeon. With no overlying orbicularis oris muscle, an unrestrained premaxilla rotates anteriorly, completely displacing the incisor-bearing portion of the alveolus from the maxillary dental arch. Orbicularis continuity must be restored over an often protuberant premaxilla. The surgeon must carefully recreate the appearance of a symmetrical philtrum and median labial tubercle. Prototypical markings for bilateral cleft lip repair are demonstrated in Fig. 45-30A,B.20

Any surgical approach to bilateral cleft lip repair would be incomplete without addressing the nasal stigmata, which include a short or absent columella, a poorly defined and underprojected nasal tip, and malpositioned lower lateral cartilages.25 Primary nasoplasty at the time of lip repair has become an increasingly common practice. Nasal skin and soft tissue are dissected free from the underlying cartilaginous framework, allowing for suture manipulation of lower lateral cartilages to improve tip symmetry, support, and projection.20

Cleft Palate Repair The primary goal of palatoplasty is to enable normal speech development. A successful palate repair is one that results in a robust, layered reconstruction of the cleft and restoration of functional velar anatomy. The two most common techniques employed for soft palate repair are intravelar veloplasty (IVV) and Furlow double-opposing Z-plasty. Paramount to each technique is the complete release of aberrant levator muscle insertions from the posterior edge of the hard palate. This maneuver untethers the velum anteriorly, enabling maximal levator muscle excursion in the superior and posterior directions postoperatively.21
SPECIFIC CONSIDERATIONS

PART II

Figure 45-26. Hallmarks of unilateral cleft lip deformity include depression of the nasal tip and flaring of the alar base on the cleft side, deviation of the caudal septum and columella toward the non-cleft side, and deficient lip height (short philtral column) on the cleft side with cephalad rotation of the cleft side of cupid’s bow.

Intravelar veloplasty requires meticulous dissection of the levator muscles with retropositioning and reconstruction of the sling mechanism in the posterior aspect of the soft palate. A Furlow double-opposing Z-plasty involves cleverly designed mirror image Z-plasties on the oral and nasal sides of the soft palate where the central limb of each Z-plasty is the cleft. The posteriorly based flap of mucosa on each surface of the palate incorporates the underlying levator muscle. Transposition of these flaps across the cleft lengthens the palate and, in a manner similar to IVV, corrects levator malposition. Lateral relaxing incisions can be utilized to relieve tension on the closure, if necessary (Fig. 45-31A–C). In experienced hands, both techniques have demonstrated excellent speech outcomes and low fistula rates. However, direct comparison between the two methods has been difficult due to ongoing evolution of the IVV technique and wide variability in the extent of dissection between performing surgeons.

Clefts involving the hard palate (Veau II–IV) often require additional maneuvers for reconstruction. Wide undermining of the nasal floor mucosa in the subperiosteal plane facilitates the nasal-side repair. As palatal mucoperiosteum is thicker and less pliable, the oral-side closure generally requires the use of relaxing incisions along the lingual side of the alveolar ridge. Additional medialization of the palatal soft tissue can be obtained by increasing isolation of the greater palatine neurovascular pedicle, which emerges from its foramen near the posterolateral aspect of the hard palate. Narrow Veau II clefts may be closed on the oral side by medialization of bilateral bipedicled mucoperiosteal flaps (von Langenbeck palatoplasty), while wider clefts may require detachment of one or both flaps anteriorly for additional medialization (Bardach two-flap palatoplasty). Lateral relaxing incisions are left open, and typically heal by secondary intention within two weeks (Fig. 45-32).

Complications of palate repair include oronasal fistula, velopharyngeal dysfunction, obstructive sleep apnea, and mid-face growth deficiency. Reported fistula rates vary widely in the literature, but increased incidence has been correlated with less experienced surgeons, wider clefts, and bilateral clefts. Few oronasal fistulae are amenable to closure with simple local tissue rearrangement. More commonly, a complete reelevation of palatal mucosa is required in order to obtain a tension-free layered closure. In the case of large or recurrent fistulae, there may be insufficient tissue available locally, and recruitment of regional healthy tissue from the buccal mucosa or tongue may be necessary.

Velopharyngeal dysfunction (VPD) is caused by incomplete closure of the velopharyngeal port, which results in air leaking through the nose during speech. Approximately 20% of patients develop VPD after primary palatoplasty. After ensuring complete release and proper orientation of levator muscles, a posterior pharyngeal flap or a sphincter pharyngoplasty may be required to decrease the size of the velopharyngeal gap, allowing

Figure 45-27. Variations in unilateral cleft lip morphology. Left unilateral incomplete cleft lip.
nasal air escape during speech. These operations carry a risk of obstructive sleep apnea, so preoperative polysomnography is indicated to rule out significant sleep-disordered breathing at baseline.

**Timeline for Repair** The longstanding debate regarding optimal timing for lip and palate repair is ongoing. Central to this controversy is the impact of early surgical intervention on speech outcomes and midface growth. Current evidence suggests earlier palate repair is better for speech but more detrimental to midface growth. Cleft care algorithms represent a compromise. Most experts perform lip repair between 3 and 6 months of age. Palate repair should be completed prior to the onset of speech development, usually around 10 to 12 months of age. The alveolar cleft is often repaired secondarily with a cancellous bone graft from the iliac crest. This operation provides bony support for the permanent teeth that will erupt adjacent to the cleft, and it is usually performed around 7 to 9 years of age. Orthognathic surgery and secondary rhinoplasty, if necessary, are delayed until skeletal maturity. The treatment timeline used at Nationwide Children’s Hospital can be seen in Fig. 45-33.
Figure 45-30. A. Bilateral cleft lip repair diagram. B. Bilateral cleft lip repair.

Figure 45-31. Furlow double opposing Z-plasty. A. Oral side markings. B. Nasal side markings. Note that the levator veli palatini muscle remains attached to the posteriorly based flap on each surface. C. Flap transposition and closure. The levator veli palatini muscle bundles, being attached to the posteriorly based flaps, are reoriented transversely and retrodisplaced as a result of flap transposition.
The Importance of Team in Cleft Care  

Children born with CL/P require expertise of medical professionals from many different disciplines. In addition to experienced craniofacial surgeons, cleft teams typically consist of otolaryngologists, pediatricians, speech pathologists, feeding specialists, pediatric dentists, orthodontists, geneticists, psychologists, nurses, and social workers. Each member is an integral part of the team and absolutely essential for the delivery of comprehensive cleft care.21

Atypical Craniofacial Clefts  

Beyond the familiar scope of clefts confined to the lip and palate, there exist myriad forms of clefting that may affect the craniofacial skeleton. Sound epidemiologic studies of these atypical craniofacial clefts have been precluded by their extreme rarity, but rough estimates place them on the order of 100 times less common than CL/P. As a result, definitive causality has not been established. With the exception of some well-defined syndromes that include atypical craniofacial clefts, genetics does not appear to play a significant part in their pathogenesis. Some extrinsic factors that have been implicated include radiation, prenatal infections, early gestational exposure to teratogenic drugs or chemicals, and amniotic bands. Metabolic derangements and vascular disturbances have also been hypothesized to play a role.27

While CL/P can be logically explained as an embryologic failure of fusion between facial processes, the location of the atypical craniofacial clefts is not well-accounted for by this theory. In the 1960s, Weston and Johnston used animal models to demonstrate the vast contributions of neural crest cells to mesenchymal development of the face. They postulated that failure of these cells to penetrate into the developing face could lead to breakdown of the surrounding epithelia and result in atypical craniofacial clefts. The last 30 years has seen continued refinement of this theory. Most recent evidence suggests that neural crest cells form developmental rests or ossification centers within the well-known facial processes. An abnormal number or impaired differentiation of these ossification centers may better explain the locations of clefts that seem to follow no known embryologic fusion plane.33

In 1974, Paul Tessier published detailed anatomic observations of a large series of children with atypical craniofacial clefts. He introduced a simple numbering system to classify these clefts based strictly on involved anatomy.28 Clefts were assigned numbers 0 to 14 as they radiate around the orbit. Numbers 0 to 7 describe facial clefts, while 8 to 14 described cranial clefts. Fig. 45-34 illustrates the paths of soft tissue clefts (above) and their corresponding skeletal clefts (below).33,35

A number 0 facial cleft and its number 14 cranial extension are midline clefts, which may be characterized by tissue deficiency or excess. Holoprosencephaly, a term used to describe a
failed cleavage of the prosencephalon into two separate cerebral hemispheres, presents as a midline tissue deficiency that causes variable degrees of hypotelorism and upper lip and nasal deformity. Mildly affected patients may have near-normal intelligence, while severely affected cases are incompatible with life. Representing the opposite end of the spectrum, patients with median cleft face dysmorphism typically present with a median clefts of the lip and/or premaxilla midline tissue excess, hypertelorism, bifid cranium, and a normal underlying CNS (Fig. 45-35A,B).33

Tessier clefts 1, 2, and 3 originate at the cupids bow. All proceed cephalad through the piriform aperture and affect the nose. While number 1 and 2 clefts spare the orbit, number 3 clefts create continuity between the orbit, maxillary sinus, nasal and oral cavities. Clefts 4, 5, and 6 begin lateral to cupids bow, spare the nose, and pass cephalad to affect the orbit and lower eyelid. The number 7 cleft, otherwise known as craniofacial microsomia, extends transversely along a line from the oral commissure to the auricular tragus. Underlying skeletal clefts can involve the mandible, maxilla, orbit, and cranium. Tessier clefts 8 through 10 continue to radiate laterally and superiorly around the orbit. Cranial extensions are numbered such that the sum of the facial cleft and its corresponding cranial extension is always 14. For example, the number 1 facial cleft continues as the number 13 cranial cleft, and the number 5 facial cleft continues as the number 9 cranial cleft.33,35 Clefts can be unilateral or bilateral and may occur in any combination. The constellation of bilateral Tessier clefts 6, 7, and 8 has been well-described within the context of Treacher Collins syndrome, in which patients exhibit malar hypoplasia, lower eyelid colobomas, and downward-slanting palpebral fissures (Fig. 45-36A–C).33

Treatment of atypical craniofacial clefts varies widely with each unique patient. Classical approaches to surgical management involved excision of atrophic soft tissue along cleft margins with reconstruction by local tissue rearrangement, with or without underlying bone grafting. Unfortunately, this methodology gives little consideration to the aesthetic units of the face, and the resulting scars often cause postoperative deformities of their own. Ortiz-Monasterio and Taylor proposed a new treatment philosophy based on the following tenants:

1. Restoration of the craniofacial skeleton
2. Reconstruction with skin and soft tissue with like color and texture
3. Generous use of tissue expanders
4. Aesthetic unit and subunit reconstruction
5. Scar location at limits of aesthetic subunits
6. Symmetrical repositioning of key facial landmarks

Fig. 45-37 demonstrates the dramatic improvement in aesthetic outcome that is attainable when abiding by this treatment philosophy.29

**Figure 45-35.** Tessier 0-14 clefts. **A.** Holoprosencephaly. Note the midline tissue deficiency, hypotelorism, and the rudimentary nose known as a “proboscis.” The degree of facial deformity in patients with holoprosencephaly typically reflects the degree to which the underlying CNS is affected. **B.** Median cleft face dysmorphism. Note the marked midline tissue excess and hypertelorism. Although this patient exhibits an obvious encephalocele, CNS function is usually normal.
Barring immediate danger to vital structures such as the eye, the timing of reconstruction can be determined on a case-by-case basis. Soft tissue clefts can be excised and closed by classical measures within the first year of life. However, bony reconstruction should be delayed until at least 5 to 6 years of age to minimize iatrogenic impairment of facial growth. Serial tissue expansion of the cheek prior to this time may be necessary to excise unfavorable scars and reorient them along aesthetic subunit boundaries. Preoperative imaging, such as computed tomography (CT) or magnetic resonance imaging (MRI), is necessary to fully characterize the defects and plan the operation. Additional preoperative workup should include anesthesia evaluation and labs, as these operations can be lengthy and accompanied by significant blood loss. Preparedness for blood transfusion is imperative.33,34

Craniofacial clefts are typically approached through a combination of bicoronal and oral vestibular incisions. Various osteotomies have been described to reposition components of...
the craniofacial skeleton, such as the orbits, maxilla, and mandible. These may be used in conjunction with bone grafts from the calvarium, ribs or iliac crest, and fixation can be achieved with standard techniques using bioresorbable plates or sutures.  

**Craniosynostosis.** The term “craniosynostosis” refers to premature fusion of one or more calvarial sutures. It occurs in up to 1 out of every 2000 live births, and single-suture, nonsyndromic patients account for 85% of cases. Of these, isolated sagittal craniosynostosis is the most common form, while lambdoidal is the least common. Normal suture maintenance is driven by underlying brain growth and a complex biochemical interplay between the suture and the underlying dura mater. Multiple genes have been implicated in the development of craniosynostosis, the most notable of which being **FGFR** and **TWIST.** Fifty percent of these present as de novo mutations, and most exhibit an autosomal dominant inheritance pattern. Environmental associations, such as maternal smoking, have been postulated, but definitive causality has not been proven.  

According to Virchow’s law, patients with craniosynostosis exhibit a predictable pattern of deformity that results from an arrest of cranial growth perpendicular to the prematurely fused suture, with a compensatory increase in growth parallel to the affected suture (Fig. 45-38). Isolated sagittal craniosynostosis, for example, results in restricted cranial growth in the transverse direction and a compensatory increase in the anterior-posterior diameter of the head with frontal and/or occipital bossing. This head shape is commonly referred to as “scaphocephaly.” Fig. 45-39 depicts various other isolated craniosynostoses and the patterns of deformity that ensue.  

All patients with craniosynostosis should be screened for intracranial hypertension. It has been estimated that up to 17% of patients with single-suture involvement may develop elevated intracranial pressure (ICP). This risk approaches 50% in patients with multisuture craniosynostosis. Signs and symptoms of increased ICP may include headache, incoordination, nausea, vomiting, lethargy, sleep apnea, developmental delay, bulging fontanelles, hydrocephalus, papilledema, or loss of vision. Facial dysmorphism and a strong family history should raise suspicion for syndromic etiology, as seen in Apert, Crouzon, Pfeiffer, and Saethre-Chotzen syndromes, among others.  

Diagnosis of craniosynostosis begins with physical exam. A recent prospective multicenter study suggests 98% accuracy of diagnosis based upon physical exam findings alone. Palpable ridges may be present on the cranium but are not pathognomonic for craniosynostosis. The much more reliable physical exam finding involves recognition of the distinct pattern of cranial growth that result from premature fusion of one or more sutures. Dysmorphic facies, suspicion for multisuture involvement, or any degree of uncertainty in the diagnosis can be clarified with adjunctive imaging. While skull plain films can provide useful information, 3D computed tomography has emerged as the new gold standard imaging modality for diagnosing craniosynostosis.  

The goals of treatment for craniosynostosis are to achieve a more normalized head shape and to treat or prevent negative impacts on development that may result from increased ICP. In general, two approaches exist: (a) strip craniectomy procedures and (b) remodeling procedures. Simply put, strip craniectomy procedures remove the synostotic suture in order to disinhibit cranial growth across the affected suture. Adjunctive techniques, such as cranial spring or distractor placement versus postoperative helmet therapy are frequently combined with strip craniectomies to improve aesthetic outcomes. Many surgeons who perform these procedures will do so as early as
6 to 12 weeks of life to take advantage of early rapid brain growth, which helps drive cranial expansion after release of the synostotic suture. In addition, younger patients have a better capacity to heal the resulting cranial defects due to the high osteogenicity of the underlying dura, which decreases substantially with age. Remodeling procedures go further to normalize head shape by complete removal, rearrangement, and replacement of abnormal areas of the calvarium. Given the limited efficacy of the aforementioned strip craniectomy techniques in patients older than 6 months of age, cranial vault remodeling is generally accepted as the definitive treatment for craniosynostosis in this age group.

Advantages of strip craniectomy procedures include shorter operative times, less blood loss, and shorter hospital stays, while disadvantages include an inability to treat complex deformities from multisuture involvement, inability to treat areas of compensatory increased cranial growth, and the necessity for secondary hardware removal procedures. Remodeling procedures offer a more definitive correction of head shape in a single surgical procedure at the cost of increased operative times, higher rate of blood transfusions, and increased length of hospital stays.

The complexity of patients with syndromic craniosynostoses, such as Crouzon or Apert syndrome, mandates multidisciplinary care from an experienced team of subspecialists. These patients may present with urgent airway obstruction, dangerously elevated ICP, and/or vision-threatening globe protrusion (Fig. 45-40A–C). Early surgical interventions, such as strip craniectomy or posterior cranial vault distraction, are designed to increase cranial volume and therefore decrease ICP. Although optimal timing of definitive reconstruction is debatable, results of cranial vault remodeling and midface advancement surgeries appear more stable and demonstrate less relapse when delayed. Hearing, speech, and feeding difficulties are common among patients with syndromic craniosynostoses. As always, the psychosocial implications of such profound facial differences make social workers and psychologists indispensable members of the team.

**Atrophy and Hypoplasia.** Two conditions that exemplify the atrophy and hypoplasia class of craniofacial anomalies are progressive hemifacial atrophy and Robin sequence. Progressive hemifacial atrophy, otherwise known as Parry-Romberg syndrome, is a rare, acquired, idiopathic atrophy of the skin, subcutaneous tissue, muscle, and occasionally bone affecting one side of the face (Fig. 45-41). With a typical onset during the first or second decade of life, this self-limiting condition progresses with an indolent course for 2 to 10 years before stabilizing, or “burning out.” The pathogenesis of Parry-Romberg syndrome is not well understood. Autoimmune processes such as scleroderma, chronic neurotropic viral infections, trigeminal neuritis, intracerebral vascular malformations, and increased sympathetic nerve activity have all been postulated to play a role. After progression of atrophy ceases, the mainstay of treatment is volume and contour restoration with autologous fat grafting. More severe cases may require microvascular transfer of free tissue, such as the parascapular fasciocutaneous flap.

Robin sequence is defined as the triad of micrognathia, glossoptosis, and airway obstruction (Fig. 45-42). Cleft palate is present in up to 90% of affected patients, though it is not an obligatory component of the diagnosis. The cause of this condition is not known, but many believe mandibular hypoplasia to be the inciting event. According to this theory, micrognathia (small jaw) prevents forward migration of the tongue during gestational development. Glossoptosis results, where the tongue remains flipped dorsally into an obstructive position within the oropharyngeal airway. The first step in management is prone positioning, which utilizes gravity to bring the mandible and tongue base forward and alleviate the upper airway obstruction. More severely affected babies may require emergent endotracheal intubation at the time of delivery in order to secure the airway.

A diagnosable syndrome can be expected in over 50% of patients born with Robin sequence. Stickler syndrome (congenital ocular, orofacial, auditory, and articular anomalies), which is the leading cause of childhood blindness due to retinal detachment, is the most commonly associated syndrome. For this reason, ophthalmology and genetics evaluations are indicated in all patients with Robin sequence. Additionally, a thorough airway evaluation by an otolaryngologist is necessary to confirm obstruction at the level of the tongue base and to rule out intrinsic airway anomalies or obstruction at lower levels of the respiratory tract.

Babies who are mildly affected can often be managed nonsurgically with prone positioning alone. Close monitoring is required because obstructive symptoms do not always follow a linear course to resolution. High caloric expenditure on
Figure 45-40. A and B. Frontal and lateral views of a young girl affected by Crouzon syndrome. Brachycephaly is appreciable on the lateral view, which results from bicoronal craniosynostosis. This patient also exhibits exorbitism and significant midface hypoplasia. C. A patient with Crouzon syndrome whose severe exorbitism has led to exposure keratitis.
increased work of breathing, in combination with reflux and feeding difficulties that are ubiquitous in this population, may manifest as poor weight gain over time. Persistent failure to thrive indicates a failure of conservative management.\(^4\)

Robin sequence patients with single-level obstruction at the tongue base who have failed conservative measures should be considered for surgical airway management.\(^4\) Tongue-lip adhesion (TLA) is designed to bring the tongue base forward and out of the airway by temporarily sewing the undersurface of the tongue to the mucosal surface of the lower lip. Adhesions are typically reversed within the first year of life as significant mandibular growth and improved muscle tone of the tongue result in a stable airway.\(^3\)

Another option to treat upper airway obstruction in patients with Robin sequence is mandibular distraction osteogenesis (MDO). In this procedure, osteotomies are made in bilateral mandibular rami, and distractor devices are applied that enable a gradual (1–2 mm/day) lengthening of the mandible. As the mandible is brought forward, the tongue base follows, resulting in enlargement of the oropharyngeal airway. Specific risks include injury to tooth buds, inferior alveolar or marginal mandibular nerves, and disruption of mandibular growth potential.\(^4\)

In Robin sequence, patients who fail or are not candidates for less invasive surgical maneuvers, tracheostomy remains the definitive option for airway control. Figure 45-43 represents an algorithm for management of children with Robin sequence proposed on the basis that TLA is less invasive and does not preclude subsequent MDO in the event of failure.\(^4\) However, one option has not been proven to be significantly better than the other, and many surgeons prefer MDO as a first-line intervention.

**Hypertrophy, Hyperplasia, and Neoplasia.** Numerous hypertrophic, hyperplastic, or neoplastic processes can affect the craniofacial region. The presence of certain vascular anomalies in the face can result in hypertrophy of surrounding bone or soft tissue.\(^19\) Patients with neurofibromatosis-1 may similarly present with hemifacial hypertrophy related to the presence of an underlying plexiform neurofibroma.\(^36\) Fibrous dysplasia is a focal error in osteoblast differentiation that leads to replacement of normal bone with a disorganized mass of bony trabeculae and fibrous tissue. Seventy percent of lesions are monostotic, and

![Figure 45-41. Child with progressive hemifacial atrophy, otherwise known as Parry-Romberg syndrome.](image)

![Figure 45-42. An infant with Robin sequence. Marked micrognathia and glossoptosis cause respiratory distress due to upper airway obstruction at the level of the tongue base. Note the presence of sternal retraction during inspiration.](image)

![Figure 45-43. Algorithm for management of children with Robin sequence.](image)
the remaining 30% are polyostotic. In the craniofacial region, fibrous dysplasia typically presents in childhood with pain and progressive asymmetry. Patients with McCune-Albright syndrome have polyostotic fibrous dysplasia, café au lait spots, and hyperfunctioning endocrinopathies, which classically manifest as precocious puberty. Lesions have a distinct “ground glass” appearance on CT scan. Small, monostotic fibrous dysplasia lesions can occasionally be resected completely and reconstructed with bone grafts. More commonly, surgical debulking and contouring is the treatment of choice.37

Vascular Anomalies. Vascular anomalies affect approximately 5.5% of the population. They can be broadly categorized as either tumors or malformations.38 Vascular tumors are characterized histologically by endothelial cell proliferation, with or without luminal structure. In contrast, vascular malformations are collections of abnormally developed vessels without significant endothelial cell turnover.39

Hemangiomas Hemangiomas are the most common vascular tumor in children, presenting in up to 20% of premature infants. Females are four times as likely to be affected as males, and darker-skinned individuals are rarely affected. These benign tumors are believed to be collections of primitive blood vessels formed from angioblasts. Hemangiomas can occur anywhere throughout the body, with the liver being the most common extracutaneous site.46

The natural history of hemangiomas is highly predictable depending on the timing of presentation and early clinical course. Infantile hemangiomas appear shortly after birth, usually between 2 weeks and 2 months of life. Cutaneous infantile hemangiomas may initially resemble a red scratch or bruise, while subcutaneous or visceral lesions go unnoticed. Rapid growth ensues over the next 9 to 12 months (“the proliferative phase”). During this time, cutaneous lesions become bright red and tense, while subcutaneous lesions may present as deep soft tissue masses with a bluish/purplish hue. After plateau of the proliferative phase, infantile hemangiomas reliably undergo a slow regression (“involution”), which is usually complete by 4 years of age. History alone can help differentiate a congenital hemangioma, which is fully formed at birth, from an infantile one. Congenital hemangiomas may exhibit rapidly involuting (RICH), noninvoluting (NICHI), or partially involuting (PICH) clinical courses. History and physical is often sufficient to diagnose a hemangioma. Doppler ultrasound has become the imaging modality of choice, while MRI is typically reserved to confirm the diagnosis in cases of uncertainty.40

Most hemangiomas can be observed and allowed to involute spontaneously. High-risk lesions that may require early intervention include ulcerated and bleeding hemangiomas; periocular hemangiomas, which can occlude the visual axis and lead to blindness; hemangiomas in the beard distribution, which place the patient at risk for upper airway obstruction (Fig. 45-44); and posterior midline lumbosacral hemangiomas, which may indicate underlying spinal dysraphism and cause cord compression. Patients with three or more hemangiomas should be screened by ultrasound for involvement of abdominal viscera, as large hepatic lesions may lead to high-output heart failure. Large segmental hemangiomas in the cranial nerve V distribution (Fig. 45-45) should raise suspicion for PHACES association (Posterior fossa malformations, Hemangiomas, Arterial anomalies, Cardiac defects, Eye anomalies, Sternal defects).46 The LUMBAR association (Lower body hemangiomas, Urogenital anomalies, Myelopathy, Bony deformities, Anorectal/Arterial malformations, Renal anomalies) should be considered in patients with large infantile hemangiomas of the lumbosacral region or lower extremities.41

Oral propranolol therapy has emerged as the first-line treatment for complicated or high-risk infantile hemangiomas. When administered during the proliferative phase, this nonselective beta adrenergic receptor blocker causes rapid involution of the hemangioma. Several randomized, controlled trials have demonstrated oral propranolol to cause a greater decrease in lesion size compared to placebo and steroid therapy.42 In addition, many clinicians believe the side effect profile of propranolol (hypoglycemia, sleep disturbances, hypotension, bradycardia, bronchospasm) to be more favorable than that of systemic steroids.43

While hemangioma involution may result in no visible sequelae, up to 50% of patients are left with a residual fibrofatty mass with atrophic, hypopigmented and/or telangiectatic overlying skin (Fig. 45-46A,B). If the residual deformity is troubling to the patient, surgical excision may be indicated.46

Vascular Malformations Vascular malformations are collections of abnormally formed vessels that demonstrate minimal endothelial cell turnover. They are present at birth and grow slowly in proportion with the patient. Vascular malformations are classified on the basis of anatomic origin of the abnormal vessels: capillary malformations (CM), venous malformations (VM), lymphatic malformations (LM), and arteriovenous malformations (AVM). These classes can be further categorized into “slow-flow” or “fast-flow” lesions (Table 45-4).46

Capillary malformations, formerly known as “port wine stains,” present at birth as flat, pink patches of skin. They typically darken with age and may develop a thickened or “cobblestoned” appearance. CMs may be found anywhere on the body, and overgrowth of underlying soft tissue or bone can occur. History and physical is sufficient to diagnose isolated CMs, but syndromic associations do exist that would warrant
Sturge-Weber syndrome often presents with CMs in the V1/V2 nerve distributions of the face and may be accompanied by vascular malformations of the underlying leptomeninges or globe. Patients are at high risk for seizure, stroke, and glaucoma, for which pharmacologic prophylaxis may be indicated. The mainstay of treatment of CMs is pulsed-dye laser therapy (Fig. 45-47A, pre procedure; Fig. 45-47B post procedure). Other surgical interventions, if necessary, are aimed at addressing soft tissue or bony overgrowth.

Venous malformations are lobulated collections of dilated veins that typically involve skin, mucosa, or subcutaneous tissue, although 50% demonstrate deeper involvement. Lesions may or may not be noted at the time of birth. VMs generally grow in proportion to the patient but may undergo accelerated growth during puberty or pregnancy. Swelling of the mass may occur with dependent positioning or Valsalva maneuvers, such as crying. On exam, superficial VMs are soft, compressible masses with a bluish hue. Firm, tender nodules may be present, which represent calcifications known as phleboliths. Deeper, intramuscular VMs may present with pain or increased extremity circumference, while lesions of the GI tract may simply present with bleeding. MRI with contrast is the imaging modality of choice, although ultrasound can be used in infants and young children to avoid sedation. Observation is indicated for asymptomatic lesions. Compression of involved extremities helps alleviate pain and swelling and prevent thrombosis and phlebolith formation. Due to the high risk of recurrence after surgical excision, the first line of treatment for symptomatic VMs is sclerotherapy. Surgery is reserved for small, well-localized lesions amenable to complete resection; extremity lesions near major peripheral nerves; or residual deformities after sclerotherapy (Fig. 45-48A, before laser; Fig. 45-48B, after laser; and Fig. 45-48C, after limited resection).

### Table 45-4: Classification of vascular malformations

<table>
<thead>
<tr>
<th>SLOW FLOW</th>
<th>FAST FLOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capillary malformations</td>
<td>Arteriovenous malformations</td>
</tr>
<tr>
<td>Venous malformations</td>
<td></td>
</tr>
<tr>
<td>Lymphatic malformations</td>
<td></td>
</tr>
</tbody>
</table>

Figure 45-45. Large segmental hemangiomas in the cranial nerve V distribution.

Figure 45-46. Twenty-year-old female with a capillary malformations of the right cheek. A. Before and (B) after pulsed-dye laser treatment.

Figure 45-47A, B...
**Figure 45-47.** A. A 3-year-old patient with an involuting hemangioma of the right cheek. B. The same patient at 8 years of age showing minimal sequelae after completion of involution.

**Figure 45-48.** A 5-year-old boy with venous malformation of the lower lip. A. Initial presentation. B. After three sclerotherapy treatments. C. Six weeks after surgical debulking of residual fibrotic tissue.
Lymphatic malformations, previously referred to as “cystic hygromas,” are collections of abnormal lymph channels that may cross multiple tissue planes and cause swelling, pain, bleeding, or bony overgrowth. LMs are classified as macrocystic, microcystic or combined. Large, macrocystic lesions can alter form and impair function locally through mass effect. Cutaneous components of LMs present as vesicles that may bleed or become infected. While superficial lesions can be diagnosed by history and physical exam alone, deeper lesions require MRI to confirm the diagnosis and assess the extent of the disease. Asymptomatic LMs can be observed. Sclerotherapy is the treatment of choice for all macrocysts. Symptomatic microcystic LMs have been treated with oral sirolimus, and draining cutaneous vesicles have been successfully ablated with CO₂ laser therapy. Recurrence after surgery is common; therefore, excision is reserved for severely symptomatic lesions no longer amenable to sclerotherapy or small, well-localized lesions where excision can be curative (Fig. 45-49A–C).46

Figure 45-49. A. Lymphatic malformation of the neck. B. After sclerotherapy with significant skin excess. C. Seven months after resection of excess skin.
Arteriovenous malformations are abnormal vascular connections between arteries and veins without intervening capillary beds. AVMs involving the skin appear pink and are warm to the touch. A palpable pulse or thrill may be present from the fast-flow shunting of blood from arterial to venous circulation. Lack of local capillaries can cause a painful, ischemic ulceration of the skin. Patients with large AVMs are at risk for development of congestive heart failure. Doppler ultrasound is the imaging modality of choice, but MRI is often obtained to provide additional information on the extent of the lesion. Observation is appropriate for asymptomatic AVMs. For symptomatic AVMs, embolization is frequently employed 24 to 72 hours prior to excision to minimize operative blood loss. Excision or embolization alone is rarely curative and highly likely to recur. Indications for surgery include small, well-localized AVMs; focal deformities that result from an AVM; or symptomatic AVMs not amenable to embolization.46

When multiple types of vascular malformations cohabit, they are collectively referred to as combined malformations. Patients with Klippel-Trenaunay syndrome demonstrate a combined capillary, venous, and lymphatic malformation of an extremity resulting in bony and/or soft tissue overgrowth (Fig. 45-50).45

<table>
<thead>
<tr>
<th>PROJECTED ADULT DIAMETER</th>
<th>CMN CLASSIFICATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1.5 cm</td>
<td>Small</td>
</tr>
<tr>
<td>≥1.5 cm and &lt;11 cm</td>
<td>Medium</td>
</tr>
<tr>
<td>≥11 cm and ≤20 cm</td>
<td>Large</td>
</tr>
<tr>
<td>&gt;20 cm</td>
<td>Giant</td>
</tr>
</tbody>
</table>

### Congenital Melanocytic Nevi

Congenital melanocytic nevi (CMN) are hyperpigmented lesions present at birth that result from ectopic rests of melanocytes within the skin. They can be distinguished histologically from acquired nevi by their extension into the deep dermis, subcutaneous tissue, or muscle.46 Depending on their size and location, CMNs may cause severe disfigurement and accompanying psychologic distress. Classification is based on projected diameter of the largest dimension on the fully-grown adult (Table 45-5)47. While CMNs are generally common (1% incidence), only 1 in 20,000 children are born with a giant lesion. At birth, CMNs often appear flat, brown and hairless. They grow in proportion with the patient and may develop color variegation, verrucous thickening, hypertrichosis, erosions, or ulcerations over time. CMNs carry an estimated 0.7% to 2.9% lifetime risk of melanoma, with the majority of cases presenting before puberty. Patients with giant CMNs, multiple satellite lesions, or trunk lesions appear to be at higher risk for malignancy. Melanomas can develop within the CMN itself, but they may also present as primary cancers at distant, extra-cutaneous sites, such as the GI tract or the central nervous system. Patients with CMNs require regular skin surveillance by a dermatologist. A biopsy is indicated for concerning changes in color or shape, nodularity, or ulceration. If melanoma is diagnosed, management should proceed in accordance with standard melanoma treatment guidelines.55

CMNs with multiple (>20) satellite lesions or midline CMNs over the trunk or calvaria should raise suspicion for neurocutaneous melanosis, a condition resulting from melanoblast proliferation in the central nervous system (CNS). In addition to the risk of CNS melanoma, patients with neurocutaneous melanosis may suffer from developmental delay, seizures, intracranial hemorrhages, hydrocephalus, cranial nerve palsies, or tethered spinal cord. High-risk patients should be evaluated by MRI between 4 and 6 months of age. While asymptomatic patients may be followed with serial MRI, patients with symptomatic neurocutaneous melanosis often succumb to their disease within 2 to 3 years of diagnosis.54

The goals in surgical management of CMN are (a) to decrease cancer risk, (b) to reduce symptoms, (c) to improve appearance, (d) to improve psychosocial health, and (e) to maintain function.55 It is important to note that the risk of melanoma is not eliminated even with complete excision of a CMN. Indeed, a definitive cancer risk reduction from surgical excision of CMNs has yet to be proven. Management paradigms have therefore shifted from complete excision and reconstruction to maximal excision and reconstruction without compromising function or aesthetic outcome.55 From serial excisions or skin grafting, to tissue expansion or free tissue transfer, plastic surgeons have drawn from the entire armamentarium in meeting the substantial reconstructive challenges posed by giant CMNs. Treatment plans must be grounded in principle: “tissue...
losses should be replaced in kind,” and “reconstruct by units.”48

Figure 45-51A–C shows an infant with a giant CMN of the posterior trunk and right flank preoperatively; at end of the first round of tissue expansion; and at the culmination of three rounds of tissue expansion, excision, and closure by local tissue rearrangement.49

Reconstructive surgery applies techniques that modify tissues in order to restore a normal function and appearance in a patient with congenital or acquired deformities. The most common causes of acquired deformities are traumatic injuries and cancer.

Figure 45-51. A. An infant with a giant CMN of the posterior trunk and right flank. B. Tissues expanders were placed under adjacent normal skin in preparation for first-stage excision. C. The same patient at 11 years of age after three rounds of tissue expansion and excision.
We will focus first on trauma. Although any anatomic region can be subjected to injuries that might require reconstructive surgery, traumatic fractures, and soft tissue damage in the head and neck and extremities are most common. The manner in which these reconstructive steps are conducted is critical. Reconstructive surgery involves the coordination of many specialties and must occur according to a particular timeline, involving complex system-based practice.

**Maxillofacial Injuries and Fractures**

Management of maxillofacial injuries typically occurs in the context of multiple trauma. Concomitant injuries beyond the face are the rule rather than the exception. The first phase of care is activation of the advanced trauma life support protocols. The most common life-threatening considerations in the facial trauma patient are airway maintenance, control of bleeding, identification and treatment of aspiration, assessment for closed head injuries, and identification of other injuries. Once the patient’s condition has been stabilized and life-threatening injuries managed, attention is directed to diagnosis and management of craniofacial injuries.

Physical examination of the face focuses first on assessment of soft tissue injuries as manifested by surface contusions and lacerations. Part of this process is intranasal and intraoral examination. Associated injuries to the underlying facial skeleton are determined by observation, palpation, and digital bone examination through open lacerations. Signs of a facial fracture include contour abnormalities, irregularities of normally smooth contours such as the orbital rims or inferior border of the mandible, instability, tenderness, ecchymosis, facial asymmetry, or displacement of facial landmarks. Traditional plain radiographs have largely been replaced by high-resolution CT, which is widely available at emergency centers that typically receive these patients. Reformatting raw scans into coronal, sagittal, and 3D views is a valuable method to elucidate and plan treatment for complex injuries.

The facial skeleton can be divided into the upper third, middle third, and lower third. The upper third is comprised bounded inferiorly by the superior orbital rim and is formed by the frontal bone. The middle third is the most complex and is formed primarily by the maxilla, nasal bones, and zygoma. The lower third is inferior to the oral cavity and is formed by the mandible. The functional structure of the midface may be understood as a system of buttresses formed by the frontal, maxillary, zygomatic, and sphenoid bones. These buttresses are oriented vertically and horizontally and distribute forces applied to the bones in order to maintain their shape and position without fracturing. There are three paired vertical buttresses called the nasomaxillary, zygomaticomaxillary, and pterygomaxillary buttresses. The horizontal buttresses of the midface pass through the superior and inferior orbital rims and hard palate. A guiding principle of facial fracture management is to restore the integrity of these buttresses.

**Mandible Fractures**

Mandibular fractures are common injuries that may lead to permanent disability if not diagnosed and properly treated. The mandibular angle, ramus, coronoid process, and condyle are points of attachment for the muscles of mastication, including the masseter, temporalis, lateral pterygoid, and medial pterygoid muscles (Fig. 45-52). Fractures are frequently multiple. Alterations in dental occlusion usually accompany mandible fractures. Malocclusion is caused by forces exerted on the mandible of the many muscles of mastication on the fracture segments. Dental occlusion is perhaps the most important basic relationship to understand about fracture of the midface and mandible. The Angle classification system describes the relationship of the maxillary teeth to the mandibular teeth. Class I is normal occlusion, with the mesial buccal cusp of the first maxillary molar fitting into the intercuspal groove of the mandibular first molar. Class II malocclusion is characterized by anterior (mesial) positioning, and class III malocclusion is posterior (distal) positioning of the maxillary teeth with respect to the mandibular teeth (Fig. 45-53). These occlusal relationships guide clinical management.

The goals of surgical treatment include restoration of dental occlusion, fracture reduction and stable fixation, and soft

![Figure 45-52. Mandibular anatomy.](#)

![Figure 45-53. Angle classification. Class I: The mesial buccal cusp of the maxillary first molar fits into the intercuspal groove of the mandibular first molar. Class II: The mesial buccal cusp of the maxillary first molar is mesial to the intercuspal groove of the mandibular first molar. Class III: The mesial buccal cusp of the maxillary first molar is distal to the intercuspal groove of the mandibular first molar.](#)
tissue repair. Nonsurgical treatment may be used in situations in which there is minimal displacement, preservation of the pretraumatic occlusive relationship, normal range of motion, and no significant soft tissue injury. Operative repair involves first establishing and stabilizing dental occlusion and holding in place with maxillomandibular fixation to stabilize the relationships between the mandible and maxilla. The simplest method for this is to apply arch bars to the maxillary and mandibular teeth then use secure them together using interdental wires. Alternatives are sometimes indicated (e.g., screws placed into the bone of the maxilla and mandible that serve as posts for spanning the maxilla and mandible with wires), especially for patients with poor dentition. Once the dental relationships are established, then the fractures can then be reduced and fixed using wire or plates and screws that are specially designed for this purpose. The fracture is surgically exposed using multiple incisions, depending on the location of the fracture and condition of the soft tissues. The fracture is visualized and manually reduced. Fixation may be accomplished using traditional interfragment wires, but plating systems are generally superior. The mandibular plating approach follows two schools of thought: rigid fixation as espoused by the Association for Osteosynthesis/Association for the Study of Internal Fixation and less rigid but functionally stable fixation (Champy technique). Regardless of the approach, it is important to release maxillomandibular fixation and begin range of motion as soon as possible to prevent temporomandibular joint ankylosis. Complications to be avoided include infection, nonunion, malunion, malocclusion, facial nerve injury, mental nerve injury, and dental fractures.

Frontal Sinus Fractures
The frontal sinus is located in the upper third of the face. It is actually a paired structure ordinarily fused in the midline immediately superior to the orbital rims. It has an anterior bony table that defines the contour of the forehead and a posterior table that separates the sinus cavity from the underlying dura of the intracranial frontal fossa. The anterior table is a relatively weak and subject to fracture when it sustains a direct forceful blow, making frontal sinus fractures relatively common in facial trauma. Each sinus drains through the medial floor into its frontonasal duct, which empties into the middle meatus within the nose.

Treatment of a frontal sinus fracture depends on the fracture characteristics as shown in the algorithm (Fig. 45-54). The diagnosis is established by physical examination and confirmed by CT scan. Closed fractures that are not depressed and causing a visible deformity may be observed. Depressed or open fractures must be explored. Fractures that involve only the anterior table are reduced and fixed using interosseous wires or miniature plates and screws. Fractures of the posterior table without disruption of the dura evidenced by leaking cerebrospinal fluid can be treated in similar fashion. When the dura is disrupted, excising the bone and mucosa or the posterior table

![Image of algorithm for the treatment of frontal sinus fracture](brunicardi_ch45_p1967-p2026.png)
and obliterating the nasofrontal duct with a local graft or flap converts with frontal sinus into the anterior frontal fossa of the cranial vault, “cranializing” it.

**Orbital Fractures**

Treatment of all orbital injuries begins with a careful examination of the globe, which often is best completed by a specialist to assess visual acuity and ocular mobility and to rule out globe injury. Fractures may involve the orbital roof, the orbital floor, or the lateral or medial walls (Fig. 45-55). The most common fracture involves the floor because this is the weakest bone. This type of fracture is referred to as an orbital a “blow-out” fracture because the cause is usually direct impact to the globe that results in a sudden increase in intraorbital pressure with failure of the orbital floor. The typical history is either a direct blow during an altercation or a sports-related event with a small ball directly striking the orbit. Because the medial floor and inferior medial wall are made of the thinnest bone, fractures occur most frequently at these locations. These injuries may be treated with observation only if they are isolated and small without signs of displacement or limitation of mobility of the globe. However, surgical treatment is generally indicated for large fractures or ones associated with enophthalmos (retrusion of the globe), which suggests increased intraorbital volume and restriction of upward gaze on the injured side, with entrapment of inferior orbital tissues or double vision (diplopia) persisting greater than 2 weeks. There are a variety of options for surgical exposure of the orbital floor, including the transconjunctival, subciliary, and lower blepharoplasty incisions. All provide good access for accurate diagnosis and treatment, which involves reducing orbital contents and repairing the floor with either autologous bone or synthetic materials. Late complications include persistent diplopia, enophthalmos, or displacement of the lower eyelid ciliary margin inferiorly (ectropion) or rolling inward (entropion). Entropion causes the eyelashes to brush constantly against the cornea and is very uncomfortable. Each of these sequelae has procedures for repair should they occur.

Orbital floor fractures can be associated with fractures of the lateral or inferior orbital rim. These are typically a component of facial fractures that extend beyond the orbit involving the zygomatic and maxillary bones and are discussed in more detail in the next section.

It is important to be aware of two adverse associated conditions seen at times in patients with orbital fractures. The first is superior orbital fissure syndrome. Cranial nerves III (oculomotor nerve), IV (trochlear nerve), and VI (abducens nerve), and the first division of cranial nerve V (VI, trigeminal nerve) pass into the orbit from the base of the skull and into the orbit through the superior orbital fissure. Direct fractures of the posterior orbit or localized swelling caused by a fracture nearby can cause compression of these nerves. Symptoms include eyelid ptosis, protrusion of the globe (proptosis), paralysis of the extraocular muscles, and anesthesia supraorbital and trochlear nerve distributions. The second condition to remember is orbital apex syndrome. This is the most severe circumstance in which superior orbital fissure syndrome is combined with signs of optic nerve (cranial nerve II) compression manifested visual changes ranging up to complete blindness. This is a medical emergency that requires immediate treatment to prevent permanent loss of function.

**Zygomaticomaxillary Complex Fractures**

The zygoma defines the lateral contour of the middle third of the face and forms the lateral and inferior borders of the orbit. It articulates with the sphenoid bone in the lateral orbit, the maxilla medially and inferiorly, the frontal bone superiorly, and the temporal bone laterally. It forms the anterior portion of the zygomatic arch, articulating with the zygomatic projection of the temporal bone. The temporalis muscle, a major muscle of mastication, passes beneath the zygomatic arch and inserts on the coronoid process of the mandible.

Fractures of the zygomatic bone may involve the zygomatic arch alone or any of its other portions and bony relationships. Isolated arch fractures manifest as a flattened, wide facial appearance with edema and ecchymosis. Typically, they are also associated with pain or limited mobility of the mandible. Nondisplaced fractures may be treated without surgery, but
displaced or comminuted fractures should be reduced and stabilized. This can be accomplished using an indirect approach from above the hairline in the temporal scalp, the so-called “Gilles approach,” or directly through a coronal incision in severe fractures.

A common fracture pattern is called the zygomaticomaxillary complex (ZMC) fracture. This involves the zygomatic arch, the inferior orbital rim, the zygomaticomaxillary buttress, the lateral orbital wall, and the zygomaticofrontal buttress. Muscle forces acting on the fracture segment tend to rotate it laterally and inferiorly, thereby expanding the orbital volume, limiting mandibular excursion, creating an inferior cant to the palpebral fissure, and flattening the malar eminence. ZMC fractures are almost always accompanied on physical examination by altered sensation in the infraorbital nerve distribution and a subconjunctival hematoma.

Treatment of displaced ZMC fractures is surgical. Each fracture site is exposed through incisions strategically placed to gain access but minimize disfiguring facial scars afterwards. These include an incision in the upper eyelid, exposing the zygomaticofrontal buttress and lateral orbital wall; a subtemporal or transcconjunctival incision in the lower eyelid, exposing the orbital floor and infraorbital rim; and a maxillary gingivobuccal sulcus incision, exposing the zygomaticomaxillary buttress. Severe fractures involving the arch require wide exposure through a coronal incision.

**Nasoorbitalethmoid and Panfacial Fractures**

Nasoorbitalethmoid (NOE) fractures are defined anatomically by a combination of injuries that involve the medial orbits, the nasal bones, the nasal processes of the frontal bone, and the frontal processes of the maxilla. If improperly treated, these injuries cause severe disfigurement and functional deficits from nasal airway collapse, medial orbital disruption, displacement of medial canthus of the eyelids, and nasolacrimal apparatus dysfunction. Telecanthus is abnormally wide separation of the medical canthus of the eyelids and is produced by a splaying apart of the nasomaxillary buttresses to which the medial canthal ligaments are attached. NOE fractures require surgical management with open reduction and internal fixation. At times, the thin bones are so comminuted that they are not salvageable and must be replaced or augmented using autologous bone grafts or synthetic materials. Each fragment is carefully identified, returned to a normal anatomic position, and fixed in place using plates and screws or intersosseous wiring all bone fragments meticulously, potentially with primary bone grafting, to restore their normal configuration. The key to the successful repair of NOE fractures is to carefully reestablish the nasomaxillary buttress and to restore the normal points of attachment of the medial canthal ligaments.

NOE fractures are typically caused by such extreme forces that they are frequently associated with intracranial injuries and multiple other facial bone fractures in a presentation referred to as a panfacial fracture. These may involve any combination of the fractures described previously. The challenge of these injuries is to reestablish normal relationships of key anatomic landmarks. A combination of salvageable bone fragments, autologous bone grafting, and synthetic materials accomplishes this.

**Posttraumatic Extremity Reconstruction**

The primary goal in posttraumatic extremity reconstruction is to maximize function. When structural integrity, motor function, and sensation can be reasonably preserved, then extremity salvage may be attempted. Otherwise, severe injuries require amputation best performed following reconstructive surgery principals that set the stage for maximizing function with prosthetics and minimizing chronic pain and risk of tissue breakdown. Microvascular surgical techniques are an essential part of extremity trauma surgery, allowing replantation of amputated parts or transfer of vascularized bone and soft tissue when tissue in zone of injury cannot be salvaged. Soft tissue techniques combined with advances in bone fixation and regeneration with distraction have proven tremendous benefit for patients with severe limb-threatening extremity trauma. Current state-of-the-art techniques require multidisciplinary cooperation between orthopedic, vascular, and plastic surgeons as presented in the algorithm (Fig. 45-56). Reconstructive techniques include the use of vascularized bone, bone distraction techniques, external fixation, nerve grafts and transfers, composite tissue flaps, and functioning muscle transfers tailored to the given defect. The future promises further advances with routine application of vascularized composite allografts, engineered tissue replacements, and computer animated prosthetics controlled intuitively by patients via sensors that are placed on the amputation stump and able to detect impulses transmitted through undamaged peripheral nerves remaining in the extremity.

Common causes of high-energy lower extremity trauma include road traffic accidents, falls from a height, direct blows, sports injuries, and gunshot wounds. As with maxillofacial trauma, the first phase of care is activation of the advanced trauma life support protocols. The most common life-threatening considerations are airway maintenance, control of bleeding, and identification of other injuries. Once the patient’s condition has been stabilized and life-threatening injuries managed, attention is directed to diagnosis and management of the extremity. Tetanus vaccine and antibiotics should be provided as soon as possible for open wounds.

Systematic evaluation of the traumatized extremity helps to ensure no important findings are missed. Physical examination to assess the neurovascular status, soft tissue condition, and location of bone fractures forms the foundation of ordering imaging studies to provide details of bone and vascular injuries. Evidence of absent pulses is an indication to consider Doppler ultrasound examination followed by angiography to detail the exact nature of the injury. The blood supply must be immediately restored to devascularized extremities. Crush injuries might be associated with compartment syndrome, in which tissue pressure due to swelling in the constricted facial compartments exceeds capillary perfusion pressure and causes nerve and muscle ischemia. In the early stages of compartment syndrome, findings include pain on passive stretch of the compartment’s musculature in a pale, pulseless extremity without evidence of direct vascular injury. Neurologic changes consisting of paresthesias followed by motor paralysis are late signs. Once recognized, decompressive fasciotomies must be performed as soon as possible to prevent permanent tissue loss. Compartment syndrome can be a late event after fracture reduction and fixation (either internal or external), so the extremity must be reevaluated regularly in the early postoperative period. This is especially true in situations where there has been a period of ischemia prior to successful revascularization.

Several scoring systems for extremity trauma severity have been suggested to aid in treatment planning. Open fractures can be classified according to a system devised by Gustilo and
Figure 45-56. Algorithm of posttraumatic extremity reconstruction.

colleagues. Grades I and II are open fractures with minimal soft tissue disruption. Grade III injuries most often require consideration of soft tissue reconstruction. Grade IIIA are open fractures with severe soft tissue injury but adequate soft tissues to repair. Grade IIIB involves a loss of soft tissue that will require some technique for tissue replacement. Grade IIIC involves a vascular injury requiring reconstruction. For the most severe injuries, the most important decision is whether to attempt extremity salvage or proceed with amputation. Patients with extensive fracture comminution, bone or soft tissue loss, wound contamination, and devascularization have a poor prognosis. Extremity salvage requires multiple operations and a prolonged period of rehabilitation and physical therapy. The loss of plantar sensation historically favored below-knee amputation, but this is no longer an absolute recommendation. A final decision to attempt salvage must be made within the context of comorbidities, socioeconomic considerations, patient motivation, and overall rehabilitative potential.

The first step in surgical management is complete debridement of all devitalized tissue. Early one-stage wound coverage and bony reconstruction is generally advocated and should be performed jointly by extremity trauma orthopedic and plastic surgical teams. It is acceptable for reconstruction to be deferred briefly if the adequacy of debridement is certain. Negative pressure wound therapy is useful after debridement and definitive reconstruction to control the wound drainage and prevent bacterial contamination. When there is segmental bone loss, it is advisable to achieve soft tissue closure prior to performing osseous reconstruction. Preparation for later restoration of the bone requires steps to prevent the soft tissue from collapsing into the space where bone is needed. A common technique for this is to fill the space with antibiotic-impregnated beads or an antibiotic spacer at the time of soft tissue restoration until definitive bony reconstruction is possible. An external fixation may be needed, if there is segmental bone loss (Fig. 45-57A,B).

The sequence for reconstruction is meticulous debridement of nonviable tissue, fracture reduction and stabilization, vascular repair if necessary, and finally restoration of the soft tissue coverage. A multidisciplinary team of specialists works together to perform these procedures in order to obtain the best outcomes. Orthopedic and plastic surgeons perform wound debridement. Orthopedic surgeons then reduce and stabilize the fractures. Vascular surgeons reconstruct damage major vessels. Finally, plastic and reconstructive surgeons perform soft tissue coverage. Ideally, each operating team completes their part of the procedure sequentially during the same anesthetic.

Choices for soft tissue coverage of open fractures include split-thickness skin grafts, temporary skin substitutes followed later by skin grafting, local rotation flaps, or free tissue transfers. Selecting the most appropriate option depends on the quality of the local tissues and location of the soft tissue defect relative to the underlying fracture and fixation hardware. The guiding principle is to be certain that the source of tissue transferred into the defect is outside of the zone of injury. When flaps are selected, either fasciocutaneous or muscular flaps may be indicated depending on tissue availability, wound bed contours, and surgeon preferences. Uneven wound surface contours are more reliably obliterated with a
pliable muscle flap. Fasciocutaneous flaps may provide more durable coverage in areas subject to abrasion or pressure from footwear, for example, on the foot or around the ankle. Some defects can be covered with flaps containing both skin and muscle if indicated. Ideal coverage for weight-bearing areas should be able to resist pressure and shear and provide sensation. Split-thickness skin grafts are reasonable for coverage of exposed healthy muscle or soft tissue. Local flaps may be used to cover smaller defects as long as uninjured tissue is located nearby. These may be designed as traditional random or axial flaps, but the most advanced techniques are based on underlying perforators that allow extremely versatile flap designs customized to the defect. These flaps are designed with a perforating vessel at the base near to the defect and a long axis extending an equal distance opposite. The flap is elevated and rotated into the defect in a motion reminiscent of an airplane propeller, which gives rise to the designation “propeller flap” for this kind of reconstruction (Fig. 45-58A, defect ulnar side of the forearm, with an external fixator; Fig. 45-58B, propeller flap; Fig. 45-58C, flap is inset; Fig. 45-58D, 6 weeks after operation).
When requirements exceed the potential for skin grafts or local flaps, tissue must be transferred from distant sites. The reconstructive choices differ based on the anatomic location of the defect and the extent of damage. This is often the case for major injuries in the middle or lower third of the leg where bones are covered with thin soft tissue and less donor tissue is available. A traditional method is to obtain tissue by creating a pedicled flap from the opposite, uninjured extremity. Cross-leg flaps remain effective, but indications are limited to circumstances where microsurgery is not possible or in young children who are less prone to risks associated with prolonged immobilization necessary for these flaps, such as joint stiffness or deep vein thrombosis. Free tissue transfer is the preferred alternative. The general principles of reconstructive microsurgery in lower extremity trauma are to select recipient vessels outside of the zone of injury, select donor tissue suitable for the defect with minimal risk of donor site morbidity, and ensure there is bone stability before reconstruction using either internal or external fixation. For example, a latissimus dorsi muscle flap provides a large amount of tissue for reconstruction, but loss of the latissimus function can make it more difficult for the patient to use crutches for ambulation during rehabilitation. Muscle or fasciocutaneous flaps each have a role in selected circumstances.51 Bone can also be added to help fracture repair.52 Free flaps can also be designed as “flow-through” flaps, which reconstruct missing segments of major vessels and provide soft tissue or bone coverage.53

After wound healing, proper physical and/or occupational therapy and rehabilitation is essential for the best long-term outcomes. This often requires many months of consistent retraining and conditioning in order to return to the functional status enjoyed by the patient before injury. Properly fitted orthotic appliances and footwear provide essential protection against pressure-related complications and can improve function. Late complications such as osteomyelitis may appear, evidenced by signs of infection months or even years after reconstruction. Very often this is caused by inadequate debridement at the time of initial surgery.

When limb salvage either is not possible or is not in the best interest of the patient, amputation is indicated. Maximizing limb length, providing durable soft tissue coverage, and managing peripheral nerves to avoid chronic pain help to ensure good functional recovery using extremity prosthetics. Ideally, local tissues are used; however, when they are unavailable or inadequate, the amputated part can be a useful source of skin grafts or tissues for microvascular free transfers to the stump, which preserves length and avoids a more proximal amputation. Transected nerves from amputation procedures can be managed using a technique called targeted muscle reinnervation (TMR). TMR surgery takes the transected peripheral nerves resulting from the amputation procedure, and a nerve transfer is then performed to freshly deinnervated motor nerves within the residual limb or stump. By performing these nerve transfers, the sensory and mixed-motor sensory nerves typically transected during amputation are given fresh motor nerves to rapidly reinnervate, which can directly aid in bioprosthetic function and improve pain control. The improvement in pain is a result of reducing phantom limb pain and symptomatic neuroma formation. This technique has shown to be a major advance over traditional traction neuroectomy techniques, which often contribute to increased phantom and residual limb pain rates and a much higher chance of symptomatic neuroma formation compared to TMR.54

Oncologic Reconstructive Surgery
Oncology-related reconstructive surgery has broad applications in the specialty of plastic and reconstructive surgery. Solid tumors necessarily destroy normal tissues, and surgical treatment involves excising the tumor with a margin of uninvolved normal tissue, which adds to the extent of tissue loss. As is illustrated in the case of a lower extremity sarcoma, reconstructive strategies are meticulously designed as an algorithm for effective functional and cosmetic restoration (Fig. 45-59). Chemotherapy and radiation have side effects and complications that can cause tissue loss, leading to functional and cosmetic deformities that can be improved with reconstructive surgery. The goal of comprehensive cancer treatment is to restore the patient to full health, which includes normal function and appearance.

Figure 45-59. Algorithm for effective functional and cosmetic restoration after resection of a lower extremity sarcoma.
Reconstructive surgery in the context of oncology has several distinctive aspects compared to the larger field of reconstructive surgery in general. The procedure must be highly reliable in order to avoid surgical complications that might interfere with adjuvant therapies.

Breast Reconstruction
Breast cancer is the most common malignancy besides skin cancer in women and the second leading cause of cancer-related death for women in the United States. Breast reconstruction is an important part of comprehensive cancer treatment. A number of studies have shown that breast reconstruction, both immediate and delayed, does not impede standard oncologic treatment, does not delay detection of recurrent cancer, and does not change the overall mortality associated with the disease.46–48

Preoperative counseling of the breast cancer patient regarding reconstruction options should include discussion of the timing and technique of reconstruction. It is important to ensure that the patient has realistic expectations of outcome and an understanding of the number of procedures that might be necessary to perform in order to obtain the best outcome. The plastic surgeon and surgical oncologist must maintain close communication to achieve optimal results.

Delayed breast reconstruction occurs any time after the mastectomy is performed, usually 3 to 6 months after the operation, depending on the patient’s circumstances and reasons for not electing immediate reconstruction. Although good outcomes can be obtained, it is more difficult to achieve a result that is similar to the preoperative breast shape and size because of established scarring of the chest wall. Nevertheless, it is a good option for patients who are undecided or not candidates for immediate reconstruction because of advanced disease or comorbidities.

Immediate reconstruction is defined as initiation of the breast reconstructive process at the time of the ablative surgery. Patients are considered candidates for immediate reconstruction who are in general good health and have stage I or II disease determined primarily by the size and location of the tumor. There are selected exceptions, such as when an extensive resection requires chest wall coverage. Breast reconstruction might be performed in these cases, but it is really incidental to achieving chest wall coverage. Disadvantages of immediate reconstruction include the potential delay of adjuvant therapy in the event of postoperative complications. Also, if there is uncertainty regarding the need to adjuvant radiation therapy, decision-making regarding immediate reconstruction is a challenge. Breast reconstructions by all techniques are adversely affected by radiation therapy, and many surgeons feel reconstruction should be delayed until at least 6 months after treatment.

Once the patient chooses to have immediate reconstruction, she must select a reconstructive technique. In patients selected for breast conservation, oncoplastic tissue rearrangement can be performed to minimize adverse effects of lumpectomy on breast appearance. For patients electing total mastectomy there are essentially three options: (a) tissue expansion followed by breast implant placement, (b) combined tissue flaps with breast implants, and (c) autologous tissue flaps only. After examining the patient, the surgeon then should describe those methods for which the patient is a satisfactory candidate. The patient should then be encouraged to choose based on her goals and an understanding of the advantages and disadvantages of each technique.

Oncoplastic Breast Reconstruction
Breast conservation therapy (BCT) consists of excision of the breast tumor with a surrounding margin of normal tissue combined with postoperative whole-breast irradiation. Although the overall survival for properly selected patients is shown to be comparable to total mastectomy and reconstruction, the breast can often be distorted and unnatural appearing after treatment. The area of the lumpectomy may create a depression with contour deformity, and contraction of the lumpectomy space over time can distort the nipple out of alignment and create an asymmetry with the contralateral breast. This is especially true for women with small breasts in whom a high percentage of breast volume is removed with the lumpectomy.

Oncoplastic surgery refers to the set of techniques developed to lessen breast deformity from a partial mastectomy. One of the most common methods of minimizing adverse effects on breast appearance is to rearrange the skin, parenchyma, and nipple location of the breast at the time of tumor extirpation using surgical techniques developed for breast aesthetic surgery. This procedure involves elevating the skin from the underlying glandular tissue, mobilizing the nipple on a vascular pedicle, and preserving as much of the vascularized glandular tissue as possible. After lumpectomy, the tissue is rearranged to shift glandular tissue into the defect and redrape the skin and nipple onto the new breast mound. After healing and completion of radiotherapy, a contralateral conventional mastopexy or breast reduction can be performed on the contralateral side to achieve symmetry.

Implant-Based Reconstruction
Immediate breast reconstruction based entirely on the use of implanted devices is initially the most expedient technique. Sometimes it is possible to place a full-size implant at the time of mastectomy when the breasts are small (volume <400 cc) and the patient is a young nonsmoker with good chest wall musculature. In most patients, however, a period of tissue expansion is required. The tissue expander is inserted beneath the pectoralis major and serratus anterior muscles at the time of the mastectomy and partially inflated. Alternatively, the tissue expander can be placed only under the pectoralis major muscle or even completely on top of the chest wall muscles then covered with acellular dermal matrix directly beneath the mastectomy skin flaps. Total muscle coverage is the traditional approach, but these alternatives may be suitable only for well-selected patients. Expansion usually requires 6 to 8 weeks to complete, and an implant exchange is performed typically 3 months later. The advantages of this technique are that it involves minimum additional surgery at the time of the mastectomy, has a recovery period essentially the same of that of the mastectomy alone, and creates no additional scarring. The disadvantages of this technique are the length of time necessary to complete the entire reconstruction (up to 1 year), the requirement for a minimum of two operative procedures, and a less predictable cosmetic result due to complete reliance on devices. Also, the patient awakens from surgery without a full-size breast and during the time of expansion must accept a breast of abnormal size and shape. Although the final shape of the breast may be satisfactory, it may lack a natural consistency due to the superficial placement of the device, especially when saline-filled implants are used. Finally, breast implants may develop late complications such as capsular contracture, infection, or extrusion. This method is ideal for a slender, small-breasted woman with minimal ptosis
who wish to avoid additional scarring and time for convalescence. It may also be suitable for women undergoing bilateral reconstruction because symmetry is more easily achieved if both breasts are restored using the same technique. Women who elect this type of immediate reconstruction must understand that breast implants do not have an unlimited service life and that additional surgery will likely be required to replace the breast implant at some time in the future.

**Tissue Flaps and Breast Implants**

The latissimus dorsi musculocutaneous flap is the most common transfer used in combination with breast implants. Other flaps may also be used, depending on patient preference and tissue availability. The principal advantage in using a tissue flap is immediate replacement of missing skin and soft tissue. In cases where there is already adequate breast skin, then a muscle only may be transferred to provide suitable implant coverage. The implant allows the final breast volume to be accurately reproduced to match the contralateral breast or, in bilateral reconstruction, adjust the breast size according to the patient’s desires. The advantages of this technique are that the implant is protected by abundant tissue, a period of tissue expansion is avoided, and the full benefit of preserving the breast skin is realized to achieve a natural-appearing breast. The disadvantage of this technique compared to implants alone is that it results in additional scarring and requires a longer period of recovery. For many patients, this approach represents an acceptable compromise between implant-only reconstruction and autologous tissue reconstruction, incorporating some of the advantages and disadvantages of each.

**Autologous Tissue Reconstruction**

Immediate reconstruction using only autologous tissue is the most elaborate method of breast reconstruction but consistently yields the most durable, natural-appearing results. Breast implants cannot match the ability of the autologous tissue to conform to the breast skin and envelop and simulate natural breast parenchyma. The most useful flap is the transverse rectus abdominis musculocutaneous (TRAM) flap, although other donor areas are also possibilities in selected cases. Autologous reconstruction is usually the best option in patients who require adjuvant radiation therapy.

The TRAM flap may be transferred to the chest using a variety of methods, depending on the circumstances of the individual patient. As a pedicled flap, it is transferred based on the superior epigastric vessels and tunneled beneath the skin to reach the mastectomy defect. As a free flap, it is based on the inferior epigastric vessels that are revascularized by microvascular anastomosis to vessels on the chest wall nearby the mastectomy defect. Often the microvascular technique using the deep inferior epigastric perforator (DIEP) flap is preferred because there is less risk of partial flap loss or localized areas of fat necrosis due to a more reliable blood supply (Fig. 45-60A, before operation on right breast; Fig. 45-60B, after mastectomy and immediate reconstruction with a DIEP flap). In immediate reconstruction with an axillary dissection, the axillary vessels are completely exposed and free of scar following the lymph node dissection in patients without previous surgery and radiation. In women being treated for recurrence with previous axillary surgery, the axillary vessels are less reliable, and plans should be made for the possibility of using the internal mammary vessels. The internal mammary vessels have become the most common recipient vessels for free tissue transfer in breast reconstruction in the contemporary era of sentinel lymph node biopsy that is used as a technique to perform axillary lymph node dissection in a more limited number of patients. Regardless of the technique used to transfer the tissue, the donor site is closed in a similar manner as an abdominoplasty, by repairing the abdominal wall and advancing the upper abdominal skin downward. The umbilicus is preserved on its vascular stalk brought to the surface through a small incision immediately above its location on the abdominal wall (Fig. 45-61A,B). Other donor sites including the buttock may be used in transferring the skin and fat supplied by the inferior gluteal artery perforator (IGAP) or the superior gluteal perforator as the main blood supply.

The advantages of using this technique are complete restoration of the breast mound in a single stage, avoidance of
potential problems associated with breast implants, and consistently superior cosmetic results. The disadvantages are the magnitude of the operation, additional scarring, risks of development of abdominal bulges, and a longer period of convalescence. Although the initial cost is greater, over the long term the total cost appears to be less because of less need for secondary procedures to exchange implants, achieve suitable cosmetic appearance, or care for implant-related problems. This is the best operation for patients who want the most natural breast restoration possible and who are less concerned about the amount of surgery, scarring, and recovery period.

**Accessory Procedures**
After complete healing of the breast mound from the initial stages of reconstruction, refinements and accessory procedures may be performed at a later time to optimize the natural appearance of the reconstructed breast. These may include soft tissue modifications of the breast mound revision, repositioning or the breast implant, scar revisions, autologous fat grafting, and nipple-areola complex reconstruction. A variety of methods have been described for nipple reconstruction. They are all based on local tissue rearrangements or skin grafts to create a projecting piece of skin and subcutaneous tissue that simulates the natural nipple (Fig. 45-62A,B). The pigmentation of the areola may be simulated with tattooing of colored pigments selected to match the normal coloration of the patient’s original anatomy.

**Trunk and Abdominal Reconstruction**
In the torso, as in most areas of the body, the location and size of the defect and the properties of the deficient tissue determine choice of reconstructive method. A distinction is made between partial-thickness and full-thickness defects when deciding between grafts, flaps, synthetic materials, or a combination of techniques. Unlike the head and the lower leg, the trunk
PART II

SPECIFIC CONSIDERATIONS

2012

harbors a relative wealth of regional transposable axial pattern flaps that allow sturdy reconstruction, only rarely requiring distant free tissue transfer. Indeed, the trunk serves as the body’s arsenal, providing its most robust flaps to rebuild its largest defects.

The chest wall is a rigid framework designed to resist both the negative pressure associated with respiration and the positive pressure from coughing and from transmitted intra-abdominal forces. Furthermore, it protects the heart, lungs, and great vessels from external trauma. Reconstructions of chest wall defects must restore these functions. When a full-thickness defect of the chest wall involves more than four, this is usually an indication for the need for rigid chest wall reconstruction usually using synthetic meshes made of polypropylene, polyethylene, or polytetrafluoroethylene, which may be reinforced with polymethylmethacrylate acrylic. In contaminated wounds, biologic materials are preferred, such as acellular dermal matrix allografts. For soft tissue restoration, the pectoralis major muscle is commonly used as a pedicled flap for coverage of the sternum, upper chest, and neck. It may be mobilized and transferred on a vascular pedicle based on the pectoral branch of the thoracoacromial artery or a vascular supply based on perforators from the internal mammary vessels. Either flap design is useful in covering the sternum after dehiscence or infection occurring as a complication of median sternotomy or with sternal resection for tumor extirpation. For the lower third of the sternum, a rectus abdominis muscle flap based on the superior epigastric vessels or the deep inferior epigastric vessels is useful. If based on the inferior blood supply, it must be transferred as a free flap with recipient vessels outside of the zone in injury. The latissimus dorsi musculocutaneous flap is useful for chest wall reconstructions in places other than the anterior midline. Similar to the pectoralis major muscle, it may be transferred on either a single blood supply that is based on the thoracodorsal vessels from the subscapular system or on vessels perforating from deeper source vessels near to the posterior midline. The serratus anterior muscle can be included on the same vascular pedicle to further increase its surface area. Finally, the trapezius muscle flap, based on the transverse cervical vessels, is generally used as a pedicled flap to cover the upper midback, base of neck, and shoulder. The superior portion of the muscle along with the acromial attachment and spinal accessory nerve must be preserved to maintain normal shoulder elevation function.

The abdominal wall also protects the internal vital organs from trauma, but with layers of strong torso-supporting muscles and fascia rather than with osseous structures. The goals of reconstruction are restoration of structural integrity, prevention of visceral herniation, and provision of dynamic muscular support. Although abdominal wall defects may occur in association with oncologic tumor resections, the most common etiology is fascial dehiscence after laparotomy. When a reconstruction plan is being formulated, careful physical examination and review of the medical history will help prevent selection of an otherwise sound strategy that, because of previous incisions and trauma, is destined for failure.

Superficial defects of the abdominal skin and subcutaneous tissue are usually easily controlled with skin grafts, local advancement flaps, or tissue expansion. Defects of the underlying musculoskeletal structures are more difficult to manage. The abdominal wall fascia requires a minimal-tension closure to avoid dehiscence, recurrent incisional hernia formation, or abdominal compartment syndrome. Prosthetic meshes are frequently used to replace the fascia in clean wounds and in operations that create myofascial defects. When the wound is contaminated, as in infected mesh reconstructions, enterocUTaneous fistulas, or viscus perforations, prosthetic mesh is avoided because of the risk of infection. The technique of component separation procedure has proven beneficial for closing large midline defects with autologous tissue and avoiding prosthetic materials. This procedure involves advancement of bilateral flaps composed of the anterior rectus fascia rectus and oblique muscles after lateral release. Midline defects measuring up to 10 cm superiorly, 18 cm centrally, and 8 cm inferiorly can be closed using this method.

Techniques based on rearranging and reinforcing abdominal wall elements might be inadequate for extremely large or full-thickness abdominal wall defects. For these defects, regional flaps or free flaps are required. Pedicled flaps from the thigh are useful, such as the tensor fasciae latae pedicled flap, based on the ascending branch of the lateral circumflex femoral vessels, or the anterolateral thigh flap, based on the descending branch of the lateral circumflex vessels. Bilateral flaps might be required.

Pelvic Reconstruction

Another important area for consideration of reconstructive surgical procedures is the perineum. The perineal region is part of the specialized part of the trunk that supports the pelvic outlet lying between the pubic symphysis, the coccyx, the inferior rami of the pubis, and the ischial tuberosities. Support is provided by the urogenital diaphragm, the deep and superficial fasciae, and the skin. Specialized anatomic structures pass through the perineum. Posteriorly is the anus, and anteriorly are the genitalia and urethra. Treatment of tumors involving this area often require a combination of surgery and radiation. The resulting loss of tissue and healing impairment coupled with the nonyielding nature of the bony pelvic outlet can result in unique reconstructive requirements that often are best addressed with tissue transfer. The reconstruction must achieve wound healing and restore support to the pelvic contents, accommodate urinary and bowel function, and finally restore the penis in men and the vagina and vulva in women. Local flaps, regional flaps, or free tissue transfer all have possible application depending on the extent of the resection and local tissue compromise.

Other Clinical Circumstances

Besides trauma and cancer, other etiologies can cause functional and cosmetic deformities due to tissue impairment for which reconstructive surgery has value. These include pressure sores, diabetic foot ulcers, and lymphedema.

Pressure Sores. A pressure ulcer is defined as tissue injury caused by physical pressure applied to the tissues from an external source at a magnitude that exceeds capillary perfusion pressure. Prolonged tissue ischemia leads to local tissue necrosis. Pressure ulcers tend to occur in people debilitated by advanced age, chronic illness, poor nutrition, prolonged immobilization, motor paralysis, or inadequate sensation. Spinal cord injury patients are especially prone to developing pressure sores. Pressure sores can also occur in healthy individuals who undergo prolonged surgical operations and parts of the body supporting the weight of the patient on the operating table (e.g., the occiput, the sacral prominence, the heels of the feet) are improperly padded.
Pressure sores are an important contributor to morbidity in patients suffering from limited mobility. Most can be prevented by diligent nursing care in an attentive, cooperative patient. Preventing pressure ulcers requires recognition of susceptible and utilizing appropriate measures to reduce pressure on areas of the body at risk. This involves frequent position changes while sitting or supine and the use of pressure-reducing medical equipment such as low-air-loss mattresses and seat cushions and heel protectors. Malnourishment, poor glucose control in diabetics, poor skin hygiene, urinary or bowel incontinence, muscle spasms, and joint contractures all increase the risk of pressure sore formation. Mitigating these factors is essential before embarking on a complex reconstructive treatment plan. Successful reconstruction also requires a cooperative and motivated patient with good social support.

Surgical treatment of pressure ulcers is based on wound depth. The staging system is summarized in Fig. 45-63. Stage I and II ulcers are treated nonsurgically with local wound care and interventions to relieve pressure on the affected area. Patients with stage III or IV ulcers should be evaluated for surgery. Important features for preoperative assessment include the extent of soft tissue infection, the presence of contaminated fluid collection or abscess, osteomyelitis, and communication with deep spaces (e.g., joint space, urethra, colon, or spinal canal). Laboratory blood tests and imaging studies help establish whether soft tissue or bone infection is present. Plain radiographs are usually adequate to rule out osteomyelitis; CT and MRI are helpful when plain films are equivocal. Necrotic tissue and abscesses should be surgically debrided without delay to prevent or treat systemic sepsis. Bone must also be excised if it appears involved, as evidenced by poor bleeding, softness, or frank purulence. Patients with high spinal cord injuries at or above the level of the fifth thoracic vertebra may experience sudden extreme elevation of blood pressure, an autonomic-mediated event called hyperreflexia. This condition must be immediately recognized and treated to prevent intracranial and retinal hemorrhage, seizures, cardiac irregularities, and death.

After adequate debridement, the pressure ulcer can be treated nonsurgically in patients who have shallow wounds with healthy surrounding tissues capable of healing secondarily with offloading pressure. Nonsurgical treatment is also best in patients for whom surgery is contraindicated because of previous surgery or comorbidities. For surgical candidates, primary closure is rarely performed because an inadequate amount of quality surrounding tissue prevents closure without tension, making the repair predisposed to failure. Split-thickness skin grafting can be useful for shallow ulcers with well-vascularized wound beds on which shear forces and pressure can be avoided after repair, a rare circumstance in most patients with pressure ulcers.

The mainstay of surgical treatment is tissue transfer following several guiding principles. Local muscle or musculocutaneous flaps are suitable for areas of heavy contamination and complex wound surface contours. Durability requires the ability to consistently off-load of the area of reconstruction postoperatively. Fasciocutaneous flaps afford more durable reconstruction when off-loading is not possible. The anatomic location is an important determinant of flap choice. Once a donor site is selected, a flap of adequate size is designed and transferred in a way that avoids suture lines in the area under pressure. Large flaps also permit readvancement if the patient experiences a recurrent ulcer in the same area. Sacral pressure sores may be managed with fasciocutaneous or musculocutaneous flaps based on the gluteal vessels. Ischial pressure sores may be managed with gluteal flaps or flaps transferred from the posterior thigh, such as the posterior thigh flap based on the descending branch of the inferior gluteal artery. Trochanteric ulcers

Stage 1
Observable pressure related alteration of intact skin whose indicators as compared to the adjacent or opposite area of the body may include changes in one or more of the following: skin temperature (warmth or coolness), tissue consistency (firm or boggy feel), and/or sensation (pain, itching). The ulcer appears as a defined area of persistent redness in lightly pigmented skin, whereas in darker skin tones the ulcer may appear with persistent red, blue of purple hues.

Stage 2
Partial thickness skin loss involving epidermis and/or dermis. The ulcer is superficial and presents clinically as an abrasion, blister, or shallow crater.

Stage 3
Full thickness skin loss involving damage or necrosis of subcutaneous tissue that may extend down to but not through underlaying fascia. The ulcer presents clinically as a deep crater with or without undermining of adjacent tissue.

Stage 4
Full thickness skin loss with extensive destruction, tissue necrosis or damage to muscle, bone, or supporting structures (for example, tendon or joint capsule). Undermining and sinus tracts may also be associated with Stage 4 pressure ulcers.

Figure 45-63. The staging system for pressure sores.
may be managed with musculocutaneous flaps based on the
tensor fasciae latae, rectus femoris, or vastus lateralis muscles
(Fig. 45-64). The obligatory loss of motor function associated
with using these flaps adds no additional functional impairment
in patients already paralyzed as a result of strokes or spinal cord
injuries.

Proper postoperative care after flap reconstruction of
pressure ulcers is critical for success. Low-pressure, air fluid-
ized beds help to off-load the affected area and prevent new
areas of involvement during the first 7 to 10 days of healing.
Other important measures are adequate nutritional support and
medications to prevent muscle spasms. Careful coordination
with patient care providers is planned preoperatively in order
to avoid gaps in care that can lead to early recurrent ulceration.
Care of the pressure ulcer patient is a labor-intensive process
that requires attention to detail by the surgeon, nurses, ther-
pists, caseworkers, and family.

Diabetic Foot Ulceration. The pathophysiology of primary
diabetic lower limb complications has three main components:
(a) peripheral neuropathy (motor, sensory, and autonomic),
(b) peripheral vascular disease, and (c) immunodeficiency.
Altered foot biomechanics and gait caused by painless col-
lapse of ligamentous support, foot joints, and foot arches
change weight-bearing patterns. Blunted pain allows cutane-
ous ulceration to begin. With breakdown of the skin barrier
function, polymicrobial infections become established. Bac-
terial invasion is often fostered by poor blood supply due to
peripheral vascular disease coupled with microangiopathy.
Finally, local host defenses may be less effective in resisting
bacteria because of poor blood supply and impaired cellular
function. Cutaneous ulcerations may progress painlessly to
involve deeper soft tissues and bone. The ultimate endpoint
of this process is such severe tissue damage that extremity
amputation is the only treatment remaining. More than 60% of
nontraumatic lower extremity amputations occur in diabetics.
The age-adjusted lower extremity amputation rate in diabetic-
s (5.0 per 1000 diabetics) was approximately 28 times that
of people without diabetes (0.2 per 1000 people). Improved
patient education and medical management, early detection of
foot problems, and prompt intervention play important roles in
improving the chances of limb preservation.

The best approach to managing diabetic patients with
lower extremity wounds is to involve a multidisciplinary team
composed of a plastic and reconstructive surgeon, a vascular
surgeon, an orthopedic surgeon, a podiatrist, an endocrinolo-
gist specializing in diabetes, a nutritionist, and a physical or
occupational therapist. This brings together the greatest level of expertise to manage bone and soft tissue issues as well as the underlying disease and medical comorbidities. Treatment begins with rigorous control of blood glucose levels and a thorough assessment of comorbidities. In addition to careful detailing of the extent of the wound and the tissues involved, physical examination documents sensory deficits and vascular status. Plain radiographs, MRI, bone scintigraphy, and angiography or duplex Doppler ultrasound imaging may be indicated. A patient with significant vascular disease may be a candidate for lower extremity endovascular revascularization or open bypass.61 Nerve conduction studies may diagnose surgically reversible neuropathies at compressive sites and aid in decisions about whether to perform sensory nerve transfers to restore plantar sensibility.60 Antibiotic and fungal therapies should be guided by tissue culture results.

Surgical management starts with debridement of devitalized tissues. Methods of wound closure are dictated by the extent and location of the remaining defect. Negative pressure wound dressings may be appropriate for superficial defects in an effort to allow secondary healing or as a temporizing measure until definitive wound closure can be achieved. Skin grafts might be indicated at times but cannot be expected to provide durable coverage in weight-bearing or high-shear areas. Local and regional flaps can be considered if the extremity is free of significant occlusive peripheral vascular or combined with vascular bypass. Microvascular free tissue transfers are appropriate when defects are large or when local flaps are not available. Combination lower extremity bypass and free flap coverage has proved beneficial for the treatment of the diabetic foot in terms of healing and reduction of disease progression (Table 45-6). Consultation with a podiatrist or an orthopedic surgeon who specializes in foot and ankle problems can be considered to improve foot biomechanics and manage bony prominences that act as pressure points on the soft tissue to reduce the risk of recurrent ulceration. Proper foot-wear (including orthotic devices and off-loading shoe inserts), hygiene, and toenail and skin care are essential.60

**Lymphedema.** Lymphedema is the abnormal accumulation of protein-rich fluid in the interstitial spaces of the tissues. It is a complex disorder with both congenital and acquired causes. No universally effective remedy has been devised, but a variety of treatment methods including reconstructive surgery have been effective in carefully selected patients.

It is important to be familiar with the fundamentals of lymph physiology in order to understand the rationale for the various forms of lymphedema treatment. Lymph fluid is formed at the capillary level where there is a net outflow of fluid and serum proteins from the intravascular space into the interstitium. In the average adult, this amounts to approximately 3 liters of fluid daily. Open-ended lymph capillaries collect this fluid where the lymphatic endothelial cells form loose intercellular connections that freely allow fluid to enter. From here, the network of specialized vascular structures gathers the extravasated fluid and transports it back into central circulation. The system is a high-volume transport mechanism that clears proteins and lipids from the interstitial space primarily by means of differential pressure gradients. Lymph fluid enters the lymph vessels driven by colloid and solute concentration gradients at the capillary level. Flow is sustained in the larger vessels through direct contractility of the lymph vessel walls and by indirect compression from surrounding skeletal muscle activity. Throughout the system, one-way valves prevent reverse flow. The lymphatic vessels course throughout the body alongside the venous system, into which they eventually drain via the major thoracic and cervical ducts at the base of the neck.

Under normal conditions, there is a balance between fluid formation and lymph transport capacity. With congenital hypoplasia or acquired obstruction, there is a reduction in transport capacity resulting in accumulation of fluid and protein in the interstitium. Localized fluid stagnation, hypertension, and valvular incompetence further degrade transport capacity and accelerate lymph fluid accumulation edema. Dissolved and suspended serum proteins, cellular debris, and waste products of metabolism elicit an inflammatory response with associated fibrovascular proliferation and collagen deposition leading to firm, non-pitting swelling characteristic of chronic, long-standing edema. Lymphoscintigraphy can help detail the lymphatic anatomy and quantify lymphatic flow. MRI can provide additional information about the larger caliber lymphatic vessels, possibly helping to identify specific points of obstruction.

Primary lymphedema is caused by congenital hypoplasia and is classified clinically based on the age of the affected individual when swelling first appears. Lymphedema present at birth is an autosomal dominant disorder sometimes referred to as Milroy’s disease. Lymphedema praecox occurs near the time of puberty but can appear up to age 35. This form tends to occur in females and usually affects the lower extremity. It accounts for more than 90% of cases. Finally, lymphedema tarda appears after the age of 35 years and is relatively rare.

Secondary lymphedema is the acquired form of the disorder and is more common than congenital causes. Worldwide the most common etiology is parasitic infestation with filarial, a highly specialized nematode transmitted by blood-eating insects.

### Table 45-6

<table>
<thead>
<tr>
<th>AREA OF DEFECT</th>
<th>RECONSTRUCTIVE OPTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forefoot</td>
<td>V-Y advancement</td>
</tr>
<tr>
<td></td>
<td>Toe island flap</td>
</tr>
<tr>
<td></td>
<td>Single toe amputation</td>
</tr>
<tr>
<td></td>
<td>Lisfranc’s amputation</td>
</tr>
<tr>
<td>Midfoot</td>
<td>V-Y advancement</td>
</tr>
<tr>
<td></td>
<td>Toe island flap</td>
</tr>
<tr>
<td></td>
<td>Medial plantar artery flap</td>
</tr>
<tr>
<td></td>
<td>Free tissue transfer</td>
</tr>
<tr>
<td></td>
<td>Transmetatarsal amputation</td>
</tr>
<tr>
<td>Hindfoot</td>
<td>Lateral calcaneal artery flap</td>
</tr>
<tr>
<td></td>
<td>Reversed sural artery flap</td>
</tr>
<tr>
<td></td>
<td>Medial plantar artery flap ± flexor digitorum brevis</td>
</tr>
<tr>
<td></td>
<td>Abductor hallucis muscle flap</td>
</tr>
<tr>
<td></td>
<td>Abductor digiti minimi muscle flap</td>
</tr>
<tr>
<td></td>
<td>Free tissue transfer</td>
</tr>
<tr>
<td></td>
<td>Syme’s amputation</td>
</tr>
<tr>
<td>Foot dorsum</td>
<td>Supramalleolar flap</td>
</tr>
<tr>
<td></td>
<td>Reversed sural artery flap</td>
</tr>
<tr>
<td></td>
<td>Thinner free flaps (e.g., temporoparietal fascia, radial forearm, groin, thinned anterolateral thigh flaps)</td>
</tr>
</tbody>
</table>
Symptomatic Lymphedema

Amenable to physiologic lymphatic procedures?

Yes
  • Responsive to nonsurgical therapy, but symptoms plateaued or worsening
  • Significant pitting edema

No
  • Minimal or no improvement with nonsurgical therapy
  • Minimal to absent pitting edema

Suitable lymphatic vessels on MRL or ICGL for LVA?

Yes
  Secondary to surgery and/or XRT?
    Yes
      LVA ± VLNT
    No
      LVA only

No
  VLNT only

Severe functional impairment? Excess soft tissue? Skin changes?

Yes
  Liposuction ± excision

No
  Consider further LVA or VLNT

Figure 45-65. Algorithm for lymphedema management.

Aesthetic, or cosmetic, surgery is an important part of the specialty of plastic surgery. The American Medical Association defines cosmetic surgery as “surgery performed to reshape normal structures of the body to improve the patient’s appearance and self-esteem.” It is a natural extension of surgical techniques for tissue modification traditionally developed for other reasons. Because aesthetic surgery primarily relates to personal appearance and attractiveness and not a particular disease process, there has been a tendency to dismiss the health value of
aesthetic surgery. Nevertheless, personal appearance plays an important role in psychosocial health. Physical attractiveness plays a role in the marketplace with well-documented influence on employment opportunities, advancement, and earnings. The multibillion industry of products and services designed to optimize appearance, which spans a wide spectrum between simple cosmetics to elaborate surgical procedures, bears testament to the perceived value by the general population.

Important work demonstrates a link between aesthetic surgery and psychosocial health. Surgery performed on the face, nose, ears, breast, and body can positively affect quality of life on multiple scales. There is a clear association between one’s personal appearance and success in the marketplace. As the primary benefits of aesthetic surgery are related to the psychosocial outcomes, it is important to assess the state of psychological health prior to offering aesthetic surgery. A variety of preoperative psychological comorbidities can adversely affect outcomes, most notably a syndrome known as body dysmorphic disorder, present in individuals who manifest a preoccupation with one or more perceived defects or flaws in physical appearance that are not observable or appear slight to others. Performing a surgical procedure to modify personal appearance in such an individual is associated with a high risk of a poor outcome.

It is important for all surgeons to have an appreciation of the methods of patient evaluation, surgical techniques, and typical outcomes that might be anticipated in aesthetic surgery. Patients seek aesthetic surgery when they are unable to achieve a personal standard of physical appearance without surgical modification of various body parts that most affect their appearance. This is especially true for features that are visible in public and strong determinants of appearance, such as the face, breasts, abdomen, and buttocks. The etiology of undesirable characteristics of form or skin quality can be familial or acquired through natural processes of aging, injury, cancer, or degeneration. Unwanted changes in appearance that result from these processes may still fall within the range of normal appearance yet fall short of the patient’s personal aesthetic ideal. Patient assessment requires an understanding of personal and cultural ideals of appearance. The surgeon must be knowledgeable about the various surgical and nonsurgical techniques that might be considered to address the patient’s concerns.

In practical terms, there are both reconstructive and cosmetic elements to almost every plastic surgery case, and the definition of “normal” structure is sometimes very subjective and difficult to quantify. Nevertheless, there are patients for whom it is a priority to make surgical changes to their bodies in the clear absence of a functional deformity. Aesthetic surgery patients present a unique challenge to the plastic surgeon because the most important outcome parameter is not truly appearance, but patient satisfaction. Optimally, a good cosmetic outcome will be associated with a high level of patient satisfaction. For this to be the case, the plastic surgeon must do a careful analysis of the patient’s motivations for wanting surgery, along with the patient’s goals and expectations. The surgeon must make a reasonable assessment that the improvements that can be achieved through surgery will meet the patient’s expectations. The surgeon must appropriately counsel the patient about the magnitude of the recovery process, the exact location of scars, and potential complications. If complications do occur, the surgeon must manage these in a manner that preserves a positive doctor-patient relationship.

Aesthetic Surgery of the Face

A thorough evaluation of the patient who presents for facial aesthetic surgery begins with acquiring a clear understanding of the patient’s primary concern regarding appearance. Examination focuses on that region but takes into consideration overall facial appearance that might be contributing to the patient’s concerns but of which the patient is unaware. The skin quality is carefully assessed as well as the location, symmetry, and position of each critical feature of facial appearance such as scalp hairline, forehead length, eyebrow shape and position, eyelid configuration, nasal proportions, and shape of the lips. Overall facial proportions are assessed, such as the prominence of the orbital rims and malar areas, the chin projection, and contours along the margin of the mandible. An appropriately performed facelift can yield an aesthetically pleasing result (Fig. 45-66).

A variety of procedures have been described for modifying facial appearance. Nonsurgical interventions topical treatments of the skin surface include chemical and laser facial peels. Injections of biocompatible materials made of processed biologic proteins (e.g., collagen, hyaluronic acid) or synthetic materials such as polymethylmethacrylate can modify the depth of facial wrinkles and fullness of facial structures such as the lips. Appearance can also be modified using neuromodulators to block facial muscle function to reduce undesirable movements of facial landmarks or deepening of facial wrinkles. Surgical interventions may be employed when the structure and position of facial features require modifications greater than what may be achieved with nonsurgical procedures. Browlift operations raise the position of the eyebrows (Fig. 45-67). Blepharoplasty is a set of procedures that modify the shape and position of the upper and lower eyelids. Facelift modifies the configuration and amount of facial skin and subcutaneous

Figure 45-66. Incisions for cervicofacial rhytidectomy.
Breast reduction surgery reduces the amount of both skin and breast tissue volume and modifies the position of the nipple on the breast mound (Fig. 45-69). The most common indication is to treat symptoms of large breasts known as macromastia, which is associated with a symptomatic triad of upper back pain, bra strap grooving, and skin rashes under the fold of the breasts. Unilateral breast reduction is often performed to achieve breast symmetry after contralateral postmastectomy breast reconstruction. As with all breast surgery, achieving a natural and cosmetically acceptable appearance is essential to a satisfactory outcome. Mastopexy techniques share many aspects with breast reduction except that breast volume is preserved and only the amount of skin and location of the nipple are modified. Fundamental to the success of the procedure is the establishment of symmetric and proper nipple position. Nipple ptosis is graded by the nipple position relative to the inframammary fold.

Many patients seek surgical intervention to increase breast size in a procedure known as augmentation mammoplasty (Fig. 45-70). Breast volume is increased by insertion of a synthetic implant specifically designed for this purpose. Modern breast implants are manufactured from various formulations of silicone polymers. The implant shell, which is in contact with the tissues, is always made from silicone elastomer. The filling material can be either silicone or saline, depending on the patient and surgeon preference. As with any surgical procedure that involves implanting synthetic materials, the surgeon must fully understand the nature of the materials and be able to inform the patient of all known risks and benefits.

The pervasive risk of breast cancer among women mandates careful consideration of the impact of any breast surgery on cancer screening, diagnosis, and treatment. Preoperative breast cancer screening consistent with current American Cancer Society guidelines should be performed for all patients undergoing elective breast reshaping surgery. After breast augmentation surgery, routine screening mammograms are no longer considered adequate. Patients with breast implants must have diagnostic mammograms where a radiologist studies the images at the time of the study to ensure they completely visualize the breast tissue.

Gynecomastia is a condition of excess breast tissue in males. It can be caused by a wide range of medical disorders, including liver dysfunction, endocrine abnormalities, genetic syndromes (e.g., Klinefelter’s syndrome), renal disease, testicular tumors, adrenal or pituitary adenomas, secreting lung carcinomas, and male breast cancer. Pharmacologic agents associated with the potential side effect of breast enlargement include marijuana use, digoxin, spironolactone, cimetidine, theophylline, diazepam, and reserpine. Although all of these possible causes must be considered in any patient presenting with gynecomastia, the majority of patients have idiopathic enlargement of the breast parenchyma, often occurring in teenagers. Surgical correction of this condition as often indicated.

**Aesthetic Surgery of the Body**

Aesthetic surgery may be applied to the torso and extremities. The most common circumstance is following massive weight loss, typically as a result of bariatric surgery. Morbid obesity stretches the skin and supporting ligaments that tether it to the underlying fascial framework. Decreasing the amount of subcutaneous fat often results in significant skin laxity that creates body contour deformities. Improvement can be achieved only through skin excision. Therefore, all body-contouring surgery structures to correct features such as deep nasolabial folds, skin redundancy along the inferior border of the mandible, and loss of definition of neck contours. Rhinoplasty involves a complex set of procedures to modify the size, shape, and airway function of the nose (Fig. 45-68).

**Aesthetic Surgery of the Breast**

Surgery to modify the shape, volume, and nipple position of the breast are among the most common aesthetic procedures.
Figure 45-68. A. Rhinoplasty anatomy. B. Preoperative appearance. C. Postoperative appearance.
represents a trade of excess skin for scar, and this must be emphasized during patient consultation. The patient willing to accept scars in exchange for improved contour is likely to be satisfied with the procedures. With the increased number of bariatric surgery procedures over the past decade, body-contouring surgery has become very popular and is emerging as a new subspecialty of plastic surgery.

Excess skin and subcutaneous tissue on the anterior abdominal wall creates a redundancy that can hang over the pubic area called an abdominal wall pannus. It can cause difficulty dressing and maintaining proper personal hygiene. A panniculectomy is a procedure that removes the redundant skin and subcutaneous tissue of the pannus. If additional contouring of the abdominal wall is performed, the procedure is known as abdominoplasty. During this procedure, not only is the pannus excised but the maximum amount of skin is excised to tighten the abdominal wall. Optimum contouring typically requires tightening of the underlying abdominal wall by suturing the midline and transposing the umbilicus as the upper abdominal skin is advanced inferiorly. At times additional skin must be excised transversely, requiring a concurrent vertical incision to remove skin in two vectors (Fig. 45-71). Possible complications include skin necrosis, persistent paresthesias of the abdominal wall, seroma, and wound separation. Necrosis of the umbilicus may complicate preservation of that structure if the stalk is excessively long or an umbilical hernia is repaired. Adding a

Figure 45-69. Inferior pedicle reduction mammaplasty.
vertical resection increases the incidence of skin necrosis, especially at the confluence of scars in the lower abdomen.

Brachioplasty, or arm lift, excises excess skin and subcutaneous tissue from the arms. It results in improved contour but leaves a visible longitudinal scar on the medial aspect of the arm. Therefore, it is reserved for patients with excessive skin in that region. The patient willing to accept the scar can be happy with the results. Complications include distal seroma and wound separation. Paresthesias in the upper arm and forearm may occur secondary to injury of sensory nerves passing through the resection area, though this rarely affects function. Incisions that cross the axilla must be designed to avoid axillary contractures that limit shoulder mobility.

Thigh and buttock lifts treat loose skin on the thighs and buttocks. A variety of methods have been described, and application requires proper patient selection in order to obtain the best outcome. The lateral thighs can be lifted simultaneously during abdominoplasty with one scar along the belt line. If the lift is continued on the posterior torso, a buttocks lift can be performed as well. This procedure is referred to as a circumferential lower body lift. Contouring the medial thighs typically requires an incision in the groin crease. Firmly anchoring the deep thigh fascia to Colles’ fascia is essential to help prevent spreading of the labia. In cases of severe excess skin on the inner thighs, a long vertical incision is necessary. Complications of thigh and buttock lift include seroma, wound separation, skin necrosis, and change in the shape of the genital region (with possible sexual dysfunction).
Suction Lipectomy

Liposuction is a technique that involves the removal of adipose tissue through minimal incisions using a hollow suction cannula system. The key consideration in determining acceptable candidates for this body contouring technique directly relies on the patient’s inherent skin elasticity, which provides the sought-after retraction of skin over the lipoaspirated adipose depot to improve area contour. Thus, assessment of skin tone is a vital part of the preoperative patient evaluation. If there is excessive skin laxity in the body area to be treated, it may worsen after liposuction and contribute to contour irregularities, voids, and abnormal appearance.

This technique can be highly effective in the correctly chosen patient as the access port sites provide minimally visible scars and can remove significant amounts of fatty tissue to improve contour. However, it is worth mentioning that liposuction is not considered a weight-loss treatment; rather, it is a tool for addressing unwanted and troublesome adipose depots. Typically, the best candidates for liposuction are individuals who are close to their goal weight and have focal adipose deposits that are resistant to diet and exercise (Fig. 45-72). The suction cannula system removes adipose tissue by avulsing fat into the small holes located within the cannula tip. As the cannula is repeatedly passed throughout the adipose planes to remove the fat, one can often visualize and feel the reduction in the fat depot area treated. In general, larger-diameter cannulas remove adipose tissue at a faster rate yet carry a higher risk of causing contour irregularities such as grooving and/or uneven removal of fat. Newer liposuction technologies employing ultrasonic or laser probes to heat and emulsify fat via cavitation before suction are gaining increasing application because they also aid in better tightening of the overlying skin envelope. However, use of these technologies also increases the chance and incidence of tissue damage and injury from the heat of the cannula and can cause burn injury to skin and underlying structures.

A major advance in the field of liposuction involves application of tumescent local anesthesia. This method involves the infiltration of very dilute lidocaine and epinephrine (lidocaine 0.05% and epinephrine 1:1,000,000) in large volumes throughout the subcutaneous tissues prior to suction removal of fatty tissue. Tumescent volumes can range from one to three times the anticipated aspirate volume. The dilute lidocaine provides sufficient anesthesia to allow the liposuction to be performed without additional agents in some instances. However, in cases where large volumes of fat are to be removed or in cases where multiple sites are to be addressed, then sedation and/or general anesthesia is often preferred. With tumescent anesthesia, the absorption of the dilute lidocaine from the subcutaneous tissue is very slow, with peak plasma concentrations occurring approximately 10 hours after the procedure. Therefore, the standard lidocaine dosing limit of 7 mg/kg may be safely exceeded. Current recommendations suggest a limit of 35 mg/kg of lidocaine with tumescent anesthesia. A very important component of the tumescent anesthetic solution is diluted epinephrine, which has proved to limit blood loss during the procedure.

Safety issues are paramount for liposuction because of potential fluid shifts postoperatively and hypothermia. If ≥5000 mL of aspirate is to be removed, the procedure should be
Figure 45-72. A and B. Preoperative photos of a 22-year-old woman with focal adipose deposits on the trunk and extremities. C. Patient 3 months after surgery.
performed in an accredited acute care hospital facility. After the procedure, vital signs and urinary output should be monitored overnight in an appropriate facility by qualified and competent staff familiar with perioperative care of the liposuction patient.

**Autologous Fat Grafting**

The concept of reinjecting fat tissue harvested by liposuction has been put into practice for decades. Key to the technique is a processing step in which the steriley collected fat is separated from the aqueous (primarily tumescent fluid) and free lipid fractions. This can be done by centrifugation and/or filtering. Ideally, the prepared adipose grafts are then injected into the tissues using specially designed blunt-tipped cannulas that provide for micrograft injection. Small aliquots of fat grafts are injected with each cannula pass to deposit the grafts within the vascularized tissues of the recipient bed. Autologous fat grafting has gained increased interest and has been applied to various areas of aesthetic and reconstructive surgery. Specific applications include fat grafting to augment areas where fat atrophy is commonplace (aging of the face or hands), to enhance breast aesthetics and/or other breast reconstruction techniques, gluteal augmentation, or to address contour deformities or irregularities caused by iatrogenic, traumatic, oncologic, or congenital processes.

**REFERENCES**

Entries highlighted in bright blue are key references.


   This is a modern publication of the classic 18th century work by Guiseppi Baronio who studied skin grafting in animals. Baronio’s work represents the first preclinical animal study of a surgical procedure. The logo of the most important professional organization dedicated to plastic surgery research, the Plastic Surgery Research Council, is based on Baronio’s illustration of a sheep with multiple grafted areas of skin on the back.


   A variety of skin substitutes are available for repairing areas of skin loss from injuries such as deep partial-thickness or full-thickness burns. This article provides a nice summary of contemporary options.


   This is the classic article studying blood supply to the skin that introduced the angiosome concept and transformed our understanding of the anatomic basis of surgical flap design. The blood supply was shown to be a continuous three-dimensional network of vessels in all tissue layers. The anatomical territory of a source artery corresponded in both the skin and deep tissues and gave rise to the angiosome concept.


   This is an excellent summary of the basic principles of wound healing. It explains the physiologic basis and rationale for various wound care methods, including dressings, negative pressure wound therapy, skin and dermal substitutes, and tissue expansion. This is basic knowledge that is important for all surgeons to understand.


   This is the definitive textbook on pediatric plastic surgery that covers each aspect in depth.


BRIEF HISTORY OF ANESTHESIA

The discovery of anesthesia is one of the seminal American contributions to the world. Along with infection control and blood transfusion, anesthesia has enabled surgery to occupy its fundamental place in medicine. Before the advent of anesthesia in the 1840s, many substances and methods had been tried in the search for pain relief and better operating conditions. Patients were typically restrained by several attendants, and only the most stoic could tolerate the screams heard in the operating theater.

Beginnings

Horace Wells (1815–1848), a dentist, first pursued using nitrous oxide for the relief of pain in surgical procedures in 1844. After experimenting on himself, Wells attempted to demonstrate the analgesic effects of nitrous oxide for a dental procedure at Harvard Medical School in 1845. The public demonstration was a failure, at least partially, due to improper administration of the gas. Wells never recovered from his humiliating experience and eventually committed suicide. However, he does hold a place in history as the first person to recognize and use the only anesthetic from the 1800s that is still in use today—nitrous oxide.

Ether Day

William Morton (1819–1868) was a dentist and partner of Horace Wells. After taking a course in anesthesia from Wells, Morton left the partnership in Hartford, Connecticut, and established himself in Boston. He continued his interest in anesthesia, but using diethyl ether instead of nitrous oxide. Ether proved a good choice. He practiced the administration of ether on a dog and then used it when extracting teeth from patients in his office. On October 16, 1846, Morton gave the first public demonstration of ether as an anesthetic for Johns Collins Warren, a distinguished surgeon and one of the founders of Massachusetts General Hospital. In attendance in the surgical amphitheater were several surgeons, medical students, and a newspaper reporter. After induction of anesthesia, Warren successfully removed a vascular mass from the patient’s neck with no ill effects (Fig. 46-1). The description of this public demonstration of ether was published in the Boston Medical and Surgical Journal (now The New England Journal of Medicine). The stature of Warren lent considerable credence to the advent of surgical anesthesia. The news spread rapidly, and surgeons around the world were quick to adopt this new invention. Massachusetts General Hospital has restored and preserved the original amphitheater where the demonstration took place, now called the Ether Dome. The description of the public demonstration of ether was voted as the most important article published in the history of The New England Journal of Medicine in its first 200 years.

The Modern Era

Anesthesia has developed rapidly over the past century. Inhaled anesthetics, initially discovered fortuitously by observation, have been synthetically produced and remain the mainstay of anesthetic maintenance. The advent of the hollow syringe and needle and discovery of rapidly acting intravenous anesthetics allowed for rapid induction of anesthesia. Development of endotracheal intubation and mechanical ventilation revolutionized the delivery of inhaled anesthetics. The discovery of local anesthetics led to the development of peripheral nerve blocks...
Key Points

1. The discovery of anesthesia was one of the most important advances and has enabled surgery to occupy its fundamental place in medicine.

2. Advances in anesthetic monitoring have made the administration of anesthesia safer than ever. Types of cardiovascular monitors include arterial catheters, central venous and pulmonary artery catheters, and transesophageal echocardiography.

3. A detailed preoperative evaluation should be performed on each patient when circumstances allow, with special attention devoted to functional status. The American College of Cardiology/American Heart Association guidelines for preoperative evaluation can guide workup.

4. The American Society of Anesthesiologists has developed specific guidelines for preoperative fasting to mitigate the risk of aspiration of gastric contents; individual patients may need more stringent preoperative fasting periods and/or rapid sequence inductions.

5. The American Society of Anesthesiologists has developed an algorithm for management of the difficult airway. Notably, in patients in whom both intubation and ventilation are impossible, the algorithm calls for placement of a laryngeal mask airway as the next step.

and spinal anesthesia. Concurrently, physiologic monitoring techniques have advanced to make the administration of anesthesia safer than ever.

Initially, anesthesia was given by medical students, nurses, and dentists, but eventually became a physician specialty of medicine of its own. The American Board of Anesthesiology was formed in 1938. Over the past 50 years, anesthesiology has increasingly specialized and also spread outside the operating room into critical care, pain management, and perioperative medicine.

BASIC PHARMACOLOGY

Pharmacokinetics and Pharmacodynamics
Pharmacodynamics is the study of what a drug does to the body; pharmacokinetics is the study of what the body does to a drug.
The conduct of anesthesia is predicated upon the pharmacodynamics and pharmacokinetics of the drugs used.\(^4\)

**Administration, Distribution, Metabolism, and Elimination**

*Administration* of a drug affects its pharmacokinetics, as there will be different rates of drug entry into the circulation. For example, medications administered via the oral route are subject to first-pass effect of the portal circulation; this can be bypassed with the IV, nasal, or sublingual route. Other routes of drug administration include transdermal, intramuscular, subcutaneous, or inhalation.

*Distribution* is the delivery of a drug from the systemic circulation to the tissues. Once a drug has entered the systemic circulation, the rate at which it will enter the tissues depends on the blood flow into that tissue, as well as the molecular size of the drug, lipid solubility, capillary permeability, polarity, plasma protein and tissue binding, and volume of distribution, the fluid volume in which the drug distributes. The distribution or redistribution of drugs can play a critical role in shaping their clinical use. For instance, clinically, the effect of propofol is terminated by its redistribution into fatty tissues and not metabolism of the drug.

*Metabolism* is the permanent breakdown of original compounds into smaller metabolites. Drug *elimination* varies widely; some drugs are excreted unchanged by the body, some decompose via plasma enzymes, and some are degraded in the liver. Many drugs rely on multiple pathways for metabolism and elimination (i.e., metabolized by liver enzymes and then excreted by the kidney).

*Context-sensitive half time* is the time required for blood concentrations of a drug to decrease by 50% after its discontinuation, which is determined by the interaction of the duration of administration, distribution and accumulation, and metabolism and excretion. Fig. 46-2 illustrates the context sensitive half-time for commonly used anesthetics and opioids.\(^4\)

**Pharmacodynamics**

Pharmacodynamics, or how the plasma concentration of a drug translates into its effect on the body, depends on biologic variability, receptor physiology, and clinical evaluations of the actual drug. An agonist is a drug that causes a response (activates a receptor). A full agonist produces the full receptor/tissue response, and a partial agonist elicits less than the maximum response induced by a full agonist. An antagonist is a drug that blocks agonist mediated responses. An additive effect means that a second drug acts with the first drug and will produce an effect that is equal to the algebraic summation of both drugs. A synergistic effect means that two drugs interact to produce an effect that is greater than expected from the two drugs’ algebraic summation. Tolerance, desensitization, or tachyphylaxis occurs when a larger than expected dose is required to produce a response. Tolerance usually results from chronic drug exposure, either through enzyme induction (e.g., alcohol) or depletion of neurotransmitters (e.g., cocaine).\(^4\)

**Potency, Efficacy, Lethal Dose, and Therapeutic Index**

The *potency* of a drug is the dose required to produce a given effect, such as pain relief or a change in heart rate. The average sensitivity to a particular drug can be expressed through the calculation of the effective dose; ED50 would have the desired effect in 50% of the general population. The *efficacy* of any therapeutic agent is its power to produce a desired effect. Two drugs may have the same efficacy but different potencies.

Dose-response curves show the relationship between the dose of a drug administered (or the resulting plasma concentration) and the pharmacologic effect of the drug. The lethal dose (LD50) of a drug produces death in 50% of animals to which it is given, and the toxic dose (TD50) is the dose that elicits a toxicity in 50% of humans to which it is given. The ratio of the toxic dose and effective dose, TD50/ED50, is the therapeutic index. A drug with a high therapeutic index is safer than a drug with a low or narrow therapeutic index.\(^4\)

**ANESTHETIC AGENTS**

**Inhaled Anesthetics**

Inhaled anesthetics have greatly advanced since the original demonstration with ether. Modern agents provide faster induction and emergence and provide all of the major characteristics of general anesthesia: unconsciousness, analgesia, and muscle relaxation.

*Minimum alveolar concentration (MAC)* is a measure of anesthetic potency. It is the ED50 of an inhaled agent (i.e., the dose required to prevent movement in response to skin incision in 50% of patients). The higher the MAC, the less potent an agent is. Advantages and disadvantages of inhaled anesthetics are shown in Table 46-1.\(^5\)

**Nitrous Oxide.** Nitrous oxide has a low solubility and is a weak anesthetic agent, but it has the most rapid onset and offset. Because
Advantages and disadvantages of inhaled anesthetics

<table>
<thead>
<tr>
<th>ANESTHETIC</th>
<th>ADVANTAGES</th>
<th>DISADVANTAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nitrous Oxide</td>
<td>No odor, taste, or pungency</td>
<td>Airspace expansion</td>
</tr>
<tr>
<td></td>
<td>Rapid uptake and elimination</td>
<td>Increased nausea</td>
</tr>
<tr>
<td></td>
<td>Analgesic effect</td>
<td>and vomiting</td>
</tr>
<tr>
<td></td>
<td>Minimal cardiovascular depression</td>
<td>Inhibits methionine synthase</td>
</tr>
<tr>
<td></td>
<td>Minimal biotransformation</td>
<td>Environmental pollutant</td>
</tr>
<tr>
<td></td>
<td>Inexpensive</td>
<td>Supports combustion</td>
</tr>
<tr>
<td>Isoflurane</td>
<td>Good muscle relaxation</td>
<td>Slow uptake and elimination</td>
</tr>
<tr>
<td></td>
<td>Bronchodilation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stable heart rate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inexpensive</td>
<td></td>
</tr>
<tr>
<td>Sevoflurane</td>
<td>Rapid uptake and elimination</td>
<td>Breakdown to compound A in circuit</td>
</tr>
<tr>
<td></td>
<td>Not pungent</td>
<td>More expensive than isoflurane</td>
</tr>
<tr>
<td>Desflurane</td>
<td>Rapid uptake and elimination</td>
<td>Airway irritant</td>
</tr>
<tr>
<td></td>
<td>Very low biotransformation</td>
<td>Requires electric/heated vaporizer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Expensive</td>
</tr>
</tbody>
</table>

Table 46-1


Barbiturates Barbiturates used in anesthesia include thiopental and methohexital. These drugs act as agonists at the γ-aminobutyric acid (GABA) receptor, which inhibit excitatory synaptic transmission. Clinically, they produce a rapid, smooth induction of general anesthesia and wear off quickly. They cause hypotension and myocardial depression in a dose-dependent manner. Barbiturates are seldom used in modern anesthesia, with the exception of methohexital, which is still commonly used during electroconvulsive therapy.

Propofol Propofol is an alkylated phenol that inhibits synaptic transmission through its effects at the GABA receptor. It has a short duration, rapid recovery, and low incidence of nausea and vomiting. Consequently, it is the induction agent of choice. Propofol causes hypotension in a dose-dependent manner, and it should be used cautiously in patients with cardiac disease or hypovolemia. Continuous infusion of propofol is commonly used for sedation in the intensive care unit setting. Continuous infusions of propofol are also used for moderate-to-deep sedation for many procedures and are also commonly incorporated into total intravenous anesthetics (TIVA), most commonly for neurosurgical procedures. Propofol is an irritant and frequently causes pain on injection. Propofol also has anticonvulsant properties.

Benzodiazepines Benzodiazepines are most commonly used to reduce anxiety and produce amnesia. Midazolam, which has a rapid onset and relatively short duration of action, is by far the most commonly used benzodiazepine in anesthesia. Lorazepam and diazepam are still sometimes used as anxiolytics or amnestic. Benzodiazepines act as agonists at the GABA$_\text{A}$ receptor. They produce sedation, vasodilation, and respiratory depression in a dose-dependent manner. They should be used with caution when given with opioids because a synergistic reaction causing respiratory depression is common. Oral midazolam is commonly used for axiolyis in children. Benzodiazepines are excellent anticonvulsants and only rarely cause allergic reactions. Benzodiazepines should be administered cautiously in older adult patients due to the heightened risk of delayed awakening and postoperative delirium.

Etomidate Etomidate is an imidazole derivative used for IV induction. Its rapid and almost complete hydrolysis to inactive metabolites results in rapid offset. Like the IV agents mentioned earlier, etomidate acts on the GABA receptor. Etomidate has little direct effect on cardiac output and heart rate; induction doses thus cause less reduction in blood pressure than seen with propofol. Etomidate is associated with pain on injection. Notably, etomidate causes adrenal suppression, although whether a single dose of etomidate given at induction causes clinically relevant adrenal suppression remains controversial.

Dexmedetomidine Dexmedetomidine is an IV α2-adrenergic agonist, administered as a continuous infusion, and has both sedative and analgesic properties. It is useful for sedation in an intensive care unit setting and as an adjunct to general anesthesia, especially as part of a total intravenous anesthetic. Side effects include hypotension and bradycardia in a dose-dependent manner. It does not cause respiratory depression at commonly used doses and is thus particularly useful for procedural sedation for patients at high risk of respiratory complications. It is synergistic with opioids and thus can be used to facilitate an opiate-sparing anesthetic.

Ketamine Ketamine differs from the aforementioned IV agents in that it produces analgesia as well as amnesia. Its principal
action is on the N-methyl-D-aspartate (NMDA) receptor. It is a dissociative anesthetic, producing a cataleptic gaze with nystagmus. Patients may experience delirium and hallucinations while regaining consciousness. The addition of benzodiazepines has been shown to reduce the incidence of these side effects. Ketamine typically increases heart rate and blood pressure, which may cause myocardial ischemia in patients with coronary disease. Ketamine is often used in acutely hypovolemic patients to maintain blood pressure via sympathetic stimulation. Importantly, ketamine is a direct myocardial depressant in patients who are catecholamine depleted, and it can produce profound hypotension and low cardiac output in such patients. Ketamine is a bronchodilator and is sometimes used as an induction agent in asthmatic patients. It can increase intracranial pressure and intraocular pressure, and thus its use in patients with trauma to the head and neck is controversial. Ketamine can be administered intramuscularly to induce anesthesia in patients who would not tolerate an inhalational induction or IV placement, such as patients with developmental delay.

**Opioid Analgesics** The commonly used opioids—morphine, codeine, hydromorphone, meperidine, and the fentanyl-based compounds—act on μ-receptors in the brain and spinal cord. The main side effects of opioids are euphoria, sedation, constipation, and respiratory depression, which also are mediated by μ-receptors in a dose-dependent fashion. Although opioids have differing potencies required for effective analgesia, equianalgesic doses of opioids result in equal degrees of respiratory depression. Thus, there is no completely safe opioid analgesic, and no reason to suppose that one opioid is safer than another. The synthetic opioid fentanyl and its analogues sufentanil, alfentanil, and remifentanil are used in the operating room. They differ pharmacokinetically in their lipid solubility, tissue binding, and elimination profiles and thus have differing potencies and durations of action. Fentanyl, which is highly lipidsoluble, accumulates in tissues and exhibits a steep increase in its context-sensitive half-time with infusions. Remifentanil is remarkable in that it undergoes rapid hydrolysis that is unaffected by sex, age, weight, or renal or hepatic function, even after prolonged infusion. Alfentanil and sufentanil are seldom used, having largely been replaced by remifentanil in the modern era. Morphine and meperidine have active metabolites that are renally excreted and thus should be used with caution or avoided in patients with renal insufficiency.

Naloxone, an opioid antagonist, can be used to rapidly reverse the effects of opioids, and is commonly used to rescue patients from opioid-associated respiratory depression. Naloxone is poorly absorbed orally and is also often combined with oral opioids to prevent abuse by injection use of the combined drug. Methylaltrexone and alvimopan are both peripheral opioid antagonists and can reverse the opioid side effect of constipation without affecting analgesia.¹⁶,¹⁷

**Nonopioid Analgesics** Ketorolac is a parenteral nonsteroidal anti-inflammatory drug (NSAID) that produces analgesia by reducing prostaglandin formation via inhibition of the enzyme cyclooxygenase (COX). Intraoperative use of ketorolac reduces postoperative need for opioids. Ketorolac along with other NSAIDs can cause major side effects, including bleeding, platelet dysfunction, and acute kidney injury and should be used cautiously in elderly patients or patients with renal insufficiency.

Acetaminophen is an analgesic drug and antipyretic; its site of action is in the central nervous system. Use of acetaminophen has been shown to reduce opioid requirements postoperatively. Long available in an orally administered form as well as a rectal suppository, an intravenous formulation of acetaminophen is now available which has become widely used in the postoperative setting.¹⁸

**Lidocaine** is a local anesthetic commonly used for local infiltration, nerve blocks, or epidural infusions. Recently, intravenous infusions of lidocaine have been shown to be beneficial in the perioperative period. A large meta-analysis of 42 trials with 2800 patients showed that intravenous lidocaine infusions modestly reduced postoperative pain, reduced opioid requirements, and shortened time to recovery of bowel function for patients undergoing abdominal surgery.¹⁹

**Local Anesthetics** Local anesthetics act on sodium channels to block transmission of neural impulses. They are divided into two groups based on their chemical structure: the amides and the esters. In general, the amides are metabolized in the liver, and the esters are metabolized by plasma cholinesterases, which yield metabolites with slightly higher allergic potential than the amides. Amides include lidocaine, bupivacaine, mepivacaine, prilocaine, and ropivacaine. Lidocaine has a fairly rapid onset and is shorter acting. Ropivacaine and bupivacaine have a slower onset and are longer lasting. All three are commonly used for local infiltration and regional nerve blocks. Amides are 95% metabolized in the liver, with a minority excreted unchanged in the kidneys. Prilocaine and mepivacaine are seldom used in anesthesia at present. Esters include cocaine, procaine, chloroprocaine, tetracaine, and benzocaine. Esters are hydrolyzed in the blood by plasma esterases.

When used in large quantities over a short period of time, local anesthetic levels can rise in the blood and cause central nervous system (CNS) toxicity and cardiovascular toxicity. Symptoms of CNS toxicity include Restlessness, tinnitus, and slurred speech and can progress to seizures and coma. Cardiovascular toxicity may manifest as hypotension, conduction abnormalities leading to heart block, and ventricular arrhythmias, and it may lead to cardiac arrest. The type of local anesthetic used affects the risk of developing toxicity; bupivacaine is most often associated with cardiovascular toxicity. Other risk factors for local anesthetic systemic toxicity include cumulative dose, site of injection, and preexisting renal, hepatic, or cardiac disease in the patient.²⁰ In addition to treating symptomatology, local anesthetic systemic toxicity can be treated with intravenous administration of lipid emulsion.²¹

**Neuromuscular Blockers** While general anesthetics provide muscle relaxation, they usually do so at a much deeper anesthetic depth than required for amnesia and hypnosis. For this reason, neuromuscular blockers are commonly administered to attain adequate relaxation at levels of anesthesia sufficient to produce hypnosis, amnesia, and analgesia. Neuromuscular blockers block conduction at the neuromuscular junction of skeletal muscle.²²

The two categories of neuromuscular blockers in use are depolarizing and nondepolarizing blockers. Characteristics of neuromuscular blockers currently used are summarized in Table 46-2. Succinylcholine is the only depolarizing agent used currently. It binds to acetylcholine receptors on the postjunctional membrane in the neuromuscular junction and causes depolarization of muscle fibers. The rapid onset (less than 60 seconds) and rapid offset (5–8 minutes) of succinylcholine make it ideal for management of the airway in certain
situations.23 Succinylcholine has several adverse effects including transient hyperkalemia, which can be severe or even fatal for patients with burns and denervating injuries. Succinylcholine can cause bradycardia, which can be severe in children. It is also associated with transient increases in intracranial and intraocular pressure. The depolarization caused by succinylcholine causes skeletal muscles to fasciculate, which in turn, can result in postoperative myalgias. Succinylcholine is a known trigger of malignant hyperthermia in susceptible individuals. Succinylcholine is broken down by pseudocholinesterase; patients who are homozygous for pseudocholinesterase deficiency will have prolonged muscle paralysis. Over the past several decades, advancement in anesthetic monitoring has made administration of anesthesia safer than ever. The goal of anesthetic monitoring is to continuously monitor the patients’ cardiovascular status, pulmonary status, respiratory physiology, anesthetic depth, concentration of gases administered, and temperature. The American Society of Anesthesiology (ASA) has established standards for basic intraoperative monitoring that are listed in Table 46-3. Types of anesthetic monitoring are listed in Table 46-4.

**Cardiovascular monitoring** includes continuous ECG monitoring as well as blood pressure monitoring, which is to be measured and recorded at least every 5 minutes. Blood pressure monitoring can be done using noninvasive blood pressure cuff measurements or invasively using an arterial catheter. Other cardiovascular monitors include monitoring of central venous pressure, pulmonary artery pressure, and cardiac output. In high-risk situations,23 succinylcholine has several adverse effects including transient hyperkalemia, which can be severe or even fatal for patients with burns and denervating injuries. Succinylcholine can cause bradycardia, which can be severe in children. It is also associated with transient increases in intracranial and intraocular pressure. The depolarization caused by succinylcholine causes skeletal muscles to fasciculate, which in turn, can result in postoperative myalgias. Succinylcholine is a known trigger of malignant hyperthermia in susceptible individuals. Succinylcholine is broken down by pseudocholinesterase; patients who are homozygous for pseudocholinesterase deficiency will have prolonged neuromuscular blockade, typically lasting for several hours.

There are several nondepolarizing neuromuscular blocking agents in clinical use. Long-acting agents including pancuronium are no longer widely used. Intermediate-duration neuromuscular blockers include the steroid-based drugs vecuronium and rocuronium, which are metabolized by the liver as well as by the kidney, and the benzylisoquinoline drugs atracurium and cisatracurium, which undergo breakdown in plasma known as Hofmann elimination. All nondepolarizers reversibly bind to the postsynaptic terminal in the neuromuscular junction and prevent acetylcholine from depolarizing the muscle. Muscle blockade occurs without fasciculation and without the subsequent side effects seen with succinylcholine. Neuromuscular blockade with nondepolarizing drugs is typically reversed. Failure to adequately reverse neuromuscular blockade is associated with an increased risk of perioperative respiratory failure and death. Reversal agents include acetylcholinesterase inhibitors including neostigmine, edrophonium, or pyridostigmine that are given concurrently with muscarinic-anticholinergic, almost always atropine or glycopyrrolate. Recently, sugammadex, a chelating agent, has been approved for use as a reversal agent for reversal of neuromuscular blockade by the steroid paralytics rocuronium and vecuronium. When given at a very high dose, sugammadex can even rapidly reverse the effect of an intubating dose of a steroid neuromuscular blocker.

### ANESTHETIC MONITORING

Over the past several decades, advancement in anesthetic monitoring has made administration of anesthesia safer than ever. The goal of anesthetic monitoring is to continuously monitor the patients’ cardiovascular status, pulmonary status, respiratory physiology, anesthetic depth, concentration of gases administered, and temperature. The American Society of Anesthesiology (ASA) has established standards for basic intraoperative monitoring that are listed in Table 46-3. Types of anesthetic monitoring are listed in Table 46-4.

**Cardiovascular monitoring** includes continuous ECG monitoring as well as blood pressure monitoring, which is to be measured and recorded at least every 5 minutes. Blood pressure monitoring can be done using noninvasive blood pressure cuff measurements or invasively using an arterial catheter. Other cardiovascular monitors include monitoring of central venous pressure, pulmonary artery pressure, and cardiac output. In high-risk situations, succinylcholine has several adverse effects including transient hyperkalemia, which can be severe or even fatal for patients with burns and denervating injuries. Succinylcholine can cause bradycardia, which can be severe in children. It is also associated with transient increases in intracranial and intraocular pressure. The depolarization caused by succinylcholine causes skeletal muscles to fasciculate, which in turn, can result in postoperative myalgias. Succinylcholine is a known trigger of malignant hyperthermia in susceptible individuals. Succinylcholine is broken down by pseudocholinesterase; patients who are homozygous for pseudocholinesterase deficiency will have prolonged neuromuscular blockade, typically lasting for several hours.

There are several nondepolarizing neuromuscular blocking agents in clinical use. Long-acting agents including pancuronium are no longer widely used. Intermediate-duration neuromuscular blockers include the steroid-based drugs vecuronium and rocuronium, which are metabolized by the liver as well as by the kidney, and the benzylisoquinoline drugs atracurium and cisatracurium, which undergo breakdown in plasma known as Hofmann elimination. All nondepolarizers reversibly bind to the postsynaptic terminal in the neuromuscular junction and prevent acetylcholine from depolarizing the muscle. Muscle blockade occurs without fasciculation and without the subsequent side effects seen with succinylcholine. Neuromuscular blockade with nondepolarizing drugs is typically reversed. Failure to adequately reverse neuromuscular blockade is associated with an increased risk of perioperative respiratory failure and death. Reversal agents include acetylcholinesterase inhibitors including neostigmine, edrophonium, or pyridostigmine that are given concurrently with muscarinic-anticholinergic, almost always atropine or glycopyrrolate. Recently, sugammadex, a chelating agent, has been approved for use as a reversal agent for reversal of neuromuscular blockade by the steroid paralytics rocuronium and vecuronium. When given at a very high dose, sugammadex can even rapidly reverse the effect of an intubating dose of a steroid neuromuscular blocker.

### Table 46-2

<table>
<thead>
<tr>
<th>AGENT</th>
<th>TYPE</th>
<th>INTUBATING DOSE</th>
<th>CONSIDERATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Succinylcholine</td>
<td>Depolarizer</td>
<td>1 mg/kg</td>
<td>Can cause severe hyperkalemia, contraindicated in burns, denervating conditions, excessive or prolonged use can lead to phase II block.</td>
</tr>
<tr>
<td>Rocuronium</td>
<td>Nondepolarizer</td>
<td>0.6 mg/kg</td>
<td>Primarily hepatic metabolism, can be reversed with sugammadex or acetylcholinesterase inhibitor.</td>
</tr>
<tr>
<td>Vecuronium</td>
<td>Nondepolarizer</td>
<td>0.1 mg/kg</td>
<td>Primarily hepatic metabolism, can be reversed with sugammadex or acetylcholinesterase inhibitor.</td>
</tr>
<tr>
<td>Cisatracurium</td>
<td>Nondepolarizer</td>
<td>0.1 mg/kg</td>
<td>Hoffman degradation, can be reversed with an acetylcholinesterase inhibitor.</td>
</tr>
</tbody>
</table>

RSI = rapid sequence induction

### Table 46-3

<table>
<thead>
<tr>
<th>American Society of Anesthesiologists standards for basic intraoperative monitoring</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standards</strong></td>
</tr>
<tr>
<td>Standard 1: Qualified anesthesia personnel shall be present in the room throughout the conduct of all general anesthetics and regional and monitored anesthesia care.</td>
</tr>
<tr>
<td>Standard 2: Oxygenation, ventilation, circulation, and temperature shall be continually evaluated</td>
</tr>
<tr>
<td>Oxygenation&lt;br&gt; Inspired gas oxygen analyzer&lt;br&gt; Pulse oximetry&lt;br&gt; Monitoring of patient clinical status</td>
</tr>
<tr>
<td>Ventilation&lt;br&gt; Auscultation&lt;br&gt; Observation of the patient&lt;br&gt; Observation of reservoir bag&lt;br&gt; End-tidal carbon dioxide analysis</td>
</tr>
<tr>
<td>Circulation&lt;br&gt; Continuous electrocardiogram display&lt;br&gt; Heart rate and blood pressure recorded at least every 5 minutes&lt;br&gt; Evaluation of circulation: auscultation of heart sounds, palpation of pulse, pulse oximetry, blood pressure monitoring with noninvasive means or intra-arterial catheter pressure measurement</td>
</tr>
<tr>
<td>Temperature&lt;br&gt; Core and/or skin temperature</td>
</tr>
</tbody>
</table>

Table 46-4

Types of anesthesia monitors and their properties

<table>
<thead>
<tr>
<th>TYPE OF MONITOR</th>
<th>WHAT IS MEASURED</th>
<th>INVASIVENESS</th>
<th>POTENTIAL FOR COMPLICATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical examination</td>
<td>Heart sounds, breath sounds, pulse, color, mental status, etc</td>
<td>Noninvasive</td>
<td>–</td>
</tr>
<tr>
<td>Pulse oximetry</td>
<td>Arterial oxygen saturation</td>
<td>Noninvasive</td>
<td>–</td>
</tr>
<tr>
<td>Arterial catheter</td>
<td>Blood pressure, acid/base status</td>
<td>Invasive</td>
<td>++</td>
</tr>
<tr>
<td>Noninvasive blood pressure measurement</td>
<td>Blood pressure</td>
<td>Noninvasive</td>
<td>+/-</td>
</tr>
<tr>
<td>Electrocardiography</td>
<td>Cardiac rhythm, rate, ST segments</td>
<td>Noninvasive</td>
<td>–</td>
</tr>
<tr>
<td>Capnography</td>
<td>Ventilatory, circulatory status</td>
<td>Noninvasive</td>
<td>–</td>
</tr>
<tr>
<td>Electroencephalogram, bispectral index, etc</td>
<td>Brain function, depth of anesthesia</td>
<td>Noninvasive</td>
<td>+/-</td>
</tr>
<tr>
<td>Temperature probe</td>
<td>Body temperature</td>
<td>Noninvasive to invasive</td>
<td>+/-</td>
</tr>
<tr>
<td>Central venous pressure, pulmonary artery pressure</td>
<td>Cardiac function, volume status</td>
<td>Invasive</td>
<td>+++</td>
</tr>
<tr>
<td>Transesophageal echocardiogram</td>
<td>Cardiac function, volume status</td>
<td>Invasive</td>
<td>+++</td>
</tr>
</tbody>
</table>


anesthetics such as liver transplantation and cardiac surgery, transesophageal echocardiography (TEE) is employed to monitor myocardial function and volume status. Intraoperative TEE can also be used to guide surgeons when performing complex cardiac surgeries, including cardiac valve replacements.

Monitoring of oxygenation and ventilation includes use of continuous pulse oximetry, monitoring of exhaled end-tidal carbon dioxide (ETCO2), and monitoring of fraction of inspired oxygen. End tidal CO2 monitoring also provides important information about systemic perfusion. During cardiac arrest, there is no delivery of CO2 to the lungs, and the end-tidal CO2 is thus very low or zero; a sudden spike in end tidal CO2 during cardio-pulmonary resuscitation correlates with return of spontaneous circulation. Modern ventilators also measure peak and plateau inspiratory airway pressure and minute ventilation. Adequacy of oxygenation and ventilation can also be confirmed by arterial blood gas analysis.

Temperature monitoring is performed using a temperature probe, usually inserted in the esophagus or nasopharynx. Core body temperature can be measured with temperature sensing Foley catheters. Temperature can also be measured at the skin.

Several monitors exist that measure depth of anesthesia, including the bispectral index (BIS) monitor and the SedLine monitor. While these monitors were designed to prevent awareness under anesthesia, a multicenter trial of over 6000 patients showed that titrating anesthetic concentration to the BIS monitor was not superior to titrating anesthetic depth to end-tidal anesthetic concentration with goal MAC greater than 0.7. Peripheral nerve stimulators should be used to monitor depth of neuromuscular blockade. A train-of-four monitor delivers four successive stimuli over 2 seconds. Presence of four twitches without fade with a ratio of the height of the first twitch to the height of the fourth twitch at least 0.9 suggests adequate reversal of neuromuscular blockade. The presence of one or two twitches (absence of the last two or three) is generally sufficient for the relaxation required for almost any kind of abdominal or thoracic operation.

PERIOPERATIVE EVALUATION AND PREPARATION

The ASA has adopted basic standards for the evaluation of patients before surgery. These standards require the anesthesiologist to evaluate the medical status of the patient, develop a plan of anesthetic care, and discuss this plan with the patient and/or the patient’s legal guardian.

A preoperative evaluation includes an appropriately detailed medical history, current drug therapy, appropriate physical examination, and review of laboratory and specific testing results. Based on these findings, the anesthesiologist may conclude that a patient is not in optimal medical condition to undergo elective surgery. These findings and opinions are then discussed with the patient’s primary physician or surgeon, and the surgery may be delayed (or cancelled) until the patient’s medical condition is further evaluated and optimized.

The medical history obtained at the preoperative visit should include the patient’s previous exposure and experience with anesthesia, as well as any family history of problems with anesthesia. History of atopy is an important aspect of this evaluation in that it may predispose patients to form antibodies against antigens that may be represented by agents administered during the perioperative period. Concurrent medications should be fully evaluated when circumstances allow, and adverse interactions with agents administered during the perioperative period need to be considered. A review of the function of major organ systems should also be performed. The physical examination is targeted primarily at the central nervous system, cardiovascular system, lungs, and airway.

Laboratory testing should be based on the patient’s condition and the proposed procedure. Otherwise healthy patients usually do not need laboratory testing for minor procedures. Preoperative testing may be necessitated by findings on physical examination; for example, an electrocardiogram should be obtained if an irregular heart rhythm is noted, and an echocardiogram may be indicated if a new murmur is observed on
auscultation. Chest imaging or pulmonary function testing may be indicated if abnormalities are noted on pulmonary examination and may be pertinent to the administration of the anesthetic or the recovery from anesthesia and surgery. Urine pregnancy testing is typically performed on the day of surgery in women of childbearing age.

ASA Physical Status Assessment
The ASA classification system is a scale used to risk-stratify patients for anesthesia and surgery. The scale, ranging from physical status I to VI, is shown in Table 46-5. Patients undergoing emergent surgery are denoted by an “E”; for example, an otherwise healthy patient undergoing an appendectomy for appendicitis would be classified as ASA IE. Mortality has shown to increase with increasing ASA physical status, and it has been shown to be higher for patients undergoing emergency surgery.27

Airway Evaluation
Airway examination can identify most patients in whom management of the airway and conventional endotracheal intubation may be difficult. It is vitally important to recognize such patients before administering medications that induce apnea. The Mallampati classification (Fig. 46-3) is based on the structures visualized with maximal mouth opening and tongue protrusion in the sitting position.28 Patients with higher Mallampati classification, in combination with other airway abnormalities, can be difficult to intubate. Other predictors of difficult intubation include short neck, immobility of the neck, a large overbite, a small mandible, or the inability to shift the lower incisors in front of the upper incisors. The thyromental distance, the distance from the thyroid cartilage to the tip of the chin should be greater than 6 cm; thyromental distance of less than 6 cm has been associated with difficult intubation.30 Obesity is also a risk factor for difficult intubation, and neck circumference has been identified as a risk factor for both difficult intubation as well as difficult mask ventilation.31

Cardiovascular Disease
Cardiac risk is widely regarded as the most important risk associated with anesthesia and surgery, and it has been the focus of an enormous amount of scholarship over the past four decades. The revised cardiac risk index incorporates six patient and surgical factors to assess a patient’s risk of major adverse cardiac events in the perioperative period: history of ischemic heart disease, congestive heart failure, cerebrovascular disease, diabetes requiring insulin, chronic kidney disease with baseline creatinine greater than 2, and whether the surgery is in a high-risk area, namely major vascular, intraperitoneal, or intrathoracic. In 2014, the American College of Cardiology and the American Heart Association published guidelines for perioperative workup and management of patients with cardiovascular disease; a simplified version is seen in Fig. 46-4.

Notably, this guideline stresses the importance of functional status in determining need for further evaluation; patients with good functional status can typically proceed to surgery without additional evaluation. Functional capacity is measured in metabolic equivalents (METs), with patients unable to attain 4 METs considered to have poor functional status. Activities representing 4 METs including walking up a flight of stairs, climbing a hill, or walking on level ground at 3 to 4 miles per hour.32

![Mallampati Classification Diagram](image_url)

**Figure 46-3.** The Mallampati classification.

<table>
<thead>
<tr>
<th>CLASS</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Soft palate, fauces, uvula, pillars</td>
</tr>
<tr>
<td>2</td>
<td>Soft palate, fauces, portion of uvula</td>
</tr>
<tr>
<td>3</td>
<td>Soft palate, base of uvula</td>
</tr>
<tr>
<td>4</td>
<td>Hard palate only</td>
</tr>
</tbody>
</table>

These recent guidelines have minimized the role of routine screening. Preoperative electrocardiograms and stress testing are unnecessary for asymptomatic patients undergoing low-risk surgery.

Special attention is required for patients with coronary stents. Elective surgery should be delayed a stent has been inserted, which allows time for it to stabilize and the risk of in-stent thrombosis to decrease. The ACC/AHA guidelines recommend delaying elective surgery for 30 days after bare metal stent placement and for 1 year after drug eluting stent placement. Dual antiplatelet therapy should be continued for urgent or emergent procedures that take place before the minimum recommended waiting period.\textsuperscript{32} For semi-elective surgeries in patients with drug-eluting stents, where the risk of delaying surgery is greater than the risk of in-stent thrombosis, ACC/AHA guidelines recommend surgery be delayed for 180 days.

Current recommendation is that β-blockers and statins should be continued in patients who are on them chronically. β-Blockers may be started in the perioperative period for patients with multiple RCRI risk factors or who are at intermediate or high risk for myocardial ischemia. If started in the perioperative period, β-blockers should be started long enough before surgery to ascertain their safety, and not on the day of surgery.\textsuperscript{32,33} Recent large randomized trials have demonstrated excess risk of mortality and stroke simultaneously with decreased risk of myocardial events in moderate- and high-risk patients who are newly treated with β-blockers in the periprocedural setting.\textsuperscript{34,35}

Implanted cardiac devices including pacemakers and implantable cardioverter-defibrillators also have important perioperative implications. A 2011 ASA practice advisory stressed the importance of determining whether electromagnetic interference is likely to occur during the planned procedure, determining the current function and necessity of the implanted device, determining whether reprogramming or temporary disabling of the device is advantageous, having alternative therapy available for the time that the device is unavailable, and restoring device function in the postoperative period.\textsuperscript{36}

**Pulmonary Disease**

Chronic pulmonary disease is an increasingly recognized cause of morbidity and mortality in surgical patients. For patients with asthma or chronic obstructive pulmonary disease, exercise tolerance and the frequency and severity of exacerbations should be evaluated. A focused history, including prior admissions and intubations for exacerbations, should be obtained. Treatment with bronchodilators in the perioperative setting is appropriate, although there is no literature to either guide this care or to document a benefit from it. Most inhaled anesthetics act as bronchodilators.\textsuperscript{37} Desflurane can be an airway irritant, and it is often avoided in patients with reactive airway disease.

The incidence of obstructive sleep apnea (OSA) has risen with the incidence of obesity. In 2014 the ASA published guidelines for perioperative management of patients with OSA. These guidelines highlight the importance of identifying patients with obstructive sleep apnea during preoperative evaluation and obtaining a sleep study if appropriate. They also highlight the importance of the development of protocols by anesthesiologists and surgeons to manage OSA in the perioperative setting. There is consensus that these patients should not be extubated until they are completely awake, and that they should be treated with
noninvasive positive pressure ventilation in the postoperative period as indicated.38

Renal Disease
Management of anesthesia in patients with chronic renal insufficiency requires close attention to perioperative fluid management and acid-base and electrolyte homeostasis. Doses of opioids and neuromuscular agents are typically reduced and dosing intervals increased to compensate for decreased renal excretion. Cisatracurium is often chosen as the muscle relaxant in patients with severe renal insufficiency because its elimination is unchanged by renal failure. Sugammadex, a reversal agent for steroid-based neuromuscular blockers, is not currently recommended for use in patients with advanced chronic kidney disease or end-stage renal disease.

Hepatic Disease
Hepatic dysfunction has many causes, and it can compel significant changes in anesthetic care. First, anesthetic agents metabolized in the liver can accumulate in these patients and may have a longer duration of effect. To mitigate this, short-acting agents are strongly preferred in these patients. Hypoalbuminemia can paradoxically increase the free plasma levels of drugs, which can also exaggerate their effects. In patients with substantial ascites, high intrabdominal pressure may increase the risk of passive gastric reflux, and thus many such patients are managed as if they have a full stomach, regardless of how long they have been NPO. In patients with significant hepatocellular dysfunction and/or portal hypertension, the combination of thrombocytopenia and coagulation factor deficiency not only increase the risk of bleeding associated with surgery, they also are relative or absolute contraindications to a variety of anesthetic techniques, such as subarachnoid blocks and epidural anesthesia. Presence of esophageal varices increases the risk of gastric tubes and transesophageal echocardiography.

Endocrine Disease
Perioperative management of the diabetic patient can be especially challenging. A hemoglobin A1c level should be obtained if a recent level is not available, as an increased A1c level is associated with an increase in perioperative complications including wound infections.39-41 Several institutions have implemented protocols for glucose management for diabetic patients undergoing surgery, although recommendations differ on appropriate target glucose levels.39,42,43 Perhaps the most important thing for a practitioner to know and remember is that the difference between type 1 and type 2 diabetes is important, and that these two different diseases require different approaches to their management. In general, patients with type 2 diabetes have a lower risk of becoming hypoglycemic, tend to have higher blood sugars at baseline, and tolerate higher levels of serum glucose without significant acute hazard. Patients with type 1 diabetes, who are deficient in insulin production and thus require insulin administration to prevent ketosis, are far more likely to become hypoglycemic when subjected to stress, and they are also at risk for developing ketoacidosis with hyperglycemia. Patients with type 1 diabetes merit more careful monitoring of their blood sugars in the perioperative setting than patients with type 2 diabetes.44-47

Preoperative Fasting
The ASA has developed specific guidelines for preoperative fasting to mitigate the risk of aspiration of gastric contents. Table 46-6 shows guidelines for preoperative food and fluid intake for elective procedures. Individual patients may need lengthier fasting times than the guidelines indicate.

Notably, a rapid sequence induction and intubation should be considered in patients who are at higher risk for aspiration such as those with very symptomatic gastroesophageal reflux, achalasia, gastroparesis, or dysmotility, regardless of fasting status.48

Patients With Advanced Directives
Patients with do not resuscitate (DNR) and/or do not intubate (DNI) orders present a unique challenge. Patients or their power-of-attorney may choose to rescind these directives in the perioperative period, maintain them as originally ordered, or modify them to allow for a limited resuscitation. Both the ASA and the American College of Surgeons recommend that preoperative discussions with the patient and their family clarify the patient’s wishes, and both societies emphasize that policies that mandate uniform enforcement or disregarding of all DNR orders take away patients’ right to self-determination.49,50

Risk Estimation
Several risk calculators have been developed to estimate perioperative morbidity and mortality. The American College of Surgeons National Surgical Quality Improvement Program (NSQIP) surgical risk calculator allows for estimation of risk of eight different adverse outcomes, including mortality.51 The Society of Thoracic Surgeons risk calculator estimates risk of morbidity and mortality after cardiac surgical procedures.52 While such risk estimators can provide widely differing estimates or may be inaccurate in certain situations,53 they are still invaluable aids for discussions with patients and their families about high-risk surgery.

INTRAOPERATIVE MANAGEMENT

General Anesthesia
General anesthesia remains the cornerstone of anesthesia practice; many surgical procedures cannot be done under regional techniques or monitored anesthesia care with sedation. The induction of general anesthesia can precipitate catastrophic anesthetic complications. Many different techniques can be used to induce general anesthesia, each with significant advantages and disadvantages.

Intravenous induction, used primarily in adults, quickly produces unconsciousness, and depending on the agent used, apnea as well. Propofol, the most common induction agent currently used, can cause hypotension due to its myocardial depressing and vasodilatory properties. Hypertension and
Tachycardia commonly occur during laryngoscopy or other significant airway stimulation.

The goal of a rapid sequence induction (RSI) is to achieve secure protection of the airway with a cuffed endotracheal tube without ever mask ventilating a patient. It is intended to prevent vomiting and aspiration, and it is routinely employed in patients at heightened risk for aspiration. There are no randomized controlled trials that demonstrate any kind of outcome benefit of rapid sequence induction in such patients, but it is nevertheless routinely employed for this purpose in the United States.

Pediatric patients are often not amenable to preoperative IV catheter placement. Hence, inhalation induction of anesthesia is commonly used in children, with IV placement occurring after induction. Even among children, however, patients at heightened risk for aspiration or with a full stomach may be best managed with preoperative IV placement and an IV induction. Patients with developmental delay may not be amenable to preoperative IV placement or inhalational induction of anesthesia. In such patients, intramuscular administration of an agent such as ketamine is often required to induce anesthesia.

**Airway Management.** Most anesthesiologists prefer to secure the airway of a patient undergoing general anesthesia, and this is usually accomplished immediately after anesthesia has been induced. The airway may be managed in several ways, including by face mask, with a laryngeal mask airway (LMA), or, most definitively, by endotracheal intubation with a cuffed endotracheal tube. Nasal and oral airways can help establish a patent airway in a patient being ventilated with a mask by creating an air passage behind the tongue.

The LMA is a cuffed supraglottic oral airway that is inserted through the oropharynx and ideally positioned just above the glottis opening. It is passed blindly, and the inflated cuff creates a seal around the laryngeal inlet. An LMA does not protect against aspiration and should generally not be used in patients with a high risk of aspiration.

Tracheal intubation requires a skilled operator and proper equipment. In most elective anesthetics, attempts to intubate the trachea are facilitated by the administration of muscle relaxants in a patient who is already under a general anesthetic. Intubation is typically performed under direct visualization with a laryngoscope, watching the endotracheal tube pass through the vocal cords into the trachea. To obtain a direct line of sight, the patient is placed in the sniffing position. The neck is flexed at the lower cervical spine and extended at the atlanto-occipital joint. This flexion and extension are amplified during laryngoscopy. Laryngoscope blades can be curved (Macintosh) or straight (Miller) blades. Laryngoscopic views are typically reported in a classification system developed by Cormack and Lehane (Fig. 46-5).

**Management of the Difficult Airway.** Some patients have physical characteristics or a history suggestive of difficulty in placing an endotracheal tube. A short neck, limited neck mobility, small interincisor distance, short thyromental distance, and high Mallampati classes may all represent a challenge to endotracheal intubation. Several tools have been developed to assist in management of the difficult airway.

The GlideScope, a video laryngoscope, allows for visualization of the larynx on a video screen (Fig. 46-6). Having more of a bend than a standard curved Macintosh blade, it can be advantageous for visualizing and intubating the trachea in patients with large tongues or relatively anterior glottis openings. Placement of the endotracheal tube once the larynx has been visualized can still be challenging. A recent study of ICU patients requiring intubation showed that video laryngoscopy did not improve first-pass orotracheal intubation success rate and was associated with higher rates of severe life-threatening complications.

The intubating laryngeal mask airway (ILMA) is an advanced form of LMA designed to maintain a patent airway and facilitate tracheal intubation. The ILMA can be placed in anticipated or unexpectedly difficult airways as an airway rescue device and as a guide for intubating the trachea. The device itself is substantially more rigid than other laryngeal mask airways, and includes a handle which the operator can use to displace the opening of the device. A specially manufactured endotracheal tube can be passed blindly through the ILMA into the larynx, or the ILMA can be used as a conduit for a flexible fiberoptic scope. Experience with airway management in general and the use of this device in particular is essential for its effective use in emergency situations; operators with little experience will enjoy little success with this device.

The flexible fiberoptic intubation bronchoscope is the gold standard for difficult intubation. It is indicated in difficult or compromised airways where neck extension is not desirable or in cases with risk of dental damage. The flexible bronchoscope allows excellent visualization of the airway and glottic opening. This technique can be used for oral and nasal intubation, for awake or asleep intubation, and for intubation in the awake, spontaneously ventilating patient whose airway has been treated with topical local anesthetic.

The ASA has developed an algorithm for management of the difficult airway (Fig. 46-7). Notably, in patients in whom...
1. Assess the likelihood and clinical impact of basic management problems:
   - Difficulty with patient cooperation or consent
   - Difficulty mask ventilation
   - Difficult supraglottic airway placement
   - Difficult laryngoscopy
   - Difficult intubation
   - Difficult surgical airway access

2. Actively pursue opportunities to deliver supplemental oxygen throughout the process of difficult airway management.

3. Consider the relative merits and feasibility of basic management choices:
   - Awake intubation vs. intubation after induction of general anesthesia
   - Noninvasive technique vs. invasive techniques for the initial approach to intubation
   - Video-assisted laryngoscopy as an initial approach to intubation
   - Preservation vs. ablation of spontaneous ventilation

4. Develop primary and alternative strategies:
   
   **Awake Intubation**
   - Airway approached by Noninvasive intubation
     - Succeed
       - Invasive Airway Access
     - FAIL
       - Cancel Case
         - Consider feasibility of other options
         - Invasive airway access
   
   **Intubation After Induction of General Anesthesia**
   - Initial intubation attempts successful
   - Initial intubation attempts UNSUCCESSFUL
     - FROM THIS POINT ONWARDS
     - CONSIDER:
       - 1. Calling for help.
       - 2. Returning to spontaneous ventilation.
       - 3. Awakening the patient.

   **Face Mask Ventilation Adequate**
   - Ventilation adequate, intubation unsuccessful
     - Alternative approaches to intubation
       - Successful Intubation
       - FAIL after multiple attempts
         - Invasive airway access
         - Consider feasibility of other options
         - Awake patient
   
   **Face Mask Ventilation Not Adequate**
   - CONSIDER/ATTEMPT SGA
     - SGA ADEQUATE
     - SGA NOT ADEQUATE or NOT FEASIBLE
       - Emergency noninvasive airway ventilation
         - Emergency intubation
           - Successful ventilation
           - FAIL

   *Confirm ventilation, tracheal intubation, or SGA placement with exhaled CO₂.*

   **a.** Other options include (but are not limited to): surgery utilizing face mask or supraglottic airway (SGA) anesthesia (e.g., LMA, ILMA, laryngeal tube), local anesthesia infiltration, or regional nerve blockade. Pursuit of these options usually implies that mask ventilation will not be problematic. Therefore, these options may be of limited value if this step in the algorithm has been reached via the Emergency Pathway.

   **b.** Invasive airway access includes surgical or percutaneous airway, jet ventilation, and retrograde intubation.

   **c.** Alternative difficult intubation approaches include (but are not limited to): video-assisted laryngoscopy, alternative laryngoscope blades, SGA (e.g., LMA or ILMA) as an intubation conduit (with or without fiberoptic guidance), fiberoptic intubation, intubating stylet or tube changer, light wand, and blind oral or nasal intubation.

   **d.** Consider repreparation of the patient for awake intubation or canceling surgery.

   **e.** Emergency noninvasive airway ventilation consists of a SGA.

---

both intubation and ventilation are impossible, the algorithm calls for placement of an LMA with ventilation attempted through the LMA.

Monitored Anesthesia Care
Monitored anesthesia care (MAC) is when a patient undergoes a procedure under local anesthesia under the care of an anesthesiologist who can provide sedation as indicated. Sedation is administered to a level that allows the patient to maintain airway reflexes and breath spontaneously. Advantages of MAC anesthesia include reduced invasiveness, as the airway is not manipulated, and faster recovery. ASA standard monitors must be used, including capnography, which allows for rapid detection of apnea or hypoventilation. In some instances, most commonly for gastroenterology procedures, patients are administered an intravenous anesthetic that is often classified as a MAC, even though the patient is so deeply anesthetized that they have no significant response to significant airway stimulation; the airway is monitored by the anesthesia provider and secured if necessary.

Regional Anesthesia/Acute Pain
Regional anesthesia allows for selective blockade and is an excellent anesthetic option for several different types of procedures. Regional anesthesia can also provide excellent postoperative pain control. Regional techniques include neuraxial blockade, including spinal and epidural anesthetics, peripheral nerve blocks, and truncal blocks.

Spinal anesthetics (also referred to as subarachnoid blocks) can be used for lower extremity, lower abdominal, pelvic, and urologic and gynecologic procedures. A small caliber needle (typically 25-gauge or smaller) is inserted into the intrathecal space in the cauda equina and below the conus medullaris, and a small volume of local anesthetic is injected. Duration and level of the block in the spinal cord is affected by the anesthetic used, the dose employed, and baricity of solution injected. Complications can include hypotension, bradycardia, postdural puncture headache, injury to local structures or nerves, and hematoma formation. The American Society of Regional Anesthesia publishes guidelines regarding safe intervals to perform neuraxial anesthetics after the administration of anticoagulant and antiplatelet agents.

An epidural catheter can be used as a primary anesthetic for a procedure, or it can be placed preoperatively and used in conjunction with a general anesthetic for postoperative pain control. Epidural catheters can be placed in the thoracic or lumbar spine and can remain in place for days after surgery. A dilute local anesthetic and/or opioid is administered through the catheter to provide analgesia. Complications of epidural anesthesia are similar to that of spinal anesthesia. In addition to improved pain control, benefits of epidural anesthesia include reduced pulmonary complications and decreased duration of postoperative ileus.

Peripheral nerve blockade can also be used to provide surgical anesthesia as well as postoperative analgesia, particularly for surgeries of the upper or lower extremities. The nerve or plexus of interest is located with ultrasound and/or peripheral nerve stimulator, and local anesthetic is injected around the nerve. Single-shot nerve blocks allow for surgical anesthesia and immediate postoperative analgesia and can last for several hours. Flexible catheters can also be placed in proximity to nerves to allow for continuous infusion and blockade that can continue for several days. Complications of peripheral nerve blocks include injury to nerves or nearby structures and local anesthetic systemic toxicity.

Recently, truncal blocks have become more commonly and widely used for the treatment of postoperative pain. Truncal blocks include the transversus abdominis plane (TAP) block, the rectus sheath block, the pectoral nerve block, and the serratus anterior plane block. These truncal blocks are usually done under ultrasound guidance with local anesthetic injected in the appropriate plane. Truncal blocks are typically performed as part of a multimodal approach to postoperative pain. Limited evidence suggests that use of truncal blocks decreases postoperative opioid requirements.

RECOVERY AND COMPLICATIONS

The Postanesthesia Care Unit
The advent of the modern postanesthesia care unit represents a major advance in the safety of perioperative care, as the close monitoring that occurs there can prevent or expedite the management of a variety of serious complications. Ventilation, oxygenation, hemodynamics, temperature, nausea, and pain are closely monitored in the PACU, with close attention also given to urine output, ongoing bleeding, and drainage. To be discharge eligible, patients should have returned to their baseline mental status, be oxygenating and ventilating adequately, have adequate pain control, and have stable vital signs. There are multiple scoring systems that can be used assess suitability for discharge from PACU. Postoperative hemorrhage, hypertension or hypotension, myocardial ischemia, arrhythmias, and altered mental status commonly manifest in the postoperative care unit. Postoperative nausea and vomiting (PONV) occurs in 20% to 30% of surgical cases, and it is a common cause of increased PACU length of stay and increased cost of PACU stay. For this reason, many or most patients undergoing general anesthesia receive prophylactic antiemetics.

Enhanced Recovery After Surgery Pathways
Enhanced recovery after surgery (ERAS) pathways are multimodal perioperative care pathways designed to hasten recovery after elective surgery. These pathways may include preoperative education and counseling, preoperative optimization, limiting preoperative bowel preparation, limiting preoperative fasting, providing multimodal analgesia (including regional anesthesia) as appropriate, minimizing intraoperative fluid administration, and early mobilization. ERAS pathways have been shown to reduce duration of hospitalization and reduce cost of perioperative care.

Acute Postoperative Pain
The management of postoperative pain has changed dramatically in the modern era, with multimodal approaches and regional techniques reducing the use of opiates. Regardless, opioids remain the mainstay of intraoperative and postoperative analgesia, especially for larger and more invasive procedures. Patients with chronic pain or opiate tolerance can present a unique challenge in the perioperative period, and they can benefit from regional and multimodal approaches. Such patients may benefit from early involvement of an acute pain medicine specialist.

For the past decade, pain has been described as the fifth vital sign, and physicians have been strongly encouraged to
aggressively treat pain in their patients. Generally speaking, physicians sought to accomplish this goal through the more liberal use of opiates. Opioid prescription has thus soared in recent years. Some studies suggest that 3% to 7% of surgical patients prescribed opioids in the postoperative period continue to use them for a prolonged period after surgery, suggesting that opioid abuse often begins in the postoperative period. The U.S. Centers for Disease Control and Prevention has recently declared that prescription drug abuse is an epidemic. Physicians and surgeons will have to continue to seek the fine line between adequate pain control and prescribing patterns that enable dependence.

Malignant Hyperthermia
Malignant hyperthermia (MH) is a hereditary, life-threatening, hypermetabolic disorder, developing during or after receiving general anesthesia. The clinical incidence of MH ranges from 1:10000 to 1:250,000. A genetic predisposition and exposure to one or more triggering agents are necessary to evoke MH. Triggering agents include all volatile anesthetics (e.g., isoflurane, sevoflurane, and desflurane) and the depolarizing neuromuscular blocker succinylcholine. Volatile anesthetics and/or succinylcholine cause a rise in the myoplasmic calcium concentration in susceptible patients, causing persistent muscle contraction, the production of large quantities of carbon dioxide and lactic acid, and a relentless increase in body temperature.

MH is often an autosomal dominant disorder associated with several gene loci, predominantly the ryanodine receptor gene RYR1. MH can be diagnosed with the caffeine-contracture halothane test, which requires a muscle biopsy. Genetic testing can be helpful after an episode of MH. There is no simple, reliable blood screening test yet available for diagnosis.

The classic MH crisis entails a hypermetabolic state with tachycardia and increased end-tidal CO₂. Relentless muscle contraction causes respiratory and metabolic acidosis, as well as rhabdomyolysis, arrhythmias, hyperkalemia, and even sudden cardiac arrest. Hyperthermia typically occurs after the episode is well under way. Treatment must be aggressive and begin as soon as a case of MH is suspected. Volatile anesthetics should be stopped immediately and dantrolene given at an initial dose of 2.5 mg/kg intravenously. The national MH hotline should be contacted for help in managing any patient with MH. Patients should be monitored in the intensive care setting for possible recrudescence of MH.

Cardiovascular Complications
Hemodynamic perturbations are a common in the perioperative period. Arrhythmias may begin before, during, and after an anesthetic, and are particularly common after cardiothoracic and esophageal surgery. Hypotension may be due to anemia, hypovolemia, myocardial ischemia or dysfunction, or other less common events such as pulmonary embolism and anaphylactic reactions. Hypertension is also common, particularly when antihypertensive regimens are altered in the perioperative period.

Respiratory Failure
Respiratory insufficiency and failure occur frequently in the postoperative period. Respiratory depression can occur as a consequence of residual neuromuscular blockade, residual inhaled anesthetics, or opioids. Mechanical airway obstruction can be ameliorated by rescue maneuvers, or insertion of an oral or nasal airway. It is imperative to evaluate and treat anesthesia-related causes of respiratory insufficiency. The opioid antagonist naloxone can be given for opioid related respiratory insufficiency, and an additional reversal agent can be given for residual neuromuscular weakness. Respiratory failure can reasonably be managed with noninvasive positive pressure ventilation in many cases; patients who fail this or are unlikely to benefit from it should be intubated. High-flow nasal cannula is increasingly being used for postextubation respiratory failure, and several studies have demonstrated its benefit in postoperative patients.

Neurologic and Psychiatric Complications
Perioperative neurologic and psychiatric complications include stroke, both ischemic and hemorrhagic, postoperative delirium, and postoperative cognitive dysfunction. Treatment of perioperative stroke may be difficult as initiation of anticoagulation or thrombolysis may not be safe after surgery. Postoperative delirium is common and transient. Treatment includes reorientation, treatment of pain, workup for metabolic, hemodynamic, or respiratory perturbations, and consideration of the side effects of the anesthetics and analgesics administered. Antipsychotics such as haloperidol can be useful in the treatment of postoperative delirium. Postoperative cognitive dysfunction (POCD) is a decline in cognitive function that may last days or may persist for months. Risk factors for persistent POCD include advanced age, history of prior stroke, and lower educational level. POCD rates among older adult patients have been shown to be as high as 40% at hospital discharge and 12% 3 months after surgery. Notably, a causal link has not been established between administration of anesthesia and development of POCD, suggesting that it may be the physiologic stress of the perioperative experience that may lead to a decline in cognitive function in such patients.

CONCLUSION
The practice of anesthesia has improved dramatically over the past century. Advances in training, pharmacology, anesthesia equipment, and monitoring have not only made anesthesia dramatically safer but have also allowed ever sicker patients to benefit from surgery.

REFERENCES

Entries highlighted in bright blue are key references.


39. Stryker LS, Abdel MP, Morrey ME, et al. Elevated postoperative blood glucose and preoperative hemoglobin A1C are...


This page intentionally left blank
INTRODUCTION

As the population ages, an increasing number of older adults will develop surgical diseases. The segment of the U.S. population age 65 and older is expected to double by 2050. Older adults present unique challenges in surgical management and decision-making. The accumulation of comorbid conditions and physiologic vulnerability that occurs with age put older adults at high risk for major morbidity and mortality after surgery. It is essential that surgeons approach this population with a new set of skills and knowledge to provide optimal care for this vulnerable population. A comprehensive understanding of the unique vulnerabilities of older adults—geriatric syndromes and risk factors—are required to accurately estimate surgical risk, inform surgical decision-making, and guide perioperative management. In this chapter, we will (a) discuss the physiologic conditions that are common in older adults that are essential for surgical risk assessment of older adults considering surgery, (b) describe best practices in perioperative care in the older adult, and (c) discuss special considerations and common pitfalls for surgical conditions that are common in the older adult.

GERIATRIC SYNDROMES

Approximately one-third of older individuals undergoing vascular and urologic surgery suffer from geriatric syndromes. This term is used to describe clinical conditions that do not fit into discrete disease categories but that can substantially negatively impact quality of life and result in disability. Geriatric syndromes often involve multiple underlying factors and organ systems (i.e., multiple causation of a unified manifestation) and include frailty, falls, delirium, malnutrition, dizziness, syncope, urinary incontinence, and pressure ulcers among others. These syndromes can be present before surgery and/or develop as a result of surgery and hospitalization. Furthermore, close attention to both the presence and development of geriatric syndromes among surgical candidates is important and often overlooked, making caring for older surgical patients unique compared to their younger and healthier counterparts.

Frailty

Frailty is among the most widely studied geriatric syndrome in the surgical literature. Frail older individuals are at high risk for adverse events in the face of stressors such as surgery. They are more likely to experience surgical complications, delayed recovery, falls, and to develop functional impairment. Frailty is also associated with a higher risk of death. It is believed that frailty is a chronic, progressive condition that represents a spectrum; less frail individuals may be responsive to strategies or interventions to ameliorate its clinical manifestations, while more frail individuals may demonstrate an irreversible predeath condition with limited life expectancy.

Frailty has been shown to be independently predictive of poor postoperative outcomes. Makary et al studied 594 older patients presenting for elective surgery at a university hospital and demonstrated that frail individuals were at increased risk of postoperative complications (OR 2.54; 95% CI 1.12–5.77), longer length of stay (incidence rate ratio 1.69; 95% CI 1.28–2.23), and discharge to a skilled or assisted living facility after previously living at home (OR 20.48; 95% CI 5.54–75.68). Similar findings have been echoed throughout the surgical literature inclusive of vascular, colorectal, cardiac, urologic, and other types of procedures.

Definitions of frailty fall into two broad models; a phenotypic model and a deficits accumulation model. The phenotypic model was originally described by Linda Fried using data from the Cardiovascular Health Study, which is an observational study of community-dwelling men and women age 65 years and
older. Individuals in this cohort underwent baseline evaluations and had 4 to 7 years of follow-up with annual examinations and surveillance for the following outcomes: incident disease, hospitalization, falls, disability, and mortality. Based on observations in these individuals over time, the following criteria were identified to define frailty: weight loss, exhaustion, physical activity, walk time, and grip strength (Table 47-1). The presence of one or two of these factors is associated with intermediate risk for poor outcomes, i.e., a “prefrail” phenotype, and the presence of three or more of these factors is associated with high risk for poor outcomes, i.e., a “frail” phenotype. This study additionally demonstrated that frailty is strongly associated

### Table 47-1

**Criteria used to define frailty**

- **Weight loss:** “In the last year, have you lost more than 10 pounds unintentionally (i.e., not due to dieting or exercise)?” If yes, then frail for weight loss criterion. At follow-up, weight loss was calculated as: (Weight in previous year – current measured weight)/(weight in previous year) = K. If K ≥0.05 and the subject does not report that he/she was trying to lose weight (i.e., unintentional weight loss of at least 5% of previous year’s body weight), then frail for weight loss = Yes.

- **Exhaustion:** Using the CES-D Depression Scale, the following two statements are read. (a) I felt that everything I did was an effort; (b) I could not get going. The question is asked “How often in the last week did you feel this way?” 0 = rarely or none of the time (<1 day), 1 = some or a little of the time (1−2 days), 2 = a moderate amount of the time (3−4 days), or 3 = most of the time. Subjects answering “2” or “3” to either of these questions are categorized as frail by the exhaustion criterion.

- **Physical Activity:** Based on the short version of the Minnesota Leisure Time Activity questionnaire, asking about walking, chores (moderately strenuous), mowing the lawn, raking, gardening, hiking, jogging, biking, exercise cycling, dancing, aerobics, bowling, golf, singles tennis, doubles tennis, racquetball, calisthenics, swimming. Kcals per week expended are calculated using standardized algorithm. This variable is stratified by gender. *Men:* Those with Kcal of physical activity per week <383 are frail. *Women:* Those with Kcals per week <270 are frail.

- **Walk Time,** stratified by gender and height (gender-specific cutoff a medium height).

<table>
<thead>
<tr>
<th>Gender</th>
<th>Cutoff for time to walk 15 feet criterion for frailty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>≥7 seconds</td>
</tr>
<tr>
<td>Height ≤173 cm</td>
<td></td>
</tr>
<tr>
<td>Height &gt;173 cm</td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>≥6 seconds</td>
</tr>
<tr>
<td>Height ≤159 cm</td>
<td></td>
</tr>
<tr>
<td>Height &gt;159 cm</td>
<td></td>
</tr>
</tbody>
</table>

- **Grip Strength,** stratified by gender and body mass index (BMI) quartiles:

<table>
<thead>
<tr>
<th>Gender</th>
<th>Cutoff for grip strength (Kg) criterion for frailty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td></td>
</tr>
<tr>
<td>BMI ≤24</td>
<td>≤29</td>
</tr>
<tr>
<td>BMI 24.1–26</td>
<td>≤30</td>
</tr>
<tr>
<td>BMI 26.1–28</td>
<td>≤30</td>
</tr>
<tr>
<td>BMI &gt;28</td>
<td>≤32</td>
</tr>
<tr>
<td>Women</td>
<td></td>
</tr>
<tr>
<td>BMI ≤23</td>
<td>≤17</td>
</tr>
<tr>
<td>BMI 23.1–26</td>
<td>≤17.3</td>
</tr>
<tr>
<td>BMI 26.1–29</td>
<td>≤18</td>
</tr>
<tr>
<td>BMI &gt;29</td>
<td>≤21</td>
</tr>
</tbody>
</table>

with several major chronic diseases, including cardiovascular disease, pulmonary disease, and diabetes; however, not all frail individuals demonstrated these associations. The same is true for disability. While there is some overlap between frailty and disability, not all frail individuals are disabled. These findings suggest that while there may be overlap between these three constructs in some individuals, frailty is a distinct process from both comorbidity and disability.12

Alternatively, the deficit accumulation model, developed by Rockwood et al, suggests that frailty is defined by discrete failures of redundant physiologic systems. The more deficits that occur, the more likely it is that adverse outcomes will result. Using data from the Canadian Study of Health and Aging, a longitudinal study of individuals age 65 and older, the authors developed a frailty index represented by the cumulative proportion of 92 accumulated deficits, which include symptoms, signs, functional impairments, and laboratory abnormalities. They demonstrated that deficits accumulated at a rate of 3% per year in their cohort, represented a gamma distribution, and increased with chronological age; they proposed that this model be used as a proxy for aging and mortality.13

While these models are helpful to conceptualize frailty, no one model is all inclusive, and each may have applicability in different settings. For example, the frailty phenotype does not include items on cognition or mood and may not be easily applicable to the busy clinical setting.14 The deficits accumulation model is ideal for use in large databases, such as the American College of Surgeons National Surgical Quality Improvement Project (ACS-NSQIP),15 and may be helpful for research and public health and policy purposes, but it is not practical for clinical care. Measurement of frailty in the clinical setting will be discussed later in this chapter.

While frailty is often defined as a geriatric syndrome, it is also plausible that other geriatric syndromes (i.e., urinary incontinence, falls, pressure ulcers, delirium, and functional decline) may demonstrate shared risk factors that lead to frailty. In turn, frailty may also cause more risk factors and more geriatric syndromes.16 Regardless of the association and directionality between frailty and other geriatric syndromes, identification of each is essential in the preoperative setting in order to help risk stratify and potentially to mitigate risk for patients considering surgical intervention.

Falls
Older adults are at markedly increased risk of falls, and one in three adults age 65 and older report falling in the last year. The incidence of falls increases with age, and close to 60% of individuals who have fallen in the last year will fall again within the following year. Falls are associated with subsequent declines in functional status, greater likelihood of nursing home placement, increased use of medical services, and development of a fear of falling. Approximately half of older individuals who fall are unable to get up, resulting in a “long lie,” which is further associated with lasting functional declines.17 In fact, falls can be so detrimental to older individuals that the Joint Commission (United States) established fall prevention as one of its national safety goals in 2015.18

Causes of falls can be multifactorial, as with other geriatric syndromes. Factors include age-related declines, chronic disease, medications, environmental factors, changing positions, routine activities, risk-taking behaviors, acute illness, or situational hazards such as the unfamiliar setting of hospitals and long-term care facilities.17 It has been shown that preoperative falls are associated with poor postoperative outcomes among patients undergoing elective surgery. One study looking at 7982 such patients found that a preoperative history of one, two, or three or more falls predicted postoperative falls at 30 days (adjusted OR 2.3, 3.6, 5.5, respectively) and 1 year (adjusted OR 2.3, 3.4, 6.9, respectively), in addition to predicting a decline in functional status at 30 days (adjusted OR 1.2, 2.4, 2.4, respectively) and 1 year (adjusted OR 1.3, 1.5, 3.2, respectively) and in-hospital complications (adjusted OR 1.2, 1.3, 2.0, respectively).18 Furthermore, preoperative falls are a major predictor of poor postoperative outcomes and may be a valuable preoperative assessment tool as part of routine preoperative care.

Delirium
Delirium is a disorder of attention and awareness that develops acutely and tends to fluctuate, as defined by the new Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) criteria. Delirium is so common among older adults that up to one third of patients age 70 and older admitted to the hospital experiences delirium, half of which have delirium on admission and the other half of which develop delirium during the hospitalization itself.19 Rates of delirium among older patients undergoing surgery ranges from 4% to 5% in cataract and urologic procedures to 50% to 60% in infrarenal AAA repair or hip fracture surgery.20 Patients who develop postoperative delirium have a two- to threefold increased risk of mortality within the first year after surgery.21

Preoperative assessment should focus on the identification of risk factors for delirium including age 70 years or older, cognitive impairment, limited physical function, history of alcohol abuse, abnormal serum sodium, potassium or glucose, intrathoracic surgery, and AAA surgery. Preoperative assessment and documentation of mental status is imperative in order to establish a baseline for postoperative comparison. Intraoperative blood loss is another risk factor for postoperative delirium, and patients with a postoperative hematocrit less than 30% are at increased risk, irrespective of baseline risk factors. Postoperatively, undertreatment of pain is an important risk factor for delirium. Management of fluid, electrolyte and metabolic abnormalities, optimization of blood loss replacement, maintenance of circadian rhythms, and cautious prescription of medication and pain management are among the most important methods by which to minimize the risk of postoperative delirium among surgical patients.20 Furthermore, many commonly used medications may induce delirium and should be avoided. They include drugs with anticholinergic properties, corticosteroids, merperidine, and sedative hypnotics. Finally, postoperative care bundles have been shown to successfully reduce the incidence of delirium. These strategies include sensory enhancement (glasses, hearing aids), early mobility, cognitive orientation and therapeutic activities, and sleep protocols. Treatment of postoperative delirium should focus on treatable etiologies (Fig. 47-1), and pharmacologic treatment should be reserved for patients who are at risk of harming themselves or others.

PREOPERATIVE ASSESSMENT
Preoperative assessment in older adults is more complex than in younger individuals, as there are many unique characteristics that require consideration. The purpose of the assessment
is not to “clear” the patient for surgery but rather to minimize risks and optimize good outcomes. The ACS NSQIP and the American Geriatrics Society (AGS) published best practice guidelines to help optimize this process. Preoperative planning should be proactive, commencing at the time of surgical decision-making, if not sooner. To this end, the ACS NSQIP/AGS produced best practice guidelines and a preoperative management checklist to provide a framework for thinking about critical issues in this patient population (Table 47-2). In addition to these issues, they emphasize the importance of planning analgesia strategies, making efforts to minimize opioid use and to prevent functional and cognitive decline, obtaining multidisciplinary consultation early and early involvement of allied health staff (i.e., physical or occupational therapy), and anticipating home health needs that may be required at discharge.

**Best Practices: Preoperative Assessment**

**Functional Assessment.** As previously discussed, poor physical function prior to surgery is associated with higher risk of major postoperative complications, increased need for intensive
Table 47-2
Immediate preoperative management checklist from the ACS NSQIP/AGS

1. Confirm and document patient goals and treatment preferences, including advance directives
2. Confirm and document patient’s health care proxy or surrogate decision-maker
3. In patients with existing advance directives, discuss new risks associated with the surgical procedure and an approach for potentially life-threatening problems consistent with the patient’s values and preferences
4. Consider shortened fluid fast (clear liquids up to two hours before anesthesia)
5. Adhere to existing best practices regarding antibiotic and venous thromboembolism prophylaxis
6. Ensure nonessential medications have been stopped and essential medications have been taken


debilitating services, increased rates of discharge to a skilled or assisted nursing facility, and higher mortality.24-26 Assessment of physical function and performance status in the preoperative setting are recommended by the ACS NSQIP/AGS best practice guidelines.22 There are several methods by which to measure physical function. Overall functional status may be ascertained by assessing the ability of an individual to perform activities of daily living (ADLs) and instrumental ADLs. ADLs include dressing, bathing, toileting, transferring, continence, and eating independently.23 Instrumental ADLs measure an individual’s ability to live independently and include the ability to perform the following tasks: shopping, laundry, mode of transportation, ability to handle finances, responsibility for won medications, food preparation, and housekeeping.27

In addition to assessment of ADLs and instrumental ADLs, the surgeon should also assess for deficits in vision, hearing, and swallowing, inquire about history of falls in the past year, evaluate for limitations in gait and mobility, and determine risk for falls, which can be performed via the timed up and go test (TUGT). This test measures gait and mobility impairment and is associated with increased risk for falls in ambulatory individuals. All that is required to perform this test are a chair, a mark 10 feet in front of the chair, and a stopwatch. Individuals are instructed to do the following while being timed:

1. Stand up from the chair (without using arm rests, if possible)
2. Walk to the mark (10 feet in front of them)
3. Turn
4. Walk back to the chair
5. Sit down in the chair

A time of ≥15 seconds indicates high risk of falls and should prompt referral to physical therapy for further assessment.

In a prospective cohort of individuals age 65 and older undergoing surgery, the TUGT times were stratified into three groups that strongly correlated with varying risk for postoperative complications and 1-year mortality. These groups were “slow” (≥15 seconds), “intermediate” (11–14 seconds), and “fast” (≤10 seconds). Postoperative complications and 1-year mortality in the slow group were significantly higher compared to those in the fast group, 52% to 77% versus 11% to 13% for complications, and 31% compared to 3% for mortality, respectively.5

Frailty Assessment
As stated earlier, frailty is an important consideration in preoperative planning for older individuals. Measurement of frailty can take several forms. One method is to apply the operational definition put forth by Fried (see Table 47-1), which has been applied to surgical patients and shown to be an independent predictor of postoperative adverse events, increased length of stay, and higher likelihood of discharge to a skilled or assisted living facility.38

The frailty phenotype, however, may be cumbersome to apply in the busy clinical setting. To this end, Robinson proposed alternative definitions and methods for frailty measurement. One such method includes the following criteria: cognitive impairment (Mini-Cog score of ≤3), poor nutrition (serum albumin ≤3), history of falls (≥1 fall in the past 6 months), and low hematocrit (<35%).26 A second definition includes functional impairment (TUGT ≥15 seconds and dependence in any ADL) and comorbidity (Charlson index score ≥3).28

Cognitive and Behavioral Assessment
Preoperative cognitive impairment is strongly linked to postoperative delirium, worse surgical outcomes, longer hospital stays, increased risk of functional decline, and even mortality. History and cognitive assessment are important to consider early on in all surgical candidates age 65 years and older in the preoperative setting. The Mini-Cog, consisting of the three-item recall and clock draw tests, can be used to complete this assessment (Table 47-3). If possible, someone who knows the patient well (such as a spouse or family member) should be interviewed about the presence and evolution of any cognitive decline in the patient. If decline is present, the patient should be referred to a primary care physician, geriatrician, or mental health specialist for further evaluation. Documentation of preoperative cognitive status will further assist in the identification of any postoperative cognitive dysfunction.22 Risk factors for postoperative delirium should also be assessed in the preoperative period and are detailed in Table 47-4.

In addition to measuring cognitive status, assessing the patient’s decision-making capacity is also important to determine the patient’s ability to provide informed surgical consent. It is helpful to ask the patient to describe, in his/her own words, the important features of the discussion, the condition and indications for surgery, and the risks, benefits, and alternatives to surgery. There are four legally-relevant criteria for decision-making capacity:22:

1. The patient can clearly indicate his/her treatment choice.
2. The patient understands the relevant information communicated by the physician.
3. The patient acknowledges his/her medical condition, treatment options, and likely outcomes.
4. The patient can engage in a rational discussion about the treatment options.

Depression should also be screened for in the preoperative setting, as up to 11% of the population age 71 years and older in the United States suffer from this condition.29 Risk factors
### Table 47-3
**Cognitive assessment: three-item recall and clock draw**

1. Get the patient’s attention and say:
   
   “I am going to say three words that I want you to remember now and later. These words are banana, sunrise, chair.
   Please say them for me now.”
   
   Give the patient three tries to repeat the words. If unable after three tries, go to next item.

2. Say all of the following phrases in the order indicated:
   
   “Please draw a clock in the space below. Start by drawing a large circle. Put all the numbers in the circle and set the hands to show 11:10 (10 past 11).”
   
   If the subject has not finished clock drawing in 3 minutes, discontinue and ask for recall items.

3. Say: “What were the three words I asked you to remember?”

**Scoring:**

- 3 item recall (0-3 points): 1 point for each correct word
- Clock draw (0-2 points): 0 points for abnormal clock; 2 points for normal clock

A normal clock has all of the following elements:

1. All numbers 1-12, each only once, are present in the correct order and direction (clockwise) inside the circle.
2. Two hands are present, one pointing to 11 and one pointing to 2.
3. Any clock missing any of these elements is scored abnormal. Refusal to draw a clock is scored abnormal.

**Total score of 0, 1, or 2 suggests possible impairment**

**Total score of 3, 4, or 5 suggests no impairment**


---

### Table 47-4
**Risk factors for postoperative delirium**

- Cognitive and behavioral disorders:
  - Cognitive impairment and dementia
  - Untreated or inadequately controlled pain
  - Depression
  - Alcohol use
  - Sleep deprivation

- Disease or illness related:
  - Severe illness or comorbidities
  - Renal insufficiency
  - Anemia
  - Hypoxia

- Metabolic:
  - Poor nutrition
  - Dehydration
  - Electrolyte abnormalities

- Functional impairments:
  - Poor functional status
  - Immobilization
  - Hearing or vision impairment

- Other:
  - Older age ≥70 years
  - Polypharmacy and use of psychotropic medications (benzodiazepines, anticholinergics, and antihistamines)
  - Risk of urinary retention or constipation, presence of urinary catheter


---

for depression in older adults include female gender, disability, bereavement, and sleep disturbance. Poor health status, cognitive impairment, living alone, and new medical illness may also contribute to depression.30

The Patient Health Questionnaire-2 (PHQ-2) can be used to screen for depression via the following two questions31:

1. *In the past 12 months, have you ever had a time when you felt sad, blue, depressed, or down for most of the time for at least two weeks?*

2. *In the past 12 months, have you ever had a time, lasting at least two weeks, when you didn’t care about the things that you usually cared about or when you didn’t enjoy the things that you usually enjoy?*

If a patient answers yes to either question, further reevaluation with a primary care physician, geriatrician, or psychiatrist is recommended.

Screening for alcohol and substance abuse is also recommended, as alcohol use is common among older adults. Up to 13% of men and 8% of women ≥65 years consume at least 2 drinks per day and 14.5% of men and 3.3% of women consume 5 or more drinks per day.32 Alcohol and substance abuse are associated with increased rates of postoperative mortality and complications including pneumonia, sepsis, wound infection and disruption, and prolonged length of stay.33,34 The ACS NSQIP/AGS recommend screening for alcohol and substance abuse among older individuals with the modified CAGE questionnaire in combination with prescribing daily multivitamins, including folic acid and high dose (100 mg) oral or parenteral thiamine to patients who drink alcohol.22,35

### Medical Assessment

A thorough medical assessment should be performed in all older operative candidates and should include a cardiac evaluation, pulmonary evaluation, nutritional assessment, and medication evaluation and management where appropriate.

Cardiac adverse events are the most common cause of serious perioperative morbidity and mortality among patients undergoing noncardiac operations and occur more commonly in older adults.36,37 For these reasons, cardiac evaluation may be helpful to identify older patients with higher risk for cardiac complications who may be candidates for perioperative optimization. This evaluation should follow the American College of Cardiology and the American Heart Association (ACC/AHA) algorithm for cardiac evaluation and care. This is a step-wise approach that incorporates the following factors: (a) urgency of surgery (whether the procedure is an emergency); the (b) presence of active major cardiac risk factors (i.e., unstable coronary syndromes, decompensated heart failure, significant arrhythmias or severe valvular disease) that would necessitate referral to a cardiologist; (c) if risk factors for stable coronary artery disease are present, then calculation of risk for major adverse cardiac events using the ACS NSQIP calculator.
is recommended; (d) if the patient is at low risk for major cardiac events (<1%) then no further testing is needed; (e) if the patient is at elevated risk of major cardiac events, then determination of functional capacity with an objective measure or scale may be helpful; (f) if functional capacity is poor, then additional testing such as pharmacological stress testing, may be helpful; and finally (g) if testing does not impact care, then one should proceed to surgery or consider alternative treatment strategies. 

Routine electrocardiograms are not indicated in older patients undergoing low-risk surgery in the absence of other risk factors.

The combined effect of depletion of intravascular volume, age-related impairment of response to catecholamines, and increased myocardial relaxation time adversely affects the cardiac function of an older adult patient under stress in the perioperative period. Aging has been demonstrated to cause a decrease in cardiac output by approximately 1% per year. Older individuals fail to augment heart rate to the same extent as younger individuals. More importantly, the ability to increase cardiac output with aging is dependent on ventricular dilatation, which is determined by preload. Therefore, careful attention must be paid to volume status in the perioperative period. Over one half of all postoperative deaths in older adult patients and 11% of postoperative complications are a result of impaired cardiac function under physiologic stress. Incomplete emptying of the ventricle at end systole and subsequent reduction in ejection fraction is characteristic of the aging heart. Reduced distensibility, in addition to acute stressors, leads to impaired coronary perfusion and cardiac ischemia.

An important predictor of surgical outcomes and cardiac complications in the older adult is congestive heart failure (CHF). CHF is present in approximately 10% of patients older than 65 years and is the leading cause of postoperative morbidity and mortality. This prevalence will likely increase as percutaneous interventions and medical therapy prolongs survival from myocardial ischemia and acute myocardial infarction. Therefore, identifying correctable and uncorrectable cardiovascular disease is critical before elective surgical interventions.

Common pulmonary postoperative complications in older adults include atelectasis, pneumonia, and prolonged mechanical ventilation. These complications can contribute significantly to overall morbidity and mortality among older adults, affecting up to 15% of individuals ≥70 years in the postoperative setting. Risk factors for postoperative pulmonary complications include both patient and procedure-related factors. Patient-related factors include age >60, chronic obstructive pulmonary disease (COPD), American Society of Anesthesiologists (ASA) class II or greater, functional dependence, congestive heart failure, obstructive sleep apnea, pulmonary hypertension, current cigarette use, impaired sensorium, preoperative sepsis, weight loss >10% in 6 months, serum albumin <3.5 mg/dL, and blood urea nitrogen (BUN) ≥7.5 mmol/L (≥1.5 mg/dL). Surgical related factors include prolonged operation of greater than 3 hours, surgical site, emergency operation, general anesthesia, perioperative transfusion, and residual neuromuscular blockade after an operation. Of note, obesity, well-controlled asthma, and diabetes are not risk factors for postoperative pulmonary complications. It is recommended that patients with COPD and asthma that is not well controlled undergo preoperative optimization of pulmonary function and other general recommendations include smoking cessation, preoperative intensive inspiratory muscle training, and selective chest radiograph and pulmonary function testing.

In general, the use of routine preoperative screening combined with the high cost of unnecessary testing dispute the use of a routine battery of preoperative screening tests in all patients. Instead, it is preferable to perform selected tests in high-risk patients based on history, physical exam, known comorbidities, and the type of procedure being planned.

### Nutritional Assessment

Nutritional status should also be performed in older adults prior to surgery, as poor nutrition is potentially modifiable and related to increased risk of postoperative complications. The most common adverse events related to poor nutritional status are infectious complications (i.e., surgical site infections, pneumonia, urinary tract infections), wound complications (i.e., dehiscence and anastomotic leaks), and increased length of stay. The ACS NSQIP/AGS best practice guidelines recommend the following to screen for poor nutritional status:

1. **Document height and weight and calculate body mass index (BMI).** A BMI <18.5 kg/m² places an individual at risk and should prompt referral for full nutritional assessment.
2. **Measure baseline serum albumin and prealbumin levels.** Serum albumin <3.0 g/dL (with no evidence of hepatic or renal dysfunction) should prompt referral for full nutritional assessment.
3. **Inquire about unintentional weight loss in the last year.** Unintentional weight loss >10% to 15% in the past 6 months is associated with severe nutritional risk and should prompt assessment by a dietician.

The American Society for Parenteral and Enteral Nutrition (ASPEN) argues that measurement of serum albumin and prealbumin reflect the severity of the inflammatory response rather than true poor nutritional status. Instead, they favor a more individualized approach whereby the presence of any of the following six factors would classify an individual with malnutrition:

1. **Insufficient energy intake.** Severe malnutrition in the context of chronic illness is defined as <75% of estimated energy requirement for ≥1 month.
2. **Weight loss.** Severe malnutrition in the context of chronic illness is defined as >5% weight loss in 1 month, >7.5% weight loss in 3 months, >10% weight loss in 6 months, and >20% weight loss in 1 year.
3. **Loss of muscle mass.** Severe malnutrition is defined as severe muscle wasting of the temples, clavicles, shoulders, interosseous muscles, scapula, thigh, and calf.
4. **Loss of subcutaneous fat.** Severe malnutrition is defined as loss of subcutaneous fat (e.g., orbital, triceps, fat overlying the ribs).
5. **Localized or generalized fluid accumulation that may sometimes mask weight loss.** This can be demonstrated by fluid accumulation evident on exam (e.g., extremities, vulvar/scrotal edema, or ascites).
6. **Diminished functional status as measured by hand grip strength.** Measurements for grip strength are based on normative standards supplied by the manufacturer of the measurement device.

Regardless of the method used to measure nutritional status, attention to the diagnosis and management of malnutrition in the perioperative setting is imperative to optimize postoperative outcomes.
Psychosocial Considerations
Anxiety, depression, substance abuse, and social isolation are common, underdiagnosed conditions in older adults. Careful screening can identify these potential barriers to recovery, safe discharge after surgery, and maintenance of independence.

Medication Review
Careful review and documentation of the patient’s complete medication history is important in the preoperative setting. The review should include the use of nonprescription agents, including over-the-counter, nonsteroidal anti-inflammatory drugs (NSAIDs), vitamins, eye drops, topical agents, and herbal products.

Additionally, it is important to discontinue medications that should be avoided prior to surgery in order to minimize potential adverse events and interactions. These include discontinuation of all nonessential medications that are associated with increased surgical risk and medications with potential for drug interactions with anesthesia. Herbal medication should be stopped at least 7 days prior to surgery, and the Beers criteria should be reviewed. Medications with potential for withdrawal (e.g., selective serotonin reuptake inhibitor [SSRIs], tricyclic antidepressants, benzodiazepines, antipsychotics, monoamine oxidase inhibitors [MAOIs], β-blockers, clonidine, statins, and corticosteroids) in addition to angiotensin-converting enzyme inhibitors and angiotensin II receptor blockers should be continued unless hypertension is well controlled. Medications that are associated with increased risk for postoperative delirium should also be avoided and replaced with safer medications. For example, benzodiazepines should be stopped or reduced, where possible. Pain should be adequately controlled and meperidine should be avoided for pain management. Caution should be used when prescribing antihistamine H1 antagonists (such as diphenhydramine/Benadryl) and medications with antihistamine effects. Of note, there is no increased risk associated with neuroleptics (antipsychotics) and digoxin and no conclusive evidence against using H1 agonists, tricyclic antidepressants, antiparkinson medications, steroids, NSAIDs, and antimuscarinics in the perioperative setting.

Certain medications, such as β-blockers or statins, may be beneficial to start preoperatively in order to reduce risk of perioperative adverse events such as cardiac complications and stroke. There are ACC/AHA guidelines to support the use of perioperative β-blockers in patients who are already taking them, particularly those with independent cardiac indications such as arrhythmia or history of myocardial infarction and in patients who are undergoing intermediate risk or vascular surgery with no known coronary artery disease or with multiple clinical risk factors for ischemic heart disease. Statins should be started as soon as possible prior to surgery in patients who have known vascular disease, elevated low-density lipoprotein cholesterol, or ischemia on thallium testing. They may also be considered in patients undergoing vascular and intermediate-risk surgery.

Additionally, medications doses should be adjusted based on renal function. Older patients are at greater risk for renal impairment and chronic kidney disease; furthermore, adjustment of medications that are renally cleared is an important consideration. Glomerular filtration rate (GFR), rather than creatinine, is the best overall measure of renal function due to the fact that the ratio of GFR to creatinine decreases with increasing age.19

Finally, medication use is very common among the older population, and older individuals should be monitored for polypharmacy and potential adverse interactions. This is of real concern since 40% of individuals age 65 years and older are taking five or more medications prescribed by more than one doctor.37 Additionally, 68% of older adults use over-the-counter medications, dietary supplements, or both concurrently with prescription medications.38 Of note, polypharmacy has been associated with increased risk of cognitive impairment, morbidity, and mortality, and the risk of adverse drug reactions increases with a greater number of medications.39,40

Patient Counseling
Determination of the patient’s preferences and expectations prior to surgery is an essential component to preoperative assessment and decision making. One study demonstrated that agreement between patients and their surrogates (primary care providers and close family members) was poor, despite patients predicting that their physicians (90%) and family members (87%) would accurately represent their wishes. Instead, they found that percent of agreement ranged from 59% to 88%, suggesting that substituted judgement may not be a good proxy for an individual patient’s wishes.40 Another study looked at older patients with limited life expectancy due to cancer, congestive heart failure, or chronic obstructive pulmonary disease and found that 99% of patients would undergo a low-burden treatment to restore current health (with the alternative being death), but that 74% and 89% would forgo treatment if it resulted in severe functional or cognitive impairment, respectively.41 Furthermore, it is imperative that the surgeon have a substantive discussion with the patient prior to surgery to determine their preferences and expectations and that family members and potential decision-making surrogates be involved. The ACS NSQIP/AGS Guidelines recommend that the following four points be included in these conversations:

1. Ensure that the patient has an advance directive and designated health care proxy.
2. Discuss treatment goals and plans with the patient to ensure that the physician understands the patient’s preferences and expectations. This should be documented in the medical record.
3. The surgeon should describe the expected postoperative course and possible complications, including the potential for functional decline and need for rehabilitation or nursing home care, if relevant.
4. The physician should determine the patient’s family and social support systems. If support is insufficient, then referral to a social worker should be considered.42

The American College of Surgeons has created a checklist for the optimal preoperative assessment of the geriatric surgical patient (see Table 47-2).

PREOPERATIVE PREPARATION
In the immediate preoperative period, careful planning is essential to optimize the care of the frail older patient.

Patient Goals, Preferences, and Advance Directives
It is important that surgeons have a good understanding of patients’ goals and wishes surrounding their medical care, particularly towards the end of life. This should be established in the clinic setting prior to surgery and should be confirmed.
and documented throughout the process. Additionally, patients should be encouraged to designate a health care proxy to help with this process, should they be unable to make their own medical decisions. The healthcare team should also consider early palliative care consultation in individuals with poor prognoses who are electing to undergo surgery, particularly if they have a life expectancy of less than 6 months.

**Preoperative Fasting**

Historically, preoperative fasting began at midnight the night before elective surgery, whereby, patients were not permitted to have any oral intake of either liquids or solids. However, more recent literature suggests that there may be no clear benefit to extended periods of fasting greater than 6 hours. Based on the American Society of Anesthesiologists 2011 practice guidelines for all adults, fasting should take the form of stopping clear liquids at least 2 hours before elective procedures, stopping light food intake and/or nonhuman milk 6 hours before elective procedures and stopping fried, fatty foods and meat at least 8 hours before elective procedures. Of note, patients with comorbidities or diseases that can affect gastric emptying (i.e., diabetes, hiatal hernia) may require additional periods of fasting.

**Antibiotic Prophylaxis and Venous Thromboembolism Prevention**

Antibiotic prophylaxis for older adults should comply with standard guidelines put forth by The Society for Healthcare Epidemiology of America/Surgical Infection Society/American Society of Health-System Pharmacists/Infectious Disease Society. Older adults who receive appropriate preoperative antibiotics demonstrate a mortality benefit at 60 days. Appropriate antibiotics should be administered within 60 minutes prior to surgical incision.

Older adults are at higher risk for venous thromboembolism (VTE), making VTE risk stratification among this population essential. Older individuals undergoing orthopedic procedures (i.e., total hip or knee arthroplasty) or who have suffered a hip fracture should be treated with low molecular weight heparin (LMWH) (starting either 13 hours or more preoperatively or 12 hours or more postoperatively) for a minimum of 10 to 14 days and up to 35 days. Older adults undergoing nonorthopedic surgery (i.e., general, abdominopelvic, bariatric, vascular, plastic/reconstructive, and thoracic surgery) should have LMWH or low-dose unfractionated heparin (LDUH) and mechanical prophylaxis with intermittent pneumatic compression (IPC). Older adults undergoing craniotomy/spinal surgery or cardiac surgery should have IPC. Finally, older individuals who experienced major trauma and spinal cord injury should use LDUH/LMWH and IPC if not contraindicated.

**Surgical Prehabilitation**

One purpose of the preoperative assessment is to identify potentially modifiable risk factors in order to optimize surgical outcomes. Several prehabilitation programs have emerged in order to help meet this need and have demonstrated promising results. One of the first of such programs was referred to as the Proactive Care of Older People undergoing surgery (POPS) study in the United Kingdom’s National Health Service (NHS). This project was designed to decrease complications leading to increased hospital length of stay among at-risk older adults undergoing elective surgery. The authors performed a structured geriatric team intervention to identify at-risk patients and to then facilitate coordinated multidisciplinary care in the form of daily inpatient rounds, weekly multidisciplinary meetings, and biweekly ward rounds led by a POPS consultant/clinical nurse specialist. Outcomes in surgical patients undergoing the POPS intervention were compared to those among patients not undergoing the intervention. The POPS group had fewer postoperative complications, including lower rates of pneumonia (4% vs. 20%, \( P = 0.008 \)) and delirium (4% vs. 19%, \( P = 0.028 \)), better pain control (2% vs. 30%, \( P < 0.001 \)), lower rates of delayed mobilization (9% vs. 28%, \( P = 0.12 \)) and lower rates of inappropriate catheter use (7% vs. 37%, \( P = 0.046 \)). They also demonstrated a reduction in hospital length of stay by 4.5 days.

The Michigan Surgical Home and Optimization Program (MS SHOP) is another example of a successful prehabilitation program. This is a structured, home-based preoperative training program that targets physical, nutritional, and psychological interventions. The intervention included the following four components: (a) a home-based walking program with daily reminders and feedback; (b) incentive spirometry instructions starting one week prior to surgery; (c) education on nutrition, stress management, and care planning; and (d) resources for smoking cessation, when appropriate. Eighty-two percent of individuals enrolled in the study were actively engaged in the program. Compared to individuals who did not undergo the intervention, patients enrolled in the trial demonstrated a 31% reduction in hospital length of stay and a 28% reduction in cost. Collectively, the POPS study and the MSHOP studies are illustrative of the notion that attention to pre- and perioperative assessment in the older population result in improved postoperative outcomes that benefit the patient, hospital, and health care system.

**Palliative Care Services for Older Surgical Patients**

Among seriously ill individuals, palliative care services have the potential to increase quality of life, improve symptoms and patient satisfaction, and reduce caregiver burden. The role of palliative care in older surgical patients is not as well understood, and it is not widely understood in the surgical population. Over a decade ago, the American College of Surgeons Palliative Care Workgroup identified core competencies of surgical palliative care. The two key elements of palliative care—pain management and communication skills—are essential. For surgeons who frequently care for individuals at high risk of morbidity and mortality, there are six additional core competency domains: patient care, medical knowledge, practice-based learning, interpersonal skills, communication skills, and professionalism (Table 47-5).

Emerging data suggests that surgical patients benefit from the addition of palliative care principles and services. An interventional trial evaluating a decision-making intervention considering procedures that included information about health status and prognosis increased the likelihood of choosing less aggressive treatment options among patients with frailty (OR 3.41, 95% CI 1.39–8.39) or dementia (OR 1.66, 95% CI 1.06–2.64). In a study of preoperative care consultation in frail older adults, Ernst et al found that preoperative palliative care consultations were associated with reduced mortality. Several studies have found that postoperative palliative care improves symptoms, including uncertainty, symptom distress, and depression, and improves quality of life.
### Core Competencies in surgical palliative care

<table>
<thead>
<tr>
<th>DOMAIN</th>
<th>COMPETENCY</th>
</tr>
</thead>
</table>
| **Patient Care**              | Possess the capacity to guide the transition from curative and palliative goals of treatment to palliative goals alone based on patient information and preferences, scientific and outcomes evidence, and sound clinical judgment.  
Perfom an assessment and gather essential clinical information about symptoms, pain, and suffering.  
Perfom palliative procedures competently and with sound judgment to meet patient goals of care at the end of life.  
Provide management of pain and other symptoms to alleviate suffering.  
Communicate effectively and compassionately bad news and poor prognoses.  
Conduct a patient and family meeting regarding advance directives and end-of-life decisions.  
Exercise sound clinical judgment and skill in the withdrawal and withholding of life support. |
| **Medical knowledge**         | Acute and chronic pain management.  
Non-pain symptom management.  
Ethical and legal basis for advance directives, informed consent, withdrawal and withholding of life support, and futility.  
Grief and bereavement in surgical illness.  
Quality of life outcomes and prognostication.  
Role of spirituality at the end of life. |
| **Interpersonal and communication skills** | Surgeons must be competent and compassionate communicators with patients, families, and other health care providers. They should be effective in communicating bad news and prognosis and in redefining hope in the context of cultural diversity. The interdisciplinary nature of palliative care requires that the surgeon is skilled as both a leader and a member of an interdisciplinary team and maintains collegial relationships with other health care providers. |
| **Professionalism**           | Surgeons must maintain professional commitment to ethical and empathic care, which is patient focused, with equal attention to relief of suffering along with curative therapy. Respect and compassion for cultural diversity, gender, and disability is particularly important around rituals and bereavement at the end of life. Maintenance of ethical standards in the withholding and withdrawal of life support is essential. |
| **Systems-based practice**    | Surgeons must be aware and informed of the multiple components of the health care system that provide palliative and end-of-life care. Surgeons should be knowledgeable and willing to refer patients to hospice, palliative care consultation, pain management, pastoral care, social services, etc., and to understand resource utilization and reimbursement issues involved. |

---


---

**SPECIAL CONSIDERATIONS**

**Functional Recovery**

In the past decade, there has been increasing attention to the examination of functional outcomes after major surgery in the geriatric population. In a large prospective cohort study of patients age 60 years and older undergoing abdominal surgery, Lawrence et al found that older adults required several months to fully return to basic activities of daily living (ADLs) and up to 6 months to become independent in more complex instrumental activities of daily living (IADLs). Older adults are frequently discharged to postacute facilities even when they are functional dependent at baseline and have an uncomplicated postoperative course.66

Among frail older adults, functional decline after surgery is often substantial and sustained. Studies examining functional outcomes after surgery among nursing home residents have demonstrated that the majority of nursing home residents who undergo surgery do not return to baseline levels of function postoperatively. Among residents who underwent colectomy for cancer, 53% were dead after 1 year and over half of 1-year survivors experienced functional decline.67 For residents who undergo lower extremity bypass, half die within a year of surgery. At 1 year, 13% of the initial vascular surgery cohort was ambulatory, and 18% had maintained or improved their baseline functional status—calling into question the efficacy of this procedure in the nursing home. A study of hip fracture repair in nursing home residents found that over a third of residents died and over half of residents had died or experienced functional decline within 180 days after fracture. Residents with multiple comorbidities and advanced cognitive impairment and those who did not undergo surgical correction of the fracture experienced the worst outcomes.
Cancer Surgery
Approximately 50% of cancer diagnoses are currently made in patients age 70 years or older.66 It is predicted that the increase in the older adult population will account for up to a 50% increase in the number of patients undergoing oncologic procedures by the year 2020. The increased life expectancy of the geriatric patient coupled with the increasing incidence of cancer with advancing age will lead to an increased prevalence of malignant disease requiring surgical intervention. This is an area of great interest given that randomized clinical trials to determine the outcomes of older adult patients undergoing curative resections, as well as neoadjuvant and adjuvant therapy, are lacking. In addition, older adult patients are rarely included in clinical trials; therefore, treatment decisions are often based on individual surgeon experience and nongeriatric data, and they may be flawed by inherent biases regarding the outcome of complete oncologic resections in older adult patients. Surgeons may also be reluctant to expose older patients to the toxic effects of chemotherapy and radiation without proven efficacy in this geriatric population. This highlights the need for research targeting the specific needs of older adult patients with malignancy to aid in the development of specific treatment guidelines for various cancers within this age cohort.

Numerous studies have documented increased risk of postoperative morbidity and mortality in older adults with cancer. Evaluation of a national surgical registry found that older adults undergoing major gastrointestinal surgery have substantially higher risks of complications and death than individuals younger than 65 years. The impact of age on risk was present across all operations but had most impact in liver and rectal surgery. Surgeons are challenged to decide whether major surgery is justified in older adult patients, especially those with limited life expectancy. Effectiveness of oncologic surgery in older adult patients depends on whether a cure can be achieved safely without compromise to functional status or quality of life. Postoperative life expectancy should be improved by surgery, or, at the very least, not diminished.

Emergency Surgery
Emergent surgery carries exceptionally high risk for older adult patients. In an analysis of patients age 90 years and older, 90-day mortality after emergency gastrointestinal procedures was 54%.71 In a large cohort of patients undergoing endovascular repair for ruptured aortic aneurysm, 30-day mortality was 35% after primary aortic repair and 52% after open conversion of endovascular aortic repair.72

Frail institutionalized elders are at substantial risk for poor surgical outcomes after emergent surgery. In an analysis of over 70,000 nursing home residents who underwent emergent abdominal operations (surgery for bleeding ulcer, cholecystectomy, appendectomy, and colectomy), operative mortality was two- to threefold higher than among matched community-dwelling elders. In addition, invasive life-sustaining interventions after surgery were significantly higher in the nursing home population than among noninstitutionalized Medicare enrollees, ranging from 18% vs. 5%, respectively, after cholecystectomy to 55% vs. 43%, respectively, after ulcer surgery.

The combined effects of poor nutrition, decreased cognition, and immune impairments due to nutritional or pharmacologic factors create a treacherous circumstance for older adult patients with poorly defined symptoms or who present with more advanced disease. In acute abdominal conditions, such as acute appendicitis and acute cholecystitis, one-third of older adult patients will lack an elevated white blood cell count, one-third will lack fever, and one-third will lack physical findings of localized peritonitis.74 These deficits contribute to a threefold higher rate of perforated appendicitis and of gangrene of the gallbladder in older adult patients compared to young patients. An “unimpressive” physical exam in an older adult patient with acute onset of abdominal symptoms should never be taken as a sign of the absence of surgical disease.

Cardiovascular Surgery
With advances in cardiopulmonary bypass technique, myocardial protection, and improved perioperative care, coronary artery bypass grafting (CABG) and valve replacement operations have become safer in older patients. When considering cardiovascular surgery in elders, it is essential to consider that advanced age is not the strongest predictor of poorer outcomes or increased mortality compared to older patients. It has been demonstrated that emergency operations, preoperative New York Heart Association (NYHA) functional class 3 or greater, and chronic renal failure are the strongest independent predictors of increased operative mortality.75 In one study, preoperative renal dysfunction, cerebrovascular disease, valve surgery, and catastrophic state were independent predictors of increased mortality in older adult patients.76 Older adult patients with non–dialysis-dependent renal dysfunction had a 60% chance of death during a 5-year follow-up period compared to 25% in older adult patients without a history of renal dysfunction. Similarly, the presence of cerebrovascular disease resulted in a two-fold increase in mortality among older adult patients.76 Even patients who were 80 years of age or more did not have any significant increase in surgical risk and within this population, and the 4-year actuarial survival was 70.5% with an event-free survival of 60.6%.

There has been an increase in definitive operative intervention to older patients with operable coronary artery disease. The Society of Thoracic Surgeons reports that perioperative mortality rates range from 1.6% in patients 51 to 60 years of age to 7.7% in those 81 to 90 years of age.77 Older patients are more likely to have significant three-vessel disease accompanied by poor ejec-
tion fraction, left ventricular hypertrophy, significant valvular disease, and previous history of myocardial infarction than are younger patients.77 Older patients also are more likely to be classified as NYHA functional class 3 or higher and are more likely to present on an emergent basis, in part because of reluctance to provide elective surgical intervention because of presumptive poorer outcome. Despite the increased risk of morbidity and mortality compared to younger patients, older adult patients, including those >80 years old, can undergo CABG with acceptable mortality risk. The overall mortality rate is approximately 7% to 12% for older adult patients, including those in whom CABG is performed under emergency conditions. The mortality rate decreases to approximately 2.8% when CABG is performed electively with careful preoperative evaluation.

Valve Replacement
As the population ages, the incidence of senile calcific aortic stenosis and referral for aortic valve replacement are increasing. The operative mortality from aortic valve replacement is estimated to be between 3% and 10%, with an average of approximately 7.7%.78 If aortic stenosis is allowed to progress without operative intervention, CHF will ensue. The average survival of these patients is approximately 1.5 to 2 years. If a patient is deemed fit for operative intervention, age should not...
be a deterrent, especially considering the potential to increase life expectancy. It has been recommended that the carefully selected, minimally symptomatic octogenarian with aortic stenosis should be considered a low-risk patient and be expected to experience an uneventful operative course and expedient recovery. More importantly, if elective procedures are delayed until symptoms or left ventricular dysfunction develop, patients may suffer from unnecessary increased operative risk and mortality. Early intervention results in a demonstrable improvement in quality of life in these patients, with many improving their NYHA functional classification.

Older patients are candidates for mitral valve surgery when ischemic regurgitation is present. Surgery for mitral valve disease carries a higher morbidity and mortality risk than for aortic intervention, with an estimated mortality rate as high as 20%. Left ventricular function usually is compromised in patients requiring intervention, leading to a poorer outcome in these patients. The surgical outcome for mitral valve procedures depends on the extent of the disease, age of the patient, presence of pulmonary hypertension, and extent of coronary artery disease. The presence of comorbid conditions combined with the emergent nature of surgery in a large percentage of older patients further worsens the outcome. Therefore, a decision regarding management of mitral valve disease should be individualized to each patient. Another concern regarding older patients who are candidates for valve disease surgery is the additional need for coronary revascularization—an important contributor to morbidity and mortality from surgical intervention. To mitigate risk, an older patient with multiple comorbid conditions in need of a combined procedure should only have critically stenosed vessels bypassed. Neurologic complications from valve surgery are particularly common in older patients. It has been estimated that approximately 30% of patients >70 years old who undergo valve procedures develop either transient or permanent neurologic dysfunction. This often is a result of embolism from debris dislodged from the valve during the procedure or from a formed thrombus in the right atrium.

An important consideration in valve replacement procedures in older patients is the type of prosthesis to be used. Older patients are at increased risk from bleeding-associated anticoagulation complications. This risk is especially significant in patients who have experienced falls and minor trauma that have resulted in intracranial hemorrhage. To avoid the lifelong requirement for anticoagulants, bioprosthetic valves should be used in place of mechanical valves whenever possible. Although the bioprosthetic valves are not as durable as mechanical valves, studies demonstrate excellent structural integrity 10 years post procedure, making it an appropriate choice in older patients.

Transcatheter aortic valve implantation/replacement (TAVI/TAVR) is increasingly being used to treat aortic stenosis. Initially, this technique was reserved for individuals with high surgical risk. A systematic review of transcatheter aortic valve implantation versus surgical aortic valve replacement revealed that, compared to surgical repair, the transcatheter approach may have similar or better early and midterm outcomes, including among low- to intermediate-risk patients. Furthermore, there is increasing evidence that suggests TAVI results in acceptable long-term results in the older adult population.

Endovascular Aortic Surgery

With increasing use of screening abdominal CT scans and ultrasounds for evaluation of various abdominal complaints, abdominal aortic aneurysms (AAA) are being identified with greater frequency. The percentage of AAA rises from about 1% at age 55 to 60 years to approximately 10% in patients 80 years of age or older. Historically, very old patients were deemed poor operative candidates for the traditional open repair given the frequent presence of comorbid conditions and limited cardiopulmonary reserve to tolerate a major operation or the many hours of required operative time and general anesthesia. The dissemination of endovascular techniques for repair of AAA, however, has shifted the risk-benefit ratio for operative intervention, allowing greater life expectancy for the elective repair of this potentially life-threatening condition with the benefits of a minimally invasive approach.

Multiple studies have demonstrated that endovascular aortic repair (EVAR) is feasible and efficacious in older adult patients, including those previously considered unfit for open repair. EVAR is a minimally invasive technique in which a prosthetic graft is introduced into the aortic lumen via the common femoral artery to exclude the aortic aneurysm sac. EVAR significantly reduces operative and anesthesia times, blood loss, intensive care needs, length of stays, and major postoperative morbidity associated with open AAA repair. This procedure also can be done using epidural anesthesia for high-risk candidates who may tolerate general anesthesia poorly (Fig. 47-2).

Careful consideration of the life expectancy and the risk of rupture dictate the necessity for intervention. EVAR remains a viable option in older adult patients. Nonoperative management is justified in frail older adult patients with multiple comorbidities and reduced life expectancy whose operative risks outweigh the risk of rupture and in those who are unlikely to survive long enough to benefit from the repair.

Palliative Surgery

Palliative surgery is defined as surgical intervention targeted to alleviate a patient’s symptoms, thus improving the patient’s quality of life despite minimal impact on the patient’s survival. With an increasing number of older patients presenting with advanced disease, surgeons must be familiar with the concept of palliation to control symptoms. This concept focuses on providing the maximal benefit to the patient using the least-invasive intervention. Ideally, this intervention leads to symptom relief and preservation of the quality of life in terminal disease states by alleviating symptoms such as intractable vomiting and severe pain. The success of palliative surgery is a careful balance between achieving symptom relief without the development of new symptoms from the intervention itself. A recent meta-analysis of outcomes after palliative surgery for malignant bowel obstruction from peritoneal carcinomatosis revealed that although palliative surgery can benefit some patients, many patients experience serious complications, incomplete resolution of symptoms, and substantial hospitalization relative to the patient’s remaining survival time. It is essential to provide patients with realistic information about expected outcomes after palliative surgery to ensure that this surgical intervention is in line with their care preferences. The core competencies for surgical palliative care are shown in Table 47-5.

SUMMARY

Major surgery in older adults requires careful consideration. In addition to chronic medical conditions, many elders have geriatric syndromes that put them at high risk for increased morbidity,
mortality, and poor functional recovery after surgery. Screening for and optimization of multiple domains of vulnerability is essential to improve outcomes in this vulnerable population. Furthermore, the incorporation of palliative care principles into the surgical care of frail elders will improve patient-centered decision-making, symptom management, and quality of life.

REFERENCES

Entries highlighted in bright blue are key references.


Figure 47-2. Endovascular repair of abdominal aortic aneurysms (AAAs) has gained favor for suitable older adult patients to prevent rupture. Through minimal groin incisions, this 82-year-old patient underwent repair of an AAA and right iliac artery aneurysm and was discharged on post-op day 2.
PART II
SPECIFIC CONSIDERATIONS


This page intentionally left blank
Dedicated to the advancement of surgery along its scientific and moral side.

June 10, 1926, dedication on the Murphy Auditorium, the first home of the American College of Surgeons

**WHY ETHICS MATTER**

Ethical concerns involve not only the interests of patients but also the interests of surgeons and society. Surgeons choose among the options available to them because they have particular opinions regarding what would be good (or bad) for their patients. Aristotle described practical wisdom (Greek: *phronesis*) as the capacity to choose the best option from among several imperfect alternatives (Fig. 48-1). Frequently, surgeons are confronted with clinical or interpersonal situations in which there is incomplete information, uncertain outcomes, and/or complex personal and familial relationships. The capacity to choose wisely in such circumstances is the challenge of surgical practice.

**DEFINITIONS AND OVERVIEW**

Biomedical ethics is the system of analysis and deliberation dedicated to guiding surgeons toward the “good” in the practice of surgery. One of the most influential ethical “systems” in the field of biomedical ethics is the principalist approach as articulated by Beauchamp and Childress. In this approach to ethical issues, moral dilemmas are deliberated using four guiding principles: autonomy, beneficence, nonmaleficence, and justice.

The principle of autonomy respects the capacity of individuals to choose their own destiny, and it implies that individuals have a right to make those choices. It also implies an obligation for physicians to permit patients to make autonomous choices about their medical care. Beneficence requires that proposed actions aim at and achieve something good whereas nonmaleficence aims at avoiding concrete harm: *primum non nocere.* Justice requires fairness where both the benefits and burdens of a particular action are distributed equitably.

The history of medical ethics has its origins in antiquity. The Hippocratic Oath along with other professional codes has guided the actions of physicians for thousands of years. However, the growing technical powers of modern medicine raise new questions that were inconceivable in previous generations. Life support, dialysis, and modern drugs, as well as organ and cellular transplantation, have engendered new moral and ethical questions. As such, the ethical challenges faced by the surgeon have become more complex and require greater attention.

The case-based paradigm for bioethics is used when the clinical team encounters a situation in which two or more values or principles come into apparent conflict. The first step is to clarify the relevant principles (e.g., autonomy, beneficence, nonmaleficence, and justice) and values at stake (e.g., self-determination, quality of life). After identifying the principles and values that are affecting the situation, a proposed course of action is considered given the circumstances.

Much of the discourse in bioethics adopts this “principalist” approach in which the relevant principles are identified, weighed, and balanced, and then applied to formulate a course of action. This approach to bioethics is a powerful technique for thinking through moral problems because the four principles help identify what is at stake in any proposed course of action. However, the principles themselves do not resolve ethical dilemmas. Working together, patients and surgeons must use wise judgment to choose the best course of action for the specific case.

---

*“First do no harm.”*
Choosing wisely requires the virtue of practical wisdom first described by Aristotle (see Fig. 48-1). Along with the other cardinal virtues of courage, justice and temperance, practical wisdom is a central component of virtue ethics which complement principalist ethics by guiding choices toward the best options for treatment. Practical wisdom cannot be learned from books and is developed only through experience. The apprenticeship model of surgical residency fosters the development of practical wisdom through experience. More than teaching merely technical mastery, surgical residency is also moral training. In fact, the sociologist Charles Bosk argues that the "postgraduate training of surgeons is above all things an ethical training."3

### SPECIFIC ISSUES IN SURGICAL ETHICS

#### Informed Consent

Although a relatively recent development, the doctrine of informed consent is one of the most widely established tenets of modern biomedical ethics. During the nineteenth and early twentieth centuries, most physicians practiced a form of benign paternalism whereby patients were rarely involved in the decision-making process regarding their medical care, relying instead on the beneficence of the physician. Consensus among the wider public eventually changed such that surgeons are now expected to have an open discussion about diagnosis and treatment with the patient to obtain informed consent. In the United States, the legal doctrine of simple consent dates from the 1914 decision in Schloendorff vs. The Society of New York Hospital regarding a case in which a surgeon removed a diseased uterus after the patient had consented to an examination under anesthesia, but with the express stipulation that no operative excision should be performed. The physician argued that his decision was justified by the beneficent obligation to avoid the risks of a second anesthetic. However, Justice Benjamin Cardozo stated:

> Every human being of adult years and sound mind has a right to determine what shall be done with his body; and a surgeon who performs an operation without his patient’s consent commits an assault, for which he is liable in damages . . . except in cases of emergency, where the patient is unconscious, and where it is necessary to operate before consent can be obtained.4

Having established that patients have the right to determine what happens to their bodies, it took some time for the modern concept of informed consent to emerge from the initial doctrine of simple consent. The initial approach appealed to a professional practice standard whereby physicians were obligated to disclose to patients the kind of information that experienced surgeons customarily disclosed. However, this disclosure was not always adequate for patient needs. In the 1972

---

**Key Points**

1. The physician should document that the patient or surrogate has the capacity to make a medical decision.
2. Sufficient details regarding diagnosis and treatment options should be disclosed to the patient so that the patient can provide informed consent.
3. Living wills are written to anticipate treatment options and choices in the event that a patient is rendered incompetent by a terminal illness.
4. The durable power of attorney for healthcare identifies surrogate decision makers and invests them with the authority to make healthcare decisions on behalf of patients in the event that they are unable to speak for themselves.
5. Surgeons should encourage their patients to complete a living will and clearly identify their surrogates early in the course of treatment.
6. Earlier referrals and wider use of palliative and hospice care may help more patients achieve their goals at the end of life.
7. Seven requirements for the ethical conduct of clinical research studies have been articulated: value, scientific validity, fair subject selection, favorable risk-benefit ratio, independent review, informed consent, and respect for enrolled subjects.
8. Individuals working together on research endeavors should have clear discussions early in the planning process about authorship, and those discussions should be continued throughout the project or study.
9. Disclosure of error is consistent with recent ethical advances in medicine toward more transparency, openness with patients, and the involvement of patients in their care.

---

**Figure 48-1.** Bust of Aristotle. Marble, Roman copy after a Greek bronze original by Lysippos from 330 B.C. (From http://en.wikipedia.org/wiki/File:Aristotle_Altemps_Inv8575.jpg: Ludovisi Collection, Accession number Inv. 8575, Palazzo Altemps, Location Ground Floor, Branch of the National Roman Museum. Photographer/source Jastrow [2006] from Wikipedia.)
landmark case, *Canterbury vs. Spence*, the court rejected the professional practice standard in favor of the *reasonable person standard* whereby physicians are obliged to disclose to patients all information regarding diagnosis, treatment options, and risks that a “reasonable patient” would want to know in a similar situation. Rather than relying on the practices or consensus of the medical community, the reasonable person standard empowers the public (reasonable persons) to determine how much information should be disclosed by physicians to ensure that consent is truly informed. The court did recognize, however, that there are practical limits on the amount of information that can be communicated or assimilated.4 Subsequent litigation has revolved around what reasonable people expect to be disclosed in the consent process to include the nature and frequency of potential complications, the prognostic life expectancy,6 and the surgeon-specific success rates.4 Despite the litigious environment of medical practice, it is difficult to prosecute a case of inadequate informed consent so long as the clinician has made a concerted and documented effort to involve the patient in the decision-making process.

Adequate informed consent entails at least four basic elements: (a) the physician documents that the patient or surrogate has the capacity to make a medical decision; (b) the surgeon discloses to the patient details regarding the diagnosis and treatment options sufficiently for the patient to make an informed choice; (c) the patient demonstrates understanding of the disclosed information before (d) authorizing freely a specific treatment plan without undue influence (Fig. 48-2). These goals are aimed at respecting each patient’s prerogative for autonomous self-determination. To accomplish these goals, the surgeon needs to engage in a discussion about the causes and nature of the patient’s disease, the risks and benefits of available treatment options, as well as details regarding what patients can expect after an operative intervention including possible outcomes and complications.7-14

Certain clinical settings make obtaining informed consent challenging. For example, obtaining consent for emergency surgery can be difficult, as the clinical team is forced to make decisions with incomplete information. Emergency consent requires the surgeon to consider if and how possible interventions might save a patient’s life, and if successful, what kind of disability might be anticipated. Surgical emergencies are one of the few instances where the limits of patient autonomy are freely acknowledged, and surgeons are empowered by law and ethics to act promptly in the best interests of their patients according to the surgeon’s judgment. Most applicable medical laws require physicians to provide the standard of care to incapacitated patients, even if it entails invasive procedures without the explicit consent of the patient or surrogate. If at all possible, surgeons should seek the permission of their patients to provide treatment, but when emergency medical conditions render patients unable to grant that permission, and when delay is likely to have grave consequences, surgeons are legally and ethically justified in providing whatever surgical treatment the surgeon judges necessary to preserve life and restore health.4 This justification is based on the social consensus that most people would want their lives and health protected in this way, and this consensus is manifest in the medical profession’s general orientation to preserve life. It may be that subsequent care may be withdrawn or withheld when the clinical prognosis is clearer, but in the context of initial resuscitation of injured patients, incomplete information makes clear judgments about the patient’s ultimate prognosis or outcome impossible.

The pediatric population also presents unique challenges for the process of consent. For many reasons, children and adolescents cannot participate in the process of giving informed

![Figure 48-2. Algorithm for navigating the process of informed consent. (Modified with permission from Childers R, Lipsett A, Pawlik T. Informed consent and the surgeon, J Am Coll Surg. 2009 Apr;208(4):627-634.)](image-url)
consent in the same way as adults. Depending on their age, children may lack the cognitive and emotional maturity to participate fully in the process. In addition, depending on the child’s age, their specific circumstances, as well as the local jurisdiction, children may not have legal standing to fully participate on their own independent of their parents. The use of parents or guardians as surrogate decision makers only partially addresses the ethical responsibility of the surgeon to involve the child in the informed consent process. The surgeon should strive to augment the role of the decision makers by involving the child in the process. Specifically, children should receive age-appropriate information about their clinical situation and therapeutic options delivered in an appropriate setting and tone so that the surgeon can solicit the child’s “assent” for treatment. In this manner, while the parents or surrogate decision makers formally give the informed consent, the child remains an integral part of the process.

Certain religious practices can present additional challenges when treating minor children whose parents disallow medically indicated blood transfusions; however, case law has made clear the precedent that parents, regardless of their held beliefs, may not place their minor children at mortal risk. In such a circumstance, the physician should seek counsel from the hospital medicolegal team, as well as from the institutional ethics team. Legal precedent has, in general, established that the hospital or physician can proceed with providing all necessary care for the child.

Obtaining “consent” for organ donation deserves specific mention. Historically, discussion of organ donation with families of potential donors was performed by transplant professionals, who were introduced to families by intensivists after brain death had been confirmed and the family had been informed of the fact of death. In other instances, consent might be obtained by intensivists caring for the donor, as they were assumed to know the patient’s family and could facilitate the process. However, issues of moral “neutrality” as part of end-of-life care in the intensive care unit have caused a shift in how obtaining “consent” for organ donation is handled. Responsibility for obtaining consent from the donor family is now vested in trained “designated requestors” (or “organ procurement coordinators”) or by “independent” intensivists who do not have a therapeutic clinical relationship with the potential donor. In this way, the donor family can be allowed to make the decision regarding donation in a “neutral” environment without erosion of the therapeutic relationship with the treating physician or perceived undue pressure from the transplant team.

The process of informed consent also can be limited by the capacity of patients to assimilate information in the context of their illness. For example, despite the best efforts of surgeons, evidence suggests that patients rarely retain much of what is disclosed in the consent conversation, and they may not remember discussing details of the procedure that become relevant when postoperative complications arise. It is important to recognize that the doctrine of informed consent places the most emphasis on the principle of autonomy precisely in those clinical situations when, because of their severe illness or impending death, patients are often divested of their autonomy.

The Boundaries of Autonomy: Advance Directives and Powers of Attorney

Severe illness and impending death can often render patients incapable of exercising their autonomy regarding medical decisions. One approach to these difficult situations is to make decisions in the “best interests” of patients, but because such decisions require value judgments about which thoughtful people frequently disagree, ethicists, lawyers, and legislators have sought a more reliable solution. Advance directives of various forms have been developed to carry forward into the future the autonomous choices of competent adults regarding healthcare decisions. Furthermore, the courts often accept “informal” advance directives in the form of sworn testimony about statements the patient made at some time previous to their illness. When a formal document expressing the patient’s advance directives fails to exist, surgeons should consider the comments patients and families make when asked about their wishes in the setting of debilitating illness.

Living wills are written to anticipate treatment options and choices in the event that a patient is incapacitated by a terminal illness. In the living will, the patient indicates which treatments she wishes to permit or prohibit in the setting of terminal illness.

The possible treatments addressed often include mechanical ventilation, cardiopulmonary resuscitation, artificial nutrition, dialysis, antibiotics, or transfusion of blood products. Unfortunately, living wills are often too vague to offer concrete guidance in complex clinical situations, and the language (“terminal illness,” “artificial nutrition”) can be interpreted in many ways. Furthermore, by limiting the directive only to “terminal” conditions, it does not provide guidance for common clinical scenarios like advanced dementia, delirium, or persistent vegetative states where the patient is unable to make decisions, but is not “terminally” ill. Perhaps even more problematic is the evidence that demonstrates that healthy patients cannot reliably predict their preferences when they are actually sick. This phenomenon is called “affective forecasting” and applies to many situations. For example, the general public estimates the health-related quality of life (HRQoL) score of patients on dialysis at 0.39, although dialysis patients themselves rate their HRQoL at 0.56. Similarly, patients with colostomies rated their HRQoL at 0.92, compared to a score of 0.80 given by the general public for patients with colostomies. For these and other reasons, living wills are often unable to provide the extent of assistance they promise.

An alternative to living wills is the durable power of attorney for healthcare in which patients identify surrogate decision makers and invest them with the authority to make healthcare decisions on their behalf in the event that they are unable to speak for themselves. Proponents of this approach hope that the surrogate will be able to make decisions that reflect the choices that the patients themselves would make if they were able. Unfortunately, several studies demonstrate that surrogates are not much better than chance at predicting the choices patients make when the patient is able to state a preference. For example, a recent meta-analysis found that surrogates predicted patients’ treatment preferences with only 68% accuracy. These data reveal a flaw in the guiding principle of surrogate decision making: Surrogates do not necessarily have privileged insight into the autonomous preferences of patients. However, the durable power of attorney at least allows patients to choose the person who will eventually make prudent decisions on their behalf and in their best interests; therefore, respecting the judgment of the surrogate is a way of respecting the self-determination of the incapacitated patient.

There is continuing enthusiasm for a wider use of advance directives. In fact, the 1991 Patient Self Determination Act requires all U.S. healthcare facilities to (a) inform patients of
their rights to have advance directives, and (b) to document those advance directives in the chart at the time any patient is admitted to the healthcare facility. However, only a minority of patients in U.S. hospitals have advance directives despite concerted efforts to teach the public of their benefits and provide resources to help patients prepare and maintain them. For example, the ambitious SUPPORT trial used specially trained nurses to promote communication between physicians, patients, and their surrogates to improve the care and decision making of critically ill patients. Despite this concerted effort, the intervention demonstrated “no significant change in the timing of do not resuscitate (DNR) orders, in physician-patient agreement about DNR orders, in the number of undesirable days (patients’ experiences), in the prevalence of pain, or in the resources consumed.”

Some of the reluctance around physician-patient agreement about DNR orders may reflect patient and family anxiety that DNR orders equate to “do not treat.” Patients and families should be assured, when appropriate, that declarations of DNR/ do not intubate will not necessarily result in a change in ongoing routine clinical care. The issue of temporarily reserving DNR/ do not intubate orders around the time of an operative procedure may also need to be addressed with the family.

Patients should be encouraged to clearly identify their surrogates, both formally and informally, early in the course of treatment and before any major elective operation. Often, around the time of surgery or at the end of life, there are limits to patient autonomy in medical decision-making. Seeking an advance directive or surrogate decision maker requires time that is not always available when the clinical situation deteriorates. As such, these issues should be clarified as early as possible in the patient–physician relationship.

Withdrawing and Withholding Life-Sustaining Therapies

The implementation of various forms of life support technology raise a number of legal and ethical concerns about when it is permissible to withdraw or withhold available therapeutic technology. There is general consensus among ethicists that there are no philosophic differences between withdrawing (stopping) or withholding (not starting) treatments that are no longer beneficial. However, the right to refuse, withdraw, and withhold beneficial treatments was not established before the landmark case of Karen Ann Quinlan. In 1975, Quinlan lapsed into a persistent vegetative state requiring ventilator support. After several months without clinical improvement, Quinlan’s parents asked the hospital to withdraw ventilator support. The hospital refused, fearing prosecution for euthanasia. The case was appealed to the New Jersey Supreme Court where the justices ruled that it was permissible to withdraw ventilator support. This case established a now commonly recognized right to withdraw “extraordinary” life-saving technology if it is no longer desired by the patient or the patient’s surrogate.

The difference between “ordinary” and “extraordinary” care, and whether there is an ethical difference in withdrawing or withholding “ordinary” vs. “extraordinary” care, has been an area of much contention. The 1983 Nancy Cruzan case highlighted this issue. In this case, Cruzan had suffered severe injuries in an automobile crash that rendered her in a persistent vegetative state. Cruzan’s family asked that her tube feeds be withheld, but the hospital refused. The case was appealed to the U.S. Supreme Court, which ruled that the tube feeding could be withheld if her parents demonstrated “clear and convincing evidence” that the incapacitated patient would have rejected the treatment. In this ruling, the court essentially ruled that there was no legal distinction between “ordinary” vs. “extraordinary” life-sustaining therapies. In allowing the feeding tube to be removed, the court accepted the principle that a competent person (even through a surrogate decision maker) has the right to decline treatment under the Fourteenth Amendment of the U.S. Constitution. The court noted, however, that there has to be clear and convincing evidence of the patient’s wishes (consistent with the principle of autonomy) and that the burdens of the medical intervention should outweigh its benefits (consistent with the principles of beneficence and nonmaleficence).

In deliberating the issue of withdrawing vs. withholding life-sustaining therapies, the principle of “double effect” is often mentioned. According to the principle of “double effect,” a treatment (e.g., opioid administration in the terminally ill) that is intended to help and not harm the patient (i.e., relieve pain) is ethically acceptable even if an unintended consequence (side effect) of its administration is to shorten the life of the patient (e.g., by respiratory depression). Under the principle of double effect, a physician may withhold or withdraw a life-sustaining therapy if the surgeon’s intent is to relieve suffering, not to hasten death. The classic formulation of double effect has four elements (Fig. 48-3).

Withholding or withdrawing of life-sustaining therapy is ethically justified under the principle of double effect if the physician’s intent is to relieve suffering, not to kill the patient. Thus, in managing the distress of the dying, there is a fundamental ethical difference between titrating medications rapidly to achieve relief of distress and administering a very large bolus with the
intent of causing apnea. It is important to note, however, that although the use of opioids for pain relief in advanced illness is frequently cited as the classic example of the double effect rule, opioids can be used safely without significant risk. In fact, if administered appropriately, in the vast majority of instances the rule of double effect need not be invoked when administering opioids for symptom relief in advanced illness.28

In accepting the ethical equivalence of withholding and withdrawing of life-sustaining therapy, surgeons can make difficult treatment decisions in the face of prognostic uncertainty.24 In light of this, some important principles to consider when considering withdrawal of life-sustaining therapy include: (a) Any and all treatments can be withdrawn. If circumstances justify withdrawal of one therapy (e.g., IV pressors, antibiotics), they may also justify withdrawal of others; (b) Be aware of the symbolic value of continuing some therapies (e.g., nutrition, hydration) even though their role in palliation is questionable; (c) Before withdrawing life-sustaining therapy, ask the patient and family if a spiritual advisor (e.g., pastor, imam, rabbi, or priest) should be called; and (d) Consider requesting an ethics consult.

Although the clinical setting may seem limited, a range of options usually exists with respect to withdrawing or withholding treatment, allowing for an incremental approach, for example (a) continuing the current regimen without adding new interventions or tests; (b) continuing the current regimen but withdrawing elements when they are no longer beneficial; and (c) withdrawing and withholding all treatments that are not targeted to relieve symptoms and maximize patient comfort.34

The surgeon might consider discussing the clinical situation with the patient or proxy decision maker, identify the various therapeutic options, and delineate the reasons why withholding or withdrawing life-sustaining therapy would be in the patient’s best interest. If the patient (or designated proxy decision maker) does not agree with withholding or withdrawing life-sustaining therapy, the surgeon should consider involving consultants who have participated in the patient’s care, experts in palliative or end-of-life care or recommend a second medical opinion. If the second opinion corroborates that life-sustaining therapy should be withheld or withdrawn but the patient/family continues to disagree, the surgeon should consider assistance from institutional resources such as the ethics committee and hospital administration. Although the surgeon is not ethically obligated to provide treatment that he or she believes is futile, the surgeon is responsible for continued care of the patient, which may involve transferring the patient to a surgeon who is willing to provide the requested intervention.24

Living Donor Liver Transplantation

One unique ethical issue that deserves special mention is that of living donor liver transplantation. Living donor kidney transplantation has been practiced for almost 50 years and has become a routine part of clinical care, but living donor liver transplantation was first performed in the late 1980s with parent-to-child grafts and in the late 1990s for adult-to-adult grafts. These procedures are unique in that there are two patients, one with a diseased organ who requires intervention to be made well and one who is healthy and is being made unwell, albeit usually temporarily, during the intervention. Performing an ethical analysis of this situation requires considering both risks and benefits to each of the patients individually.

For the recipient, the benefits of receiving a living donor organ as opposed to a deceased donor organ are many: first, there is reduced risk of death on the waitlist, and second, there is a potential for improved post-transplant outcomes due to improved matching between relatives and the absence of hemodynamic instability often present before organ procurement in a deceased donor.30 Furthermore, the use of living donor organs is supported by the principal of utility, maximizing efficient use of organs.32

The benefit to the organ donor is in fulfillment of an altruistic ideal and satisfaction associated with having extended the recipient’s life, while the risks are those associated with partial hepatectomy, a procedure that is not without risks including postoperative complications and mortality, the risk of which is estimated to be 0.15%.29 The ethical concern in this case is having possibly violated the principle of nonmaleficence.

This particular ethical issue emphasizes the importance of truly informed consent. The donor should be provided with information on local complication and mortality rates and allowed sufficient time to consider the risks and benefits without pressure from healthcare workers.30 Furthermore, experienced centers have recommended that living donors have access to sufficient resources and strong support from an institution’s ethics committee, given substantial pressure exerted by the critical illness of a family member.31

Palliative Care

General Principles of Palliative Care
Palliative care is a coordinated, interdisciplinary effort that aims to relieve suffering and improve quality of life for patients and their families in the context of serious illness.33 It is offered simultaneously with all other appropriate medical treatment, and its indication is not limited to situations associated with a poor prognosis for survival. Palliative care strives to achieve more than symptom control, but it should not be confused with noncurative treatment.

The World Health Organization defines palliative care as “an approach that improves the quality of life of patients and their families facing the problems associated with life-threatening illness, through the prevention and relief of suffering by means of early identification and impeccable assessment and treatment of pain and other problems, physical, psychosocial, and spiritual.”34 Palliative care is both a philosophy of care and an organized, highly structured system for delivering care.

Palliative care includes the entire spectrum of intervention for the relief of symptoms and the promotion of quality of life. No specific therapy, including surgical intervention, is excluded from consideration. Therefore, surgeons have valuable contributions to make to palliative care. In fact, the term palliative care was coined in 1975 by Canadian surgeon, Balfour Mount. Furthermore, surgical palliative care can be defined as the treatment of suffering and the promotion of quality of life for seriously or terminally ill patients under the care of surgeons.36 The standard of palliative treatment lies in the agreement between patient and physician that the expected outcome is relief from distressing symptoms, lessening of pain, and improvement of quality of life. The decision to intervene is based on the treatment’s ability to meet the stated goals, rather than its impact on the underlying disease.

The fundamental elements of palliative care consist of pain and nonpain symptom management, communication among patients, their families, and care providers, and continuity of care across health systems and through the trajectory
of illness. Additional features of system-based palliative care are team-based planning that includes patient and family; close attention to spiritual matters; and psychosocial support for patients, their families, and care providers, including bereavement support.

Indications for palliative care consultation in surgical practice include: (a) patients with conditions that are progressive and life-limiting, especially if characterized by burdensome symptoms, functional decline, and progressive cognitive deficits; (b) assistance in clarification or reorientation of patient/family goals of care; (c) assistance in resolution of ethical dilemmas; (d) situations in which a patient/surrogate declines further invasive or curative treatments with stated preference for comfort measures only; (e) patients who are expected to die imminently or shortly after hospital discharge; (f) provision of bereavement support for patient care staff, particularly after loss of a colleague under care (Table 48-1). Although all patients, regardless of prognosis, may benefit from the services of a palliative care physician, hospice care is a specific form of palliative care intended for patients who have an estimated prognosis of 6 months or less to live. Hospice care is covered under Medicare Part A, and benefits may be continued beyond the original 6 months of estimated survival if physicians certify that life expectancy remains limited to 6 months or less.

Although most Americans indicate a preference to die at home, nearly 75% die in an institutional setting. Earlier referral and wider use of the hospice benefit may help more patients achieve their goal of dying at home.

Concepts of Suffering, Pain, Health, and Healing

Palliative care specifically addresses the individual patient’s experience of suffering due to illness. Indeed, the philosophical origins of palliative care began with attention to suffering and the existential questions suffering engenders. More than mere technologic evolution in the management of symptoms, the early proponents of palliative care sought a revolution in the moral foundations of medicine that challenged the assumptions that so often seemed to result in futile invasive intervention, and identified many of the problems that were subsequently taken up by medical ethicists. This reorientation of the goals of medical care from a focus on disease and its management to the patient’s experience of illness focuses attention on the purpose of medicine and the meaning of health and healing.

Over the past half century, several concepts and theories about the nature of pain, suffering, and health have been proposed in service of the evolving conceptual framework of palliative care. For example, while considering the differences between disease-oriented and illness-oriented approaches to the care of seriously ill patients, psychiatrist Arthur Kleinman wrote, “There is a moral core to healing in all societies. [Healing] is the central purpose of medicine . . . the purpose of medicine is both control of disease processes and care for the illness experience. Nowhere is this clearer than in the relationship of the chronically ill to their medical system: For them, the control of disease is by definition limited; care for the life problems created by the disorder is the chief issue.”

The relief of pain has been the clinical foundation for hospice and palliative care. Pain is defined by the International Association for the Study of Pain as “an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage.” For purposes of interdisciplinary palliative care, Saunders’s concept of “total pain” is a more useful definition and is frequently used as the basis for palliative assessments. Total pain is the sum total of four principal domains of pain: physical, psychologic, social or socioeconomic, and spiritual. Each of these contributes to, but is not synonymous with, suffering.

Effective Communication and Negotiating the Goals of Care

Changing the goals of care from cure to palliation near the end of life can be both emotionally and clinically challenging. It depends on determination of a clear prognosis and can be aided by effective communication. Unfortunately, prognostication can be notoriously difficult and inaccurate in advanced illness, and Christakis has argued that, to a large degree, physicians have abdicated their traditional responsibility to provide clear prognosis regarding incurable disease and approaching death. However, there are validated tools for prognosis in critical illness (APACHE, MODS, etc.), and with most advanced diseases, functional status is the most powerful predictor of survival. For example, patients with advanced metastatic cancer who are resting or sleeping for 50% or more of normal waking hours and require some assistance with activities of daily living (ADL) have a projected survival of weeks, and patients who are essentially bedfast and dependent for ADL have a projected survival of days to a week or two at best. Table 48-2 shows a simple prognostic tool to aid clinicians in recognizing patients nearing the end of life.

Alternatively, the Karnofsky Performance Scale is a scale of functional status ranging from 100 (high level of function) to 0 (death). It is commonly used in palliative care to roughly assess a patient’s anticipated needs as well as prognosis. The Palliative Performance Scale is a validated expansion of the Karnofsky Performance Scale that includes five palliative-focused domains, including ambulation, activity level, self-care, intake, and level of consciousness, in addition to evidence of disease. The Missoula-Vitas Quality of Life Index is a 25-question scale specifically for palliative care and hospice patients that scores symptoms, function, interpersonal relationships, well-being, and spirituality. Updates and Spanish versions are available.

Regardless of the prognostic tool used, the prognosis should be conveyed to the patient and family. If done well, communication and negotiation with patients and families about advanced terminal illnesses can potentially avoid great
psychologic harm and help make a difficult transition easier. To communicate effectively and compassionately, it is helpful to pursue an organized process similar to the structured history and physical central to the evaluation of any patient. One such structured approach to delivering unfavorable news proposes six steps that can be easily learned by clinicians: (a) getting started by selection of the appropriate setting, introductions, and seating; (b) determining what the patient or family knows; (c) determining what the patient or family wants to know; (d) giving the information; (e) expressing empathy; and (f) establishing expectations, planning, and aftercare (Table 48-3). Success with this approach to breaking bad news is critically dependent upon the clinician’s ability to empathically respond to the patient’s (and family’s) reaction to the news. The empathic response does not require the surgeon to share the same emotions of the patient, but it does require the surgeon to identify the patient’s emotion and accurately reflect that awareness back to the patient. Such effective communication may be facilitated by involving other members of the healthcare team who have developed relationships with the patient and their family. Patient assessment in these conversations should give the highest priority to identifying and responding to the most immediate source of distress. Relieving a pressing symptom is prerequisite for a more thorough search for other potential sources of suffering, and the assessment process itself can be therapeutic if conducted in a respectful and gentle manner.

CARE AT THE END OF LIFE

The process of dying and the care of a patient at the time of death is a distinct clinical entity that demands specific skills from physicians. The issues specific to dying and the available tools for compassionate care at the end of life are addressed in this section.

The Syndrome of Imminent Demise

In a patient who has progressed to the terminal stage of an advanced illness (e.g., cancer), a number of signs provide evidence of imminent death. As terminally ill patients progress toward death, they become increasingly bedbound, requiring assistance for all basic ADL. There is a steady decrease in desire and requests for food and fluids. More distressing to the dying patient is a progressively dry mouth that may be confused by the treating team as thirst. It is often exacerbated by anticholinergic medications, mouth breathing, and supplemental oxygen (O2) administered without humidification.

With progressive debility, fatigue, and weight loss, it is common for terminally ill patients to experience increasing difficulty swallowing. This may result in aspiration episodes and an inability to swallow tablets, requiring alternative routes for medication administration (e.g., IV, SC, PR, sublingual, buccal, or transdermal). In addition to the increased risk of aspiration, patients near death develop great difficulty clearing oropharyngeal and upper airway secretions, leading to noisy breathing or the so-called “death rattle.” As death approaches, the respiratory pattern may change to increasingly frequent periods of apnea often following a Cheyne-Stokes pattern of rapid, progressively longer breaths leading up to an apneic period. As circulatory instability develops near death, patients may exhibit cool and mottled extremities. Periods of confusion are often accompanied by decreasing urine output and episodes of fecal and urinary incontinence.

A number of cognitive changes occur as death approaches. Patients who are in the last days of life may demonstrate some signs of confusion or delirium. Agitated delirium is a prominent feature of a difficult death. Other cognitive changes that may be seen include a decreased interest in social interactions, increased somnolence, reduced attention span, disorientation to time (often with altered sleep-wake cycles), and an altered dream life, including vivid “waking dreams” or visual hallucinations. Reduced hearing and visual acuity may be an issue for some patients; however, patients who appear comatose may still be aware of their surroundings. Severely cachectic patients may lose the ability to keep their eyes closed during sleep because of loss of the retro-orbital fat pad.

Common Symptoms at the End of Life and Their Management

The three most common, major symptoms that threaten the comfort of dying patients in their last days are respiratory

---

Table 48-2

<table>
<thead>
<tr>
<th>Functional Level</th>
<th>Performance Status (ECOG)</th>
<th>Prognosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Able to perform all basic ADLs independently and some IADLs</td>
<td>2</td>
<td>Months</td>
</tr>
<tr>
<td>Resting/sleeping up to 50% or more of waking hours and requiring some assistance with basic ADLs</td>
<td>3</td>
<td>Weeks to a few months</td>
</tr>
<tr>
<td>Dependent for basic ADLs and bed-to-chair existence</td>
<td>4</td>
<td>Days to a few weeks at most</td>
</tr>
</tbody>
</table>

These observations apply to patients with advanced, progressive, incurable illnesses (e.g., metastatic cancer refractory to treatment). Basic ADL = activities of daily living (e.g., transferring, toileting, bathing, dressing, and feeding oneself); IADL = instrumental activities of daily living (e.g., more complex activities such as meal preparation, performing household chores, balancing a checkbook, shopping, etc.); ECOG = Eastern Cooperative Oncology Group functional (performance) status.

Table 48-3

<table>
<thead>
<tr>
<th>Communicating unfavorable news: important principles</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Setting: Find a quiet, private place to meet. Sit down close to the patient.</td>
</tr>
<tr>
<td>• Listen: Clarify the patient’s and/or the family’s understanding of the situation.</td>
</tr>
<tr>
<td>• “Warning shot”: Prepare patient and family and obtain their permission to communicate bad news (e.g., “I’m afraid I have bad news.”).</td>
</tr>
<tr>
<td>• Silence: Pause after giving bad news. Allow patient/family to absorb/react to the news.</td>
</tr>
<tr>
<td>• Encourage: Convey hope that is realistic and appropriate to the circumstances (e.g., patient will not be abandoned; symptoms will be controlled).</td>
</tr>
</tbody>
</table>
distress, pain, and cognitive failure. General principles that are applicable to symptom management in the last days of life include (a) anticipating symptoms before they develop; (b) minimizing technologic interventions (usually manage symptoms with medications); and (c) planning alternative routes for medications in case the oral route fails. It may be possible to cautiously reduce the dose of opioids and other medications as renal clearance decreases near the end of life, but it is important to remember that increased somnolence and decreasing respirations are prominent features of the dying process independent of medication side effects. Sudden cessation of opioid analgesics near the end of life could precipitate withdrawal symptoms, and therefore medications should not be stopped for increasing somnolence or slowed respirations.

The principles of pharmacotherapy for pain and non-pain symptoms in the palliative care setting are outlined in Table 48-4. The World Health Organization, the United States Agency for Healthcare Policy and Research, the Academy of Hospice and Palliative Medicine, and many other agencies have endorsed a “step ladder” approach to cancer pain management that can predictably result in satisfactory pain control in most patients (Table 48-5). More refractory pain problems require additional expertise, and occasionally, more invasive approaches (Tables 48-6 and 48-7).

The primary treatment of dyspnea (air hunger) in the dying is opioids, which should be cautiously titrated to increase comfort and reduce tachypnea to a range of 15 to 20 breaths per minute. Air movement across the face generated by a fan can sometimes be quite helpful. If this is not effective, empirical use of supplemental O₂ by nasal cannula (2–3 L/min) may bring some subjective relief, independent of observable changes in pulse oximetry. Supplemental O₂ should be humidified to avoid exacerbation of dry mouth. Typical starting doses of an immediate release opioid for breathlessness should be one-half to two-thirds of a starting dose of the same agent for cancer pain. For patients already on opioids for pain, a 25% to 50% increment in the dose of the current immediate release agent for breakthrough pain often will be effective in relieving breathlessness in addition to breakthrough pain.

The availability and variety of drugs should not prevent consideration of nonpharmacologic therapy. Massage therapy, music therapy, art therapy, guided imagery, hypnosis, physical therapy, pet therapy, and others play a constructive role not only for the relief of symptoms but also for promoting a sense of hope through improving function, aesthetic pleasure, and social connectedness. Talents and capacities neglected during the treatment and progression of disease can be recovered even in the most advanced stages of illness.

Pain is often less of a problem in the last days of life because the reduced activity level is associated with lower incident pain. This, combined with lower renal clearance of opioids, may result in greater potency of the prescribed agents. Severe pain crises are fortunately rare, but when they are inadequately addressed, can cause great and lasting distress (complicated grief) for loved ones who witness the final hours or days of agony. Such situations may require continuous administration of parenteral opioids. As death approaches and patients become less verbal, it is important to assess pain frequently, including the use of close observation for nonverbal signs of distress (e.g., grimacing, increased respiratory rate). Adequate dosing of opioid analgesics may require alternate route(s) of administration other than oral as patients become more somnolent or develop swallowing difficulties. Opioids should not be stopped abruptly, even if the patient becomes nonresponsive, because sudden withdrawal can cause severe distress.

Cognitive failure at the end of life is manifested in most patients by increasing somnolence and delirium. Gradually increasing somnolence can be accompanied by periods of disorientation and mild confusion, and it may respond to the reassuring presence of loved ones and caregivers with minimal need for medications. A more distressing form of delirium also can

---

**Table 48-4**

**Principles of pharmacotherapy in palliative care**

- Believe patient report of symptoms.
- Modify pathologic process when possible and appropriate.
- In terminally ill patients, avoid medications not directly linked to symptom control.
- Use a multidisciplinary approach.
- Consider nonpharmacologic approaches whenever possible.
- Engage participation of clinical pharmacist in treatment plan.
- Select drugs that can multitask (i.e., use haloperidol for agitated delirium and nausea).
- For pain, use adjuvant medications when possible (see Table 48-7).
- When using opioids, spare when possible (adjuvant medication, local or regional anesthetics, surgical interventions, etc.).
- Avoid fixed combination drugs.
- Avoid excessive cost.
- Select agents with minimum side effects.
- Anticipate and prophylax against side effects.
- For older adult patients, the hypoproteinemic, the azotemic: “Start low and go slow.”
- Oral route whenever possible and practical.
- No intramuscular injections.
- Scheduled dosing, not as needed, for persistent symptoms.
- Stepwise approach. (See the World Health Organization Analgesic Ladder for pain, Table 48-5.)
- Reassess continuously and titrate to effect.
- Use equianalgesic doses when changing opioids (see Table 48-5).
- Assess the patient’s and family’s comprehension of management plan.

---

**Table 48-5**

**The World Health Organization’s three-step ladder for control of cancer pain**

<table>
<thead>
<tr>
<th>Step 1: mild pain (visual analogue scale, 1–3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonopioid ± adjuvant medication</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 2: moderate pain (visual analogue scale, 4–6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opioid for mild to moderate pain and nonopioid ± an adjuvant</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 3: severe pain (visual analogue scale, 7–10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opioid for moderate to severe pain ± nonopioid ± an adjuvant</td>
</tr>
<tr>
<td>DRUG</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td><strong>Mild persistent pain, visual analogue scale (VAS) 1–3</strong></td>
</tr>
<tr>
<td><strong>Aspirin</strong> 600–1500 mg PO four times a day</td>
</tr>
<tr>
<td><strong>Choline magnesium trisalicylate (Trilisate)</strong> 750–1500 mg PO twice a day</td>
</tr>
<tr>
<td><strong>Ibuprofen (Advil, Motrin)</strong> 200–400 mg PO four times a day Maximum = 3200 mg/24 h</td>
</tr>
<tr>
<td><strong>Naproxen (Naprosyn)</strong> 250 mg PO twice a day Maximum = 1300 mg/24 h</td>
</tr>
<tr>
<td><strong>Moderate persistent pain, VAS 4–6</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Severe persistent pain, VAS 7–10</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Risk factors for NSAID-induced nephropathy include: advanced age, decreased glomerular filtration rate, congestive heart failure, hypovolemia, pressors, hepatic dysfunction, concomitant nephrotoxic agents. Dose reduction and hydration reduce risk. Opioids compounded with aspirin or acetaminophen are limited to treatment of moderate persistent pain because of dose-limiting toxicities of acetaminophen and aspirin. Slow-release preparations of morphine and oxycodone may be given rectally. Timed-release tablets or patches should never be crushed or cut. Opioid analgesics are the agents of choice for severe cancer-related pain. Sedation is a common side effect when initiating opioid therapy. Tolerance to this usually develops within a few days. If sedation persists beyond a few days, a stimulant (methylphenidate 2.5–5 mg PO twice a day) can be given. Initiate bowel stimulant prophylaxis for constipation when prescribing opioids unless contraindicated. Adjutant or coanalgesic agents are drugs that enhance analgesic efficacy of opioids, treat concurrent symptoms that exacerbate pain, or provide independent analgesia for specific types of pain (e.g., a tricyclic antidepressant for treatment of neuropathic pain). Coanalgesics can be initiated for persistent pain at any visual analogue scale level. Gabapentin is commonly used as an initial agent for neuropathic pain. No place for meperidine (Demerol), propoxyphene (Darvon, Darvocet, or mixed agonist-antagonist agents [Stadol, Talwin]) in management of persistent pain. Always consider alternative approaches (axial analgesia, operative approaches, etc.) when managing severe persistent pain. Note: These are not recommendations for specific patients. The inter- and intraindividual variability to opioids requires individualizing dosing and titration to effect. Adapted with permission from Cameron JL: *Current Surgical Therapy*, 9th ed. Philadelphia, PA: Elsevier; 2008.
Table 48-7

Examples of adjuvant medications for treatment of neuropathic, visceral, and bone pain

<table>
<thead>
<tr>
<th>DRUG CLASS</th>
<th>INITIAL DOSING (ADULT, &gt;60 kg)</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tricyclic antidepressants</td>
<td>Amitriptyline 10–25 mg PO before bed  &lt;br&gt; Nortriptyline 10–25 mg PO one per day &lt;br&gt; Doxepin 10–25 mg PO before bed &lt;br&gt; Imipramine 10–25 mg PO one per day</td>
<td>Sedating properties may be useful for relief of other concurrent symptoms. Side effects may precede benefit. Avoid in older adult patients due to anticholinergic side effects. Less anticholinergic effect</td>
</tr>
<tr>
<td>Anticonvulsants</td>
<td>Gabapentin 100–1200 mg PO three times a day. Titrate up rapidly as needed. Max: 3600 mg daily in divided doses &lt;br&gt; Carbamazepine 200 mg PO every 12 hours &lt;br&gt; Pregabalin starting dose 25–50 mg PO three times a day &lt;br&gt; Valproic acid 250 mg PO three times a day</td>
<td>Commonly used first-line agent. Generally well tolerated. Does not require blood level monitoring. Effective. Well studied. Requires blood monitoring. Does not require blood monitoring.</td>
</tr>
<tr>
<td>Local anesthetics</td>
<td>Lidocaine transdermal patch 5%. Apply to painful areas. Max: 3 simultaneous patches over 12 hours (each patch contains 700 mg lidocaine). Lidocaine/prilocaine topical. Apply to painful areas.</td>
<td>Systemic toxicity can result from applying more than recommended number per unit time and in patients with liver failure. Effective for postherpetic neuralgia.</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>Bisphosphonates (pamidronate, zoledronic acid) &lt;br&gt; Calcitonin nasal spray &lt;br&gt; Dexamethasone &lt;br&gt; Radionuclides (Sr-89) &lt;br&gt; Octreotide</td>
<td>For bone pain and reduced incidence of skeletal complications secondary to malignancy—best results in myeloma and breast cancer. Contraindicated in renal failure. Refractory bone pain &lt;br&gt; For bone pain, acute nerve compression, visceral pain secondary to tumor infiltration or luminal obstruction by reducing inflammatory component of tumor &lt;br&gt; For malignant bone pain secondary to osteoclastic activity. 4–6 wk delay in benefit. Requires adequate bone marrow reserve. For prognosis of more than 3 mo. Reduces GI secretions that contribute to visceral pain</td>
</tr>
</tbody>
</table>

*Recommendations are based on experience of practitioners of hospice and palliative medicine and in some instances do not reflect current clinical trials.
develop, manifested by increasing agitation that may require the use of neuroleptic medications. Increasing amounts of opioids and/or benzodiazepines may exacerbate the delirium (especially in the elderly).

**Pronouncing Death**

If the body is hypothermic or has been hypothermic, such as a drowning victim pulled from the water in the winter, the physician should not declare death until warming attempts have been made. In the hospital, hospice, or home setting, the declaration of death becomes part of the medical or legal record of the event. There are a number of physical signs of death a physician should look for in confirming the patient’s demise: complete lack of responsiveness to verbal or tactile stimuli, absence of heart beat and respirations, fixed pupils, skin color change to a waxyen hue as blood settles, gradual poikilothermia, and sphincter relaxation with loss of urine and feces. For deaths in the home with patients who have been enrolled in hospice, the hospice nurse on call should be contacted immediately. In some states, deaths at home may require a brief police investigation and report. For deaths in the hospital, the family must be notified (in person, if possible). A coroner or medical examiner may need to be contacted under specific circumstances (e.g., deaths in the operating room), but most deaths do not require their services. The pronouncing physician will need to complete a death certificate according to local regulations. Survivors may also be approached, if appropriate, regarding potential autopsy and organ donation. Finally, it is important to accommodate religious rituals that may be important to the dying patient or the family. Bereavement is the experience of loss by death of a person to whom one is attached. Mourning is the process of adapting to such a loss in the thoughts, feelings, and behaviors that one experiences after the loss. Although grief and mourning are accentuated in the immediate period around death, it is important to note that patients and families may have begun the process of bereavement well before the time of death as patients and families grieve incremental losses of independence, vitality, and control. In addition to the surviving loved ones, it is important to acknowledge that caregivers also experience grief for the loss of their patients.

**Aid in Dying**

Five European countries, Canada, and six U.S. states have legalized physician-assisted suicide, medical assistance in dying, or aid-in-dying, in some form, ranging from hospital-based programs to provision of fatal doses of medications for home self-administration. Medical assistance in dying is a complex ethical and legal issue with divergent opinions among the public and healthcare providers. While aid-in-dying laws passed in the United States vary somewhat, these laws essentially all allow physicians to prescribe a lethal dose of medication to mentally, competent, terminally ill adult patients for the purpose of achieving the end of life. Key areas of ethical consideration in this area include the benefit and harm of death; the relationship between passive euthanasia, active euthanasia, withholding treatment, and withdrawing treatment; the morality of physician and nursing participation in deliberately causing death; and the management of conscientious objection. Although surgeons outside of the critical care arena may only infrequently be asked to participate in aid-in-dying, it is important to be familiar with local legislation so that appropriate information can be provided to patients who request it.

**PROFESSIONAL ETHICS: CONFLICT OF INTEREST, RESEARCH, AND CLINICAL ETHICS**

**Conflict of Interest**

Conflicts of interest for surgeons can arise in many situations in which the potential benefits or gains to be realized by the surgeon are, or are perceived to be, in conflict with the responsibility to put the patient’s interests before the surgeon’s own. Conflicts of interest for the surgeon can involve actual or perceived situations in which the individual stands to gain monetarily by his or her role as a physician or investigator. In the academic community, monetary gain may not be the primary factor. Instead, motivators such as power, tenure, or authorship on a publication may serve as potential sources of conflict of interest. For example, the accrual of subjects in research studies or patients in surgical series may ensure surgeons better authorship or more financial gains. The dual-role of the surgeon-scientist therefore needs to be considered because the duty as surgeon can conflict with the role of scientist or clinical researcher.

**Research Ethics**

Over the last three decades in the United States, the ethical requirements for the conduct of human subject research have been formalized and widely accepted. Although detailed informed consent is a necessary condition for the conduct of ethically good human subject research, other factors also determine whether research is designed and conducted ethically. Emanuel and colleagues described seven requirements for all clinical research studies to be ethically sound: (a) value—enhancement(s) of health or knowledge must be derived from the research; (b) scientific validity—the research must be methodologically rigorous; (c) fair subject selection—scientific objectives, not vulnerability or privilege, and the potential for and distribution of risks and benefits, should determine communities selected as study sites and the inclusion criteria for individual subjects; (d) favorable risk-benefit ratio—with the context of standard clinical practice and the research protocol, risks must be minimized, potential benefits enhanced, and the potential benefits to individuals and knowledge gained for society must outweigh the risks; (e) independent review—unaffiliated individuals must review the research and approve, amend, or terminate it; (f) informed consent—individuals should be informed about the research and provide their voluntary consent; and (g) respect for enrolled subjects—subjects should have their privacy protected, the opportunity to withdraw, and their well-being monitored.

**Special Concerns in Surgical Research**

A significant issue for clinical surgical research is that many surgical studies are retrospective in nature and are not commonly undertaken in a prospective, double-blind, randomized fashion. For a randomized trial to be undertaken, the researchers should be in a state of equipoise—that is, there must be a state of genuine uncertainty on the part of the clinical investigator or the expert medical community regarding the comparative therapeutic merits of each arm in a trial. To randomize subjects to receive two different treatments, a researcher must believe that the existing data are not sufficient to conclude that one treatment strategy is better than another. In designing surgical trials, surgeons usually have biases that one treatment is better than another and often have difficulty maintaining the state of equipoise. As such, it is frequently difficult to demonstrate that a
Surgical Innovation

An important issue is whether surgical innovation should be treated as research or as standard of care. Throughout history, many advances in surgical techniques and technologies have resulted from innovations of individual surgeons crafted during the course of challenging operations—such innovations and technologies have served to move the field of surgery forward. In the Korean and Vietnam wars, military guidelines for treatment of vascular injuries recommended ligation and amputation rather than interposition grafting of vascular injuries. Individual surgeons chose to ignore those guidelines and subsequently demonstrated the value of the reconstructive techniques that ultimately became the standard of care. It is debated whether modifications to an accepted surgical technique in an individual patient based on their circumstances and within the skill and judgment of an individual surgeon should require the same type of prior approval that enrollment in a clinical trial would warrant. However, if a surgeon decides to use a new technique on several occasions and to study the outcomes, Institutional Review Board approval and all other ethical requirements for research are necessary. These situations require strict oversight as well as explicit consent by the patient. In particular, when developing new and innovative techniques, the surgeon should work in close consultation with his or her senior colleagues, including the chairperson of the department. Frequently, more senior individuals can provide sage ethical advice regarding what constitutes minor innovative changes in a technique vs. true novel research.

Compared to the formalized process for new drug approval by the Food and Drug Administration, the process for a surgeon developing an innovative operation can be relatively unregulated and unsupervised.

The Ethics of Authorship

Authorship specifies who is responsible for published research. It confers both recognition for academic achievement as well as responsibility for the academic integrity of the published content. Authorship is the stock in trade of productivity for academic surgeons, and it plays a significant role in promotion and tenure. It can also be commodified in the form of intellectual property and patents in which the author and the author’s institution have vested interests. Yet it can also become a liability if a given piece of work becomes embroiled in accusations of plagiarism, data fabrication, or other academic misconduct.

In the past, criteria for authorship were unspecified: Those submitting manuscripts simply listed the authors with little or no need to substantiate their contribution to the work. Unfortunately, this informal process led to confusion and even abuse. For example, there has been a long tradition of awarding authorship to the investigator who supervised or obtained funding for research, regardless of that person’s specific contribution to the manuscript. However, current recommendations specify that supervision and funding, by themselves, are insufficient criteria for authorship, and thus such individuals should only be included as authors if they make direct contributions to the work. A more disturbing example is the practice of “ghost writing” by which senior investigators publish industry-written research under their own name to bolster their productivity while providing a luster of academic integrity to industry.

To address these conflicts of interest and to provide guidance to investigators, the International Committee of Medical Journal Editors (ICMJE) provides recommendations on criteria for authorship so that individuals who contributed to the intellectual content of a work get appropriate credit and that all those listed as authors take responsibility and are accountable for the published work. The ICMJE recommendations for authorship can be found in Table 48-8. Furthermore, the ICMJE recommends that each author should be able to identify the contribution that each other author made to the work and be confident regarding the integrity of their co-authors. The ICMJE also recommends that individuals who do not meet these criteria be acknowledged in the manuscript, providing appropriate procedures for such acknowledgement. Additionally, the ICMJE criteria for authorship

<table>
<thead>
<tr>
<th>ICMJE criteria for authorship</th>
</tr>
</thead>
</table>
| According to ICMJE best practices recommendations, authors should fulfill each of the following four criteria:
| 1. Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work |
| 2. Drafting the work or revising it critically for important intellectual content |
| 3. Final approval of the version to be published |
| 4. Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved |

Contributors who do not fulfill all four criteria should be named in the manuscript in the acknowledgment section.

specifically excludes certain types of contributions including acquisition of funding, general supervision of a research group, administrative support, writing assistance, technical editing, language editing, and proofreading. 25

Many journals have adopted these criteria, operationalizing them at the time of submission by having each author specify his or her contributions. These contributions are then disclosed in the published manuscript to further specify how credit and responsibility is shared. 77 This approach has been shown to provide valuable information and has proved feasible in several journals, including The Lancet. 78

As research becomes increasingly interdisciplinary with ever-expanding teams of contributors, it can be difficult to determine which contributions warrant full authorship rather than simple acknowledgement. Individuals working together on research endeavors should have clear discussions early in the planning process about authorship, and those discussions should be continued throughout the project or study.

Clinical Ethics: Disclosure of Errors
Disclosure of error—either in medical or research matters—is important, but often difficult (see Chapter 12). Errors of judgment, errors in technique, and system errors are responsible for most errors that result in complications and deaths. Hospitals are evaluated based on the number of complications and deaths that occur in surgical patients, and surgeons traditionally review their complications and deaths in a formal exercise known as the mortality and morbidity conference, or M&M. The exercise places importance on the attending surgeon’s responsibility for errors made, whether he or she made them themselves, and the value of the exercise is related to the effect of “peer pressure”—the entire department knows about the case—on reducing repeated occurrences of such an error. Although a time-honored ritual in surgery, the M&M conference is nonetheless a poor method for analyzing causes of error and for developing methods to prevent them. Moreover, the proceedings of the M&M conference are protected from disclosure by the privilege of “peer review,” and the details are thus rarely shared with patients or those outside the department.

A report from the United States Institute of Medicine titled “To Err Is Human” highlighted the large number of medical errors that occur and encouraged efforts to prevent patient harm. 79 Medical errors are generally considered to be “preventable adverse medical events.” 80 Medical errors occur with some frequency, and the question is what and how should patients be informed that a medical error has occurred. 81

Disclosure of error is consistent with the ethical virtue of candor (e.g., transparency and openness) and the ethical principle of respect for persons by involving patients in their care. In contrast, failing to disclose errors to patients undermines public trust in medicine and potentially compromises adequate treatment of the consequences of errors and effective intervention to prevent future errors. In addition, failure to self-disclose medical errors can be construed as a breach of professional ethics, as it is a failure to act in the patient’s best interests. Information regarding a medical error may be needed so that patients can make independent and well-informed decisions about future aspects of their care. The principles of autonomy and justice dictate that surgeons need to respect individuals by being fair in providing accurate information about all aspects of their care—even when an error has occurred.

Disclosing one’s own errors is therefore part of the ethical standard of honesty and putting the patient’s interests above one’s own. Disclosing the errors of others is more complicated and may require careful consideration and consultation. Surgeons sometimes discover that a prior operation has included an apparent error; an injured bile duct or a stenotic anastomosis may lead to the condition for which the surgeon is now treating the patient. Declaring a finding as an “error” may be inaccurate, however, and a nonjudgmental assessment of the situation is usually advisable. When clear evidence of a mistake is at hand, the surgeon’s responsibility is defined by his or her obligation to act as the patient’s agent.

REFERENCES
Entries highlighted in bright blue are key references.

60. Sharpe JT. Is there a significant moral distinction between active and passive euthanasia? Critique. 2011;5:11-16.


76. Eggert LD. Best practices for allocating appropriate credit and responsibility to authors of multi-authored articles. *Front Psychol.* 2011;2:196.


INTRODUCTION

Modern surgery can save lives, help expand economies, and offer hope to individuals and communities. Prior to the acceptance and availability of aseptic technique to prevent or decrease infections, and improved anesthesia for controlling pain, surgery as a specialty was held in very low esteem by medical doctors and the general public. Over the last 100 years, surgery has developed into a highly regarded discipline that not only provides opportunities for curing certain diseases but also fulfills a special role in preventing and mitigating disability.

Yet, surgery is currently unavailable to most people worldwide. The vast majority—90%—of the world’s population receives only 10% of the surgical care delivered. Said another way, 90% of surgical resources are consumed by the most privileged 10% of the world’s population. More than 5 billion people lack access to safe, timely, and affordable surgical care.1 Very few surgical procedures occur in countries spending less than U.S. $100 per person on health care per year compared to countries spending greater than U.S. $1000 per person (Fig. 49-1).2

Examples of disparities abound. In many countries, including the wealthiest, islands of poverty coexist within cities replete with material resources. Tertiary level hospitals operate within eyesight of slums whose inhabitants have no access to even basic care. Most of the people without access—people in rural areas and in countries with poor infrastructure—are the very people most at risk for death or disability due to lack of surgical care. Often the poor accept and endure many painful and potentially correctable fatal conditions as a fact of life.3,6 Care for trauma and obstetrical emergencies is considered a basic surgical need but is absent in many rural regions. Other chronic conditions—often equally debilitating—progress to death or serious disability due to lack of available, safe surgery and anesthesia.

Many factors contribute to the disparity in access to surgical care. Poverty, a primary risk factor for all types of diseases, is a major obstacle hindering access to surgery. Healthcare professionals, including surgeons, migrate from areas of need due to a lack of infrastructure (hospitals, roadways, and stable electrical sources), limited supplies and equipment, lack of human resources, few opportunities for professional development, and concerns for personal safety. Until recently, there has been a significant lack of information regarding the burden of surgical disease and surgery’s positive impact on communities. Current research substantiates that investment in surgical care improves economies and is an integral and necessary component of global health.7,8

Disparities in care and outcomes are multidimensional, and no simple solution exists to improve access to appropriate and affordable surgical care. Yet, five major forces are reshaping priorities and strategies leading the charge for the globalization of surgical care.

1. The epidemiologic transition of diseases from primarily infectious to more chronic conditions
2. The mobile nature of the world’s populations, allowing people to move freely between more isolated areas of the world, leading to a more integrated global community
3. Ubiquitous information access exponentially enabling widespread participation in understanding and designing innovative opportunities for high-quality surgical care
4. A revolution for equity and human rights where the world’s poor are demanding benefits to surgical care similar to those found in high-income countries (HICs)
5. Recognition of the cost-effectiveness of surgical care and its potential to build economies, demonstrating the value of including surgery in global health strategies6,12

The greatest burden of disease occurs in areas where human resources—physicians, nurses, pharmacists, and other healthcare workers—are scarce (Fig. 49-2).13 The proportion of physicians is low both in high-population areas and in areas where the population is growing most rapidly (Fig. 49-3).14,15 Fully trained surgeons and anesthesiologists comprise only a small proportion of the total number of the Human Resources in Health (HRH), and efforts to meet the
There are five major forces reshaping priorities and strategies for the globalization of surgical care:

a. The epidemiologic transition of diseases
b. The mobile nature of the world’s populations
c. Ubiquitous information access
d. A revolution for equity and human rights
e. Recognition of the cost-effectiveness of surgical care for treatment and prevention of disease

The burden of disease is greatest in areas where human resources—physicians, nurses, pharmacists, and other healthcare workers—are the least.

Surgery should be viewed as an investment rather than a cost.

The key components of the global surgery ecosystem include technology, education, community, healthcare, business, and multidisciplinary engagement between a variety of disciplines.

Understanding and addressing the necessary communication, energy, and transportation technologies along with the underlying cultural context represent the foundation critical to implementing sustainable infrastructure for appropriate surgical care.

There has been a significant shift from communicable, maternal, neonatal, and nutritional causes of disease to noncommunicable causes, many of which require surgical care.

The cost-effectiveness of surgical care has been demonstrated, and its value as a public health investment is increasingly understood by policymakers.

Developing capabilities for surgical care has the ability to promote system-strengthening in resource-poor countries and to mitigate migration of health professionals at all levels.

Academic global surgery provides a unique environment to study health systems, identify solutions and implement them collaboratively, fulfilling many institutions’ missions to strengthen multidisciplinary training, advocacy, and research.

Surgical innovations that bring value by balancing cost with quality designed for challenging energy environments will foster equity in surgical care for LMICs.

Patients and their communities in low- and middle-income countries (LMICs) bear a much greater share of the burden of cancer than high-income countries (HICs).

Globally, trauma has become a leading cause of death and disability; 90% of trauma deaths occur in LMICs.

Essential surgical services should be integrated into comprehensive health care delivery, with the potential to avert 1.5 million deaths per year in LMICs.

Surgery is gaining an increasingly recognized role for improving public health, having a role in prevention as well as treatment.

The cost-effectiveness of surgical care has been demonstrated, and its value as a public health investment is increasingly understood by policymakers.

Developing capabilities for surgical care has the ability to promote system-strengthening in resource-poor countries and to mitigate migration of health professionals at all levels.

Academic global surgery provides a unique environment to study health systems, identify solutions and implement them collaboratively, fulfilling many institutions’ missions to strengthen multidisciplinary training, advocacy, and research.

Surgical innovations that bring value by balancing cost with quality designed for challenging energy environments will foster equity in surgical care for LMICs.

The potential benefits of surgical care for economic productivity are astounding. Considering that the annual economic loss from road traffic injuries alone exceeds U.S. $500 billion globally, a panel of expert economists at the Copenhagen Consensus of 2012, including four Nobel prize laureates, prioritized strengthening surgical capacity as the eighth most cost-effective investment for addressing the world’s most pressing problems. The Consensus reconvened in 2015, synchronizing their recommendations to the United Nations’ Sustainable Development Goals, and reaffirmed that surgery-related initiatives (i.e., circumcision and skilled obstetrical support), offer the best “value-for-money” in terms of alleviating world poverty. The Lancet Commission on Global Surgery, a research and advisory working group with contributors from 110 nations, echoed these sentiments. Surgery should be viewed as an investment rather than a cost.

Much of the economic modeling and strengthening of political will related to surgical and anesthesia care has been carried out by collaborative groups and consortia including academic, nongovernmental, and other organizations. The efforts of these groups truly coalesced in 2015, when several major consensus statements from governmental organizations, such as the World Bank and the World Health Assembly, recognized the importance of surgery in public health for the first time.

The third edition of World Bank’s Disease Control Priorities (DCP3) and the World Health Assembly’s Resolution 68.15 both specifically discuss the vital nature of surgical care as part of any health system.7,21 Such resolutions represent a sea change in terms of how the global policy community views surgical care for LMICs and, indeed, the entire world.

This chapter examines the ongoing need to expand and strengthen surgical care globally, explores some of the significant challenges of global surgery, and presents potential guiding concepts along with examples of successful strategies for sustainable surgical development.

DEFINING GLOBAL SURGERY

Global Surgery Ecosystem

To understand how surgery fits into healthcare systems and to understand its unique needs, it is helpful to consider global surgery as an ecosystem. The emerging field of global surgery considers surgical care to be a fundamental component of global health. As a system with both local and international scope, global surgery encompasses not just the medical and technical aspects of surgical care, but also the societal and environmental context in which surgery is performed. Global surgery also refers to a worldwide lens through which we view challenges collaboratively; thus, global surgeons may focus on resource-limited areas where needs are profound, but the ultimate goal should be to make surgical care equitable, accessible, and affordable for every human being.22 Surgery as an ecosystem considers the diverse but interrelated systems that must be functional for quality surgical care to be delivered. Only part of these systems falls within the traditional training of surgeons. Yet, modern surgical care requires these systems to work in a coordinated fashion to support three priorities critical for expanding surgery globally—accessibility, affordability, and innovation (Fig. 49-4). Global surgery is a way to consider a “systems-based practice” beyond a single hospital or community, for the benefit of people worldwide. Many interrelated components of this surgical ecosystem originate outside the hospital.

Disparities in surgical care have geographical, socioeconomic, and cultural components. Most people who live in major cities in the northern and western hemispheres take for granted a functioning energy grid. The development of energy beyond major cities has enabled wealthier communities to imagine, and indeed, to expect healthcare to be available at all times and affordable. Yet, a lack of reliable energy sources is a major limiting factor. Communication and transportation technologies, for example, the mobile phone and air and ground travel, have dramatically progressed in high- and middle-income countries but are still rudimentary in poor countries. Many of the current disparities in health care, particularly surgical care, are due to the lack of penetration of these technologies. Understanding and addressing the necessary communication, energy, and transportation deficits as well as the underlying cultural nuances are necessary to support the sustainable development of surgical care.

Electricity is necessary for all modern surgery. Anesthesia monitoring, operating room lighting, cautery, suction, and patient warming devices all require sources of electricity that are stable, without huge electrical surges. Only in the last 50 years or so could stable electricity be expected in most wealthy cities. However, in rural areas of even wealthy countries, electricity remains unpredictable (Fig. 49-5).23

In poorer countries, the cost and availability of electricity is frequently the limiting factor for more advanced diagnostic and therapeutic technology—from laboratories that require refrigeration to radiology in all of its various branches. Modern design for surgical devices has, for the most part, not taken into account the wide range of energy environments where surgery is practiced. Fragile instruments and monitors that cannot survive the rigors of the real working environment limit the types of surgery that can be provided.
Components of the Global Surgery Ecosystem. Improvements in energy, transportation, and communication are critical to support the growth of surgical care. Building capacity for surgical care requires interaction between the various components that create a functioning, sustainable system. When surgeons think of surgery, they usually think in terms of science and hands-on technical expertise. However, global surgery requires a broader understanding of systems in other disciplines. Surgeons must work collaboratively with engineers and business leaders to develop technology that can function in lower...
resource environments. These innovations can provide a source of economic growth for the community, which in turn supports better health care (Fig. 49-6).

No sustainable surgical system in the modern age can function without specialists in bioengineering, sterile process, supply chain, hospital safety, and waste management. These often unappreciated colleagues make possible the daily practice of surgery. Similarly, specialists in anesthesia, nursing, and the diagnostic specialties of radiology, pathology, and laboratory services are fundamental to a fully functional surgical service.

**Human Resources**

Primary care physicians, nurses, midwives, or advanced care practitioners (ACPs) provide much of the basic surgical and anesthetic care in LMICs. Where regulations allow, “task sharing,” or training ACPs to deliver surgery and anesthesia services previously allowed only under the purview of fully trained specialists, can provide expanded access to care.24-26 Non-MD practitioners, known as assistant medical officers (AMOs) or “tecnicos de cirurgia” in Mozambique, often have extensive operative experience, including obstetrical care, and are the primary surgical providers in some regions.27-29 Task sharing with ACPs also occurs in the United States and other countries where they fill a need otherwise unmet by specialists even in major tertiary care centers.30 However, concerns about the quality of care, lack of adequate supervision, and the effect on prestige and professional development for specialists and ACPs, continue to be topics for debate.31,32

Migration of practitioners to economically and culturally favorable locales is universal and not restricted to low-resource countries.33,34 However, the net impact on poor countries is greater. In a 2004 study, more than 23% of U.S. physicians received their medical training from other countries; of these 64% were from low-income countries.35 Using 2013 data, another study showed annual emigration rates of sub-Saharan physicians to the United States are increasing, despite a World Health Organization Global Code of Practice in 2010 aimed at LMIC workforce retention.36 Investments in training greater numbers of doctors in these countries, including surgical specialists, have been only partially successful in meeting demand in poor countries. Until economic conditions improve or opportunities for professional development increase, and incentives enticing migration of health care workers to high-income countries abate, it is unlikely that the most skilled practitioners will remain in resource-poor areas beyond their immediate obligations.37-41

**Burden of Surgical Disease**

**Epidemiologic Transition of Disease.** The population on Earth currently stands at more than 7 billion. While the rate of growth has slowed in recent years, projections estimate that...
the population will continue to grow to 9 billion by 2050.\textsuperscript{42} Population characteristics are changing rapidly. According to United Nations’ estimates, the entire world is aging even in low-income countries, and by 2050, 2 billion people will be over the age of 60.\textsuperscript{43} Just before the year 2020, the percentage of the world’s population over age 65 years is predicted to surpass the percentage of children under age 5 years, on an unprecedented reversal of trajectories for both age demographics. While this represents a victory for infectious disease control, the dramatic increase in longevity will present new challenges in terms of treating noncommunicable disease in the older adult population.\textsuperscript{44} At the same time, Sub-Saharan Africa’s population is experiencing a much different trend: a current “baby boom” will lead the region to quadruple its population, from 960 million to 4 billion, by the year 2100.\textsuperscript{45}

Until recently, infectious diseases dominated public health strategy. Now with major scourges like polio isolated to relatively small regions of the world, and HIV and malaria decreasing in their relative impact worldwide, chronic diseases and their complications, as well as the effects of aging, are gaining dominance in health care needs. Many of these chronic diseases are best approached by surgery.\textsuperscript{6}

The lack of metrics and paucity of data identifying the unmet burden of surgical need in many countries have been obstacles facing global surgery initiatives. The 2010 Global Burden of Disease Study was the first worldwide comprehensive burden of disease evaluation since the initial 1990 epidemiologic study. Using the disability-adjusted life year (DALY), a metric that captures both premature mortality and the prevalence and severity of illnesses, disease burdens were calculated for 291 causes in 21 regions of the world (including 187 countries) for 1990, 2005, and 2010 to enable identification of significant trends over time.\textsuperscript{46} While the global DALYs remained stable from 1990 to 2010, the study identified a significant shift from communicable, maternal, neonatal, and nutritional causes of disease to noncommunicable causes (Fig. 49-7).\textsuperscript{47}

In 2015, the previous estimate by the second edition of Disease Control Priorities in 2006 of an 11% global surgical disease burden was updated to 30%, obtained from provider-based survey data from the Lancet Commission.\textsuperscript{1,10} Using country-wide population surveys (the Surgeons OverSeas Assessment of Surgical Need Survey [SOSAS]) in Sierra Leone, Rwanda, and Nepal, the overall presence of surgically treatable conditions was 11.2%, with 25.6% of deaths potentially avoidable had surgical care been available. Applying these percentages to the 48 low-income countries, as defined by the World Bank, suggests that there are 288.2 million people currently living with surgically treatable conditions; providing improved access to surgical care could prevent 5.6 million deaths per year.\textsuperscript{48-50} Untreated acute and chronic surgical conditions represent a significant unmet burden of disease that has major impact on the economies of these nations.\textsuperscript{7,16,51}

\textbf{Cancer.} Patients and their communities in LMICs bear a much greater share of the burden of cancer than HICs. The dramatic increase in the proportion of reported cancer cases in LMICs is a result of population growth, aging populations, and decreased mortality from infectious diseases. In 1970, only 15\% of newly reported cancer cases worldwide were from the developing world; by 2008, this proportion rose dramatically to 58\% and is expected to grow to 70\% by 2030.\textsuperscript{52} Since 2013, the second leading cause of death worldwide has been cancer, and an estimated 20\% of all global surgery is now cancer-related.\textsuperscript{53} Previously thought to be a
disease almost exclusive to high-income countries, nearly two-thirds of the 7.6 million cancer deaths worldwide occur in LMICs. Mortality from cancer correlates inversely with a country’s economy for certain treatable cancers, including breast, testicular, and cervical cancer—LMICs have higher case fatality rates than HICs (Fig. 49-8).52,54

For example, breast cancer case fatality rates illustrate the great disparity in outcomes between regions. Case fatality rates in East Africa reach an unacceptable 59% compared to 19% in the United States.54 In LMICs, patients have very limited access to screening. They present for care with much later stages of cancer. In Haiti, after the great earthquake in 2010, with its initial onslaught of orthopedic injuries, many aid organizations found themselves faced with the unmet underlying burden of disease, including late-stage breast cancer and other tumors (Fig. 49-9). The DCP3 has devoted an entire chapter to cancer screening in LMICs, emphasizing the importance of proper infrastructure for screening and treatment, as well as considering cost-effectiveness and ethical concerns related to screening and subsequent treatment of detected cancers.55 The number and quality of training programs in surgical oncology is also inversely related to a country’s income, leaving LMICs with few adequately trained providers. Collaborative training programs between HIC and LMIC centers, as well as tele-teaching and mobile consultation, may address this shortage in a relatively low-cost, high-impact way.56

Trauma. Trauma has become a leading cause of death (5.8 million people per year) and disability around the world; 90% of trauma deaths occur in LMICs.57 Approximately 32% more people die as a result of injuries than from malaria, tuberculosis, and HIV/AIDS combined, representing 10% of the world’s deaths (Fig. 49-10).58,59 The major causes of death from injuries are road traffic accidents (RTAs), suicides, homicides, falls, drownings, and burns; in every category except burns, almost twice as many men die compared to women.60
Over 1.25 million people die from RTAs, causing LMICs to lose 3% of their GDP; 50 million more people incur nonfatal injuries, many with resulting lifelong disabilities. Globally, RTAs are the main cause of death for young people between the ages of 15 and 29. Forty-nine percent of all traffic deaths are among pedestrians, cyclists, and motorcycles. In the United States, a patient presenting with an injury in a rural community has a higher mortality than those from an urban setting. This disparity is much more pronounced in economically disadvantaged societies, where seriously injured patients from road traffic accidents are twice as likely to die compared to similarly injured patients in a high-income setting (Fig. 49-11). Additionally, death is much more likely to occur in the prehospital settings for injured patients from low-income countries. The lack of integrated communication and emergency transportation systems contribute to prehospital risk, while the lack of infrastructure, supplies, and personnel contribute to in-hospital mortality.

The number of deaths from RTAs has remained the same between 2007 and 2013. The predicted increase in mortality from RTAs, expected from the increase in population and global motorization, did not materialize, suggesting that interventions to improve global road safety (i.e., The Decade of Action for Road Safety 2011–2020) may be having some success in preventing deaths from RTAs.

Burns. The World Health Organization estimates that 265,000 people die of burn injuries each year, mostly (95%) from LMICs; the vast majority never present for medical care. Scalds and electrical burns represent another significant source of death and disability. Women and children in LMICs are most likely to be burned in domestic kitchens; men are more likely to be burned in the workplace. The economic and social impact from long hospitalizations and from the resulting disfigurement provides a significant negative stigma causing ostracism and rejection.

Of all the forms of trauma worldwide, burns are the only type that predominantly afflict women and children. Southeast Asia accounts for 27% of burn-related deaths worldwide; 70% of people dying from burns in this region are women. Cooking on wood, charcoal, or low kerosene stoves also puts children at risk, particularly from scalding (Fig. 49-12). Small children in the WHO African region have triple the number of burn deaths as children worldwide. Contrast this with the United States, where more burns and burn deaths affect men.

People living in rural areas suffer disproportionately because there are fewer facilities capable of managing the acute and chronic aspects of burns and because the population is generally poorer. Surgical grafting and management of contractures is often best done in specialized burn centers, but these are rare in LMICs. Telemedicine has been shown to be effective in managing burns and preventing complications, and now, in the era of high resolution mobile phones, it can effectively diagnose and triage many burn patients appropriately. Telemedicine can also be useful in providing much-needed education of rural providers in basic burn care (Box: Telemedicine and Tele-education at the University of Utah).
Telemedicine and Tele-education at the University of Utah

The University of Utah’s Burn Center has utilized Project ECHO (Extension for Community Healthcare Outcomes), an HIPAA-compliant tele-education platform developed at the University of New Mexico. Project ECHO provides live, free, interactive educational materials to rural physicians, nurses, and EMTs in eight surrounding states.69 This model has been expanded internationally through the University of New Mexico to 21 countries, with the potential to drastically improve the knowledge base of providers in many resource-limited settings.70

Direct patient care has also been achieved at the University of Utah’s Burn Center, through their Telemedicine outreach program. “TeleBurn” currently provides approximately 400 video consultations per year, serving 80 sites in five surrounding states that lack specialized burn care. For a region like the Intermountain West, where travel can be limited by inclement weather and long distances, the TeleBurn program provides better access to specialty care, at lower costs to patients.71

Figure 49-11. Change in traffic fatality risk (deaths per 10,000 persons, 1975–1998). (Reproduced with permission from Intermountain Healthcare.)

Figure 49-12. Domestic kitchen: risk factor for burns in women and children in LMICs. (Used with permission from James H. Kenney, Jr.)
Essential Surgery: Current and Evolving Concepts

Dr. Jim Yong Kim, President of the World Bank, aptly stated that surgery is an “indivisible, indispensable part of health care.” The wisdom of this statement has been supported by the findings of two landmark publications in 2015: Disease Control Priorities, third edition (DCP3), and the Lancet Commission on Global Surgery 2030 (LCGS). According to the DCP3, “the provision of essential surgical procedures would avert 1.5 million deaths a year or 6.7% of all avertable deaths in LMICs,” and according to the Lancet Commission, 5 billion people do not have access to safe, affordable surgical and anesthesia care when needed. Taken together, these and other findings suggest that without the provision of accessible, affordable essential surgical care in all LMICs, the lofty goal of another generation within a generation of Global Health 2035: A World Converging Within a Generation, would be unachievable. This earlier Lancet Commission believes that, with adequate investment in global health, all countries could reduce their infectious, maternal, and childhood mortality rates down to those currently seen in the best-performing middle-income countries (e.g., the 4C countries: Chile, China, Costa Rica, and Cuba) within a generation by 2035. They also make the interesting observation that the LMICs can use their own resources for much of the funding needed.

The critical role of essential surgical and anesthesia services in global health, in general, and in saving lives and disabilities in LMICs, in particular, has been established by fact-based evidence and analysis provided by the two landmark publications of 2015, DCP3 and LCGS. Table 49-1 summarizes the key findings and recommendations of the two publications.

The DCP3 adopted a working definition of essential surgical conditions as those that (a) are primarily or extensively treated by surgery; (b) have a large health burden; and (c) can be successfully treated by a surgical procedure that is cost-effective and feasible to promote globally. Using this definition, the DCP3 identified 44 essential procedures, most of which can be performed in first-level hospitals (Table 49-2). The first-level (district) hospital is the appropriate platform to provide essential surgical service. These procedures rank among the most cost-effective of all interventions and include those that treat injuries, obstetric complications (including fistulas), abdominal emergencies, cataracts, and congenital anomalies.

The LCGS Report, based on extensive research and analysis of factual evidence, provides recommendations to improve access to safe, affordable anesthesia and surgical care in LMICs. Essential surgical services should be integrated into a comprehensive platform of healthcare delivery. At the core of delivery of essential surgery is the first (district) hospital, which must be capable of delivering three bellwether essential surgery procedures (hysterectomy, laparotomy, and treatment of an open fracture). A hospital that can provide these three procedures safely is presumed to have the necessary expertise in general and orthopedic surgery, obstetrics, and anesthesia to perform all essential surgical procedures.

The cost of untreated surgical conditions is huge and, until now, not recognized. At the present time, some 33 million individuals face catastrophic health expenditure for surgical and anesthesia care in LMICs. The LCGS estimates that it would cost U.S. $420 billion to scale up the surgical workforce.

---

**Table 49-1**

<table>
<thead>
<tr>
<th>DCP3</th>
<th>LCGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provision of essential surgery in LMICs would prevent 1.5 million deaths, or 6.7% of all avertable deaths</td>
<td>5 billion people lack access to safe, affordable surgical and anesthesia care when needed</td>
</tr>
<tr>
<td>Essential surgical procedures rank among the most cost-effective of all health interventions</td>
<td>143 million more operations are needed in LMICs, where only 6% of all worldwide procedures are now done</td>
</tr>
<tr>
<td>Effective and affordable measures (such as task-sharing) increase access to surgical care</td>
<td>33 million face catastrophic health expenditure from surgery and anesthesia care each year</td>
</tr>
<tr>
<td>Investments must be made to expand capacity building</td>
<td>Without urgent investment, LMICs will lose US $12.3 trillion in economic productivity between 2015 and 2030</td>
</tr>
<tr>
<td>Substantial disparities exist between countries in safety of surgical and anesthetic care. Feasible and affordable measures (e.g., surgical safety checklist) improve safety and quality</td>
<td>Surgery is an indivisible, indispensable part of health care. Surgical care should be part of the National Health Care System, and should be “available, accessible, safe, timely, and affordable.”</td>
</tr>
<tr>
<td>Universal coverage of essential surgery should be publicly financed early on the path to universal health coverage</td>
<td></td>
</tr>
</tbody>
</table>

**LMIC = Low and Middle Income Country**

---

**Table 49-2**

**Essential surgical procedures that can be performed in first level (District) hospitals (DCP3)**

**Obstetric Complications**
- Severe postpartum hemorrhage, obstructed labor, prolonged labor, eclampsia, prolapsed cord, fetal distress, tubal pregnancy, postabortion endometritis/myometritis, postabortion sepsis, intrauterine fetal death

**Trauma and Violence**
- Major limb fracture/injury, joint dislocation, major soft tissue injury, pneumo/hemothorax, ruptured spleen

**Acute Surgical Emergencies**
- Strangulated hernia, intestinal obstruction, intestinal perforation, appendicitis, liver abscess, major wound infection, osteomyelitis/septic arthritis

**Nonacute Surgical Conditions**
- Congenital hernia, hernia, breast cancer, chronic osteomyelitis, hydrocele, urethral stricture, prostatic hypertrophy, cataract, eye injury
to have 20 surgical, anesthetic, and obstetric providers (SAOPs) per 100,000 population in LMICs by 2030. This figure must be compared to the U.S. $20.7 trillion loss in global economy that surgical conditions would be responsible for.

Recent studies have shown that essential surgical conditions account for about 18% of the global burden of disease. Investment in essential surgical services is critical and should be done early in the path towards universal health coverage (UHC). The barriers to essential surgical services in LMICs are formidable. The shortage in surgical workforce is huge, and it is clear that the deficit cannot be satisfactorily addressed without task sharing. Infrastructure deficits (clinics, hospitals, equipment, drugs, blood banks, etc) are equally enormous. The first (district) hospital is the important platform for delivery of essential surgical services. The DCP3 estimates that it would cost U.S. $43 million annually of additional spending to provide universal coverage of essential surgery applicable to first-level hospitals worldwide.

**Outreach and Engagement**

Many models for outreach and engagement have had a positive impact on the accessibility of surgery. Organizations participating in outreach are guided by a wide range of motivations and resources (Fig. 49-13). Some organizations are purely humanitarian and service oriented; others are primarily educational. Some even use the promise of healthcare to advance political, religious, or personal agendas.

Many patients have benefited from the multitude of service-oriented volunteer “missions” providing much needed surgical care that would otherwise have been unavailable. While volunteerism and medical missions provide needed clinical surgical care for underserved populations, they may not be a sustainable solution to long-term manpower shortages for health. Comprehensive initiatives are necessary to engage local healthcare professionals and organizations, governments, and academic institutions to build sustainable capacity.

**Charitable Surgical Delivery Platforms.** A significant burden of surgical disease is addressed through charitable organizations. The DCP3 divides these charitable surgical delivery platforms into two types: temporary delivery platforms and specialty surgical hospitals (Table 49-3).77

Short-term charitable surgical platforms bring entire surgical teams along with equipment and supplies needed to operate in local facilities for a short period of time. Local physicians provide the majority of follow-up care.78-80 Self-contained platforms bring the entire surgical infrastructure (fully functional operating rooms, postoperative recovery capability) through various modes of transportation: airplanes, ships, trucks, and buses. These self-contained platforms tend to stay in-country longer, may still provide short-term care, and tend not to leave behind any physical structure.78,80,81

Little information exists on outcomes and cost-effectiveness of these temporary surgical platforms. Where no other services

---

**Figure 49-13.** Global surgery initiatives. (Reproduced with permission from Intermountain Healthcare.)
exist, they may provide needed services. However, some of these charitable organizations report higher complications rates in lower-resource settings, which seem to increase even more with complex procedures. Some of these platforms include education for local care providers along with clinical care. While some question their ability to sustainably train local surgical teams, one charitable partnership with short-term, concentrated surgical training trips over 9 years documented a countrywide transition from open cholecystectomy to laparoscopic cholecystectomy in Mongolia.82

Specialty surgical hospitals establish entire hospitals or facilities within existing hospitals. Some target specific diseases (Addis Ababa Fistula Hospital) while others provide a wide range of surgical and medical services (Pan-African Academy of Christian Surgeons [PAACS] mission hospitals); many are supported through partnerships with various charitable and governmental organizations.77

International Organizations

United Nations. Committed to maintaining international peace, developing friendly relations between nations, and promoting better standards of living (conquering hunger, disease, and illiteracy) and human rights, representatives from 51 nations in 1945 signed the United Nations (UN) Charter at the United Nations Conference on International Organization in San Francisco, California.83 There are now 193 member states.84 The UN promotes a social justice agenda advocating for worldwide health, engagement of philanthropies, and civil society in global health initiatives, and it supports the sustainable development goals (SDGs).85

Sustainable Development Goals. In September 2000, the UN led a worldwide, organized effort to set benchmarks for social, economic, and environmental development. Leaders from 189 countries agreed on eight specific “millennium development goals” (MDGs), spanning poverty, mortality, education, sustainability, and development.86

The MDGs created a framework for improvement that some criticized as unattainable; nevertheless nearly 1 billion people were lifted out of extreme poverty, and primary education for girls made measurable improvements.87 Still, many challenges, including some related to lack of surgical care, remained. In 2015, the UN General Assembly reconvened to raise the bar yet again, in what was declared a “supremely ambitious and transformative vision.”88 Eight MDGs became 17 SDGs with 169 specific targets, to be achieved by 2030 (Table 49-4).85 Most relevant to the global surgeon is SDG #3, “good health and well-being,” which builds upon the MDGs’ primarily maternal- and child-mortality focus, as well as communicable disease prevention. SDG #3 broadens the focus to nine health targets, including a one-third reduction in deaths by noncommunicable diseases, as well as halving the rate of deaths and injuries from road traffic accidents by 2020. In addition, a novel push to strengthen and retain the global health workforce and systems for protection and prevention of disease also falls squarely within the realm of the surgical provider. Finally, the SDGs have garnered praise for closely involving local stakeholders, versus the expert consensus that produced the MDGs. Funding to work towards achievement of the SDGs is also divided between wealthier and poorer nations, whereas the MDGs relied primarily on funding from HICs to support their mission.89

World Health Organization. The initial UN Conference in 1945 voted to establish a new international health organization. The Constitution of the World Health Organization (WHO) was approved and ratified in 1948.85 The first World Assembly in 1948 established malaria, tuberculosis, venereal diseases, maternal and child health, sanitary engineering, and nutrition as WHO priorities. One of the WHO’s greatest public health stories is the worldwide eradication of smallpox that began with the USSR proposal for the WHO-led program in 1958 culminating in the last identified case in Somalia in 1977.

While the disease burden from communicable diseases has abated in large part from these successful international cooperative interventions, little has been done to address the growing global burden of surgical disease. Despite the laudable aims of the 1978 Declaration of Alma Ata, which expressed the need for urgent action for the world community to protect and promote health for all people, the declaration did so by crowning primary health care as the key to achieving the goal of health for all—which was then accepted by the member countries in the World Health Organization.90 Although the Alma Ata slogan

<table>
<thead>
<tr>
<th>Table 49-3</th>
<th>Examples of charitable surgical delivery platforms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Temporary Delivery</strong></td>
<td></td>
</tr>
</tbody>
</table>
| **Short-Term Trips** | World Surgical Foundation  
Kenya Orthopedic Program  
APRIDE Medical Outreach Group |
| **Self-Contained Mobile Surgical Platforms** | Mercy Ships  
Cinterandes Foundation |
| **Specialty Surgical Hospitals** | Addis Ababa Fistula Hospital  
Aravind Eye Hospital  
Mission Hospitals (PAACS’*) |

*Pan-African Association of Christian hospitals

<table>
<thead>
<tr>
<th>Table 49-4</th>
<th>Sustainable development goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No poverty</td>
</tr>
<tr>
<td>2</td>
<td>Zero hunger</td>
</tr>
<tr>
<td>3</td>
<td>Good health and well-being</td>
</tr>
<tr>
<td>4</td>
<td>Quality education</td>
</tr>
<tr>
<td>5</td>
<td>Gender equality</td>
</tr>
<tr>
<td>6</td>
<td>Clean water and sanitation</td>
</tr>
<tr>
<td>7</td>
<td>Affordable and clean energy</td>
</tr>
<tr>
<td>8</td>
<td>Decent work and economic growth</td>
</tr>
<tr>
<td>9</td>
<td>Industry, innovation, and infrastructure</td>
</tr>
<tr>
<td>10</td>
<td>Reduced inequalities</td>
</tr>
<tr>
<td>11</td>
<td>Sustainable cities and communities</td>
</tr>
<tr>
<td>12</td>
<td>Responsible consumption and production</td>
</tr>
<tr>
<td>13</td>
<td>Climate action</td>
</tr>
<tr>
<td>14</td>
<td>Life below water</td>
</tr>
<tr>
<td>15</td>
<td>Life on land</td>
</tr>
<tr>
<td>16</td>
<td>Peace, justice, and strong institutions</td>
</tr>
<tr>
<td>17</td>
<td>Partnerships for the goals</td>
</tr>
</tbody>
</table>
“health for all by 2000” did not materialize, it did galvanize efforts for global partnerships for healthcare improvements and poverty reduction. In 2015, the World Health Assembly (WHA) published resolution WHA 68.15, which proclaimed surgical and anesthesia care as a crucial component of primary care worldwide—for the first time in history. The resolution urged member states to complete nine actions, including prioritizing a core set of emergency and essential surgery and anesthesia services at the primary care level, ensuring access to essential medications and infection control techniques, and developing policies for providers’ minimum skills, among others. Additionally, the Director-General of the WHO was asked to complete ten actions related primarily to policy- and advocacy-related endeavors at the international level. The resolution was voted in unanimously by 194 member states.91

The Violence and Injury Prevention Program (VIP) and the Global Initiative for Emergency and Essentials Surgical Care (GIEESC) are two programs related to surgery within the WHO that began before 2008. But as a response to a growing recognition of the significant unmet surgical need, in 2008 the WHO for the first time included basic surgery as a component of primary health care (Fig. 49-14).92

**The Global Initiative for Emergency and Essential Surgical Care.** The Clinical Procedures (CPR) team in the WHO Department of Essential Health Technologies (EHT) convened a multidisciplinary group of experts from various surgical disciplines, professionals, and civic leaders from national and international organizations, as well as representatives from various WHO departments, in December 2005 in Geneva, Switzerland to formally organize the Global Initiative for Emergency and Essential Surgical Care (GIEESC).93 GIEESC’s main aim was to assist member states with capacity strengthening in the safe and appropriate use of emergency and essential surgical care (procedures, equipment) at resource-limited healthcare facilities through training and education programs. The training program was built around the WHO Integrated Management of Emergency and Essential Surgical Care (IMEESC) tool kit.94 The tool kit included best practice protocols, guidelines on policies, training curriculum, emergency equipment, teaching slides, and monitoring and evaluation instructions. Additionally, low-cost editions of the manual Surgical Care at the District Hospital have been made available in local languages. As of 2015, GIEESC had over 2100 members in 140 countries.95 A Mongolian edition facilitated early expansion of GIEESC throughout the country. Mongolia has improved basic infrastructure, human resources, and capabilities; and the use of the tool kit system has led to its incorporation into the countrywide healthcare plan (Box: Mongolia GIEESC).

**Mongolia GIEESC**

The WHO situational analysis tool, developed in 2007 to assess the availability of emergency and essential surgical care (EESC) at individual health facilities, has been utilized to document limited infrastructure, human resources, procedures, equipment, and supplies available for even basic EESC in many countries.96 For example, there were no trained surgeons or anesthetists at 44 first-referral hospitals in Mongolia.1 Only 66% of the facilities had electricity, and 45% had running water (Fig. 49-15).

Most facilities lacked any policy for EESC, disaster preparedness, basic equipment to provide EESC, or any

---

**Figure 49-14.** Emergency and essential surgery: an integral component of primary care. (Reproduced with permission from The World Health Report 2008—primary Health Care (Now More Than Ever). http://www.who.int/whr/2008/en/)
access to training for EESC. Adopting a health systems strengthening approach to rectify these glaring deficiencies, Mongolia implemented a nationwide EESC program involving 14 of the 21 provinces (Aimags) from 2004 to 2010 (Fig. 49-16). In 6 years, dramatic improvements in short-term process measures were identified using the WHO Monitoring and Process form: 57.1% increase in availability of emergency rooms; 59.1% increase in the supply of emergency tool kits; and a 73.6% increase in the recording of emergency cases (Figs. 49-17 and 49-18). More importantly, countrywide morbidity and mortality dropped significantly (Fig. 49-19).

Violence and Injury Prevention. The Violence and Injury Prevention (VIP) program promotes numerous activities to assist countries to prevent and mitigate the consequences of violence and injury. While injury prevention is paramount, VIP provides guidance for strengthening trauma systems in countries of all economic levels to improve emergency care and rehabilitation. VIP encourages development of systematic data collection and analysis to better guide appropriate interventions. Prevention programs include the WHO Helmet initiative, while the Essential Trauma Care Project (EsTC) creates standards for the care of injured patients and promotes systematic capacity building. VIP advocates support for the UN Decade of Action for Road Safety 2011–2020 and initiatives to achieve SDG 3.6,
halving the number of global deaths and injuries from road traffic accidents by 2020, and SDG 11.2, which aims to provide access to safe, affordable, accessible, and sustainable transport systems for all by 2030.99

WHO Safe Surgery Saves Lives Initiative. Surgeons have always sought ways to prevent perioperative complications. Aseptic technique, one of the greatest forms of prevention in surgical care, requires vigilant reinforcement to prevent serious wound infections. In resource-limited areas inadequate perioperative monitoring, lack of critical medications, and poor documentation place patients at increased risk for serious complications. The WHO Safe Surgery Saves Lives Initiative is a worldwide attempt to prevent perioperative complications.100 The initiative identified 10 basic and essential objectives that can help prevent perioperative injuries (Table 49-5).101 A simple, three-stage checklist (initiated as the patient enters the operating room, just before the procedure, and just prior to the patient leaving the room) implemented in eight high-, middle-, and low-income countries found a 50% reduction in the failure to meet basic safety standards resulting in a 50% decrease in mortality (Fig. 49-20).102

Global Surgery and Public Health

Surgical care is increasingly recognized as an integral component of public health. Traditional teaching portrays surgery as the antithesis of public health: treating the individual instead of the community, reactionary instead of preventive, and too expensive especially for countries with developing economies. Yet in reality, surgery and public health share many priorities and would benefit from greater integration in many areas (Fig. 49-21). For example, providing access to obstetrical care or birth attendants for every delivery could

Figure 49-17. Surgical procedures performed 1 to 2 years post training (13 Soum hospitals evaluated). (Reproduced with permission from Henry JA, Orgoi S, Govind S, et al: Strengthening surgical services at the soum (first-referral) hospital: the WHO emergency and essential surgical care (EESC) program in Mongolia. World J Surg. 2012 Oct;36(10):2359-2370.)
prevent the majority of vesicovaginal fistulas and markedly decrease the most common cause of maternal death—hemorrhage—for entire communities. Ninety percent of mortality from injury occurs in LMICs, providing another area for surgical teams to lead preventative, population-based strategies to improve public health. Male circumcision is another example of a well-documented preventative, minor surgical procedure, capable of reducing the transmission of HIV.


Figure 49-19. Surgical morbidity and mortality: Mongolia 2001–2009. (Reproduced with permission from Intermountain Healthcare.)
devotes an entire volume to essential surgery, emphasizing its importance as a key part of health worldwide. There are three significant developments helping to accelerate the integration of surgery and public health:

1. Improved understanding of the burden of surgical disease and its significant component of the overall burden of global disease
2. Recognition that surgery has a primary, secondary, and tertiary preventative role (Table 49-6)
3. Documentation that surgical care can be cost-effective for community-based healthcare

Even after Learmonth presented his landmark lecture in 1949 “The Contributions of Surgery to Preventive Medicine” at the University of London’s Heath Clark Lecture series, surgery has been neglected as a component of public health.  

<table>
<thead>
<tr>
<th>Table 49-5</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Ten basic and essential objectives for safe surgery (WHO*):</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Operate on the correct patient at the correct site</td>
</tr>
<tr>
<td>2. Use method known to prevent harm from anesthetic administration, while protecting the patient from pain</td>
</tr>
<tr>
<td>3. Recognize and effectively prepare for life-threatening loss of airway or respiratory function</td>
</tr>
<tr>
<td>4. Recognize and effectively prepare for risk of high blood loss</td>
</tr>
<tr>
<td>5. Avoid inducing any allergic or adverse drug reaction known to be a significant risk for the patient</td>
</tr>
<tr>
<td>6. Consistently use method known to minimize risk of surgical site infection</td>
</tr>
<tr>
<td>7. Prevent inadvertent retention of instruments or sponges in surgical wounds</td>
</tr>
<tr>
<td>8. Secure and accurately identify all surgical specimens</td>
</tr>
<tr>
<td>9. Effectively communicate and exchange critical patient information for the safe conduct of the operation</td>
</tr>
<tr>
<td>10. Establish routine surveillance of surgical capacity, volume, and results</td>
</tr>
</tbody>
</table>


Surgical safety checklist

| Figure 49-20. Surgical safety checklist. (Reproduced with permission from WHO surgical safety checklist, 2009, http://whqlibdoc.who.int/publications/2009/9789241598590_eng_Checklist.pdf © World Health Organization 2009 All rights reserved.) |
|---|---|---|
| Before Induction of anaesthesia | Before Skin Incision | Before Patient Leaves Operating Room |
| (with at least nurse and anaesthetist) | (with nurse, anaesthetist and surgeon) | (with nurse, anaesthetist and surgeon) |

- Has the patient confirmed his/her identity, site, procedure, and consent? [ ] Yes [ ] Not applicable |
- Is the site marked? [ ] Yes [ ] Not applicable |
- Is the anaesthesia machine and medication check complete? [ ] Yes |
- Is the pulse oximeter on the patient and functioning? [ ] Yes |
- Does the patient have a: Known allergy? [ ] No [ ] Yes |
- Difficult airway or aspiration risk? [ ] No [ ] Yes, and equipment/assistance available |
| Risk of >500 ml blood loss (7 ml/kg in children)? [ ] No [ ] Yes, and two IVs/central access and fluids planned |
- Confirms all team members have introduced themselves by name and role. |
- Confirms the patient’s name, procedure, and where the incision will be made. |
- Has antibiotic prophylaxis been given within the last 60 minutes? [ ] Yes [ ] Not applicable |
- Anticipated critical events to Surgeon: |
  - What are the critical or non-routine steps? |
  - How long will the case take? |
  - What is the anticipated blood loss? |
- Anticipated critical events to Anaesthetist: |
  - Are there any patient-specific concerns? |
- Anticipated critical events to Nursing team: |
  - Has sterility (including indicator results) been confirmed? |
  - Are there equipment issues or any concerns? |
- Is essential imaging displayed? [ ] Yes [ ] Not applicable |

Nurse verbally confirms: |
- The name of the procedure |
- Completion of instrument, sponge and needle counts |
- Specimen labeling (read specimen labels aloud, including patient name) |
- Whether there are any equipment problems to be addressed |

To surgeon, anaesthetist and nurse: |
- What are the key concerns for recovery and management of this patient? |

This checklist is not intended to be comprehensive. Additions and modifications to fit local practice are encouraged.

Based on the WHO Surgical Safety Checklist http://whqlibdoc.who.int/publications/2009/9789241598590_eng_Checklist.pdf © World Health Organization 2009 All rights reserved.
Strategies for Integration of Global Surgery and Public Health. Three areas stand out as opportunities for integration of global surgery and public health: education, professional societies, and multinational health policy organizations. From an education standpoint, several universities in HICs have developed formal programs for the study of surgery and public health.108,109 Diseases commonly present in very late stages in LMICs and in disadvantaged populations in developed countries. Many morbid conditions could have been cured while localized in their earlier stages and likely eradicated by a local surgical procedure. Early recognition and treatment of surgically correctable diseases is a critical preventive role for surgery. Many surgical procedures are not only a form of tertiary prevention, but are also forms of primary prevention (Table 49-7).110

Assigning Disease Priorities. Global surgery interventions can be prioritized to identify those conditions in which clinicians and public health professionals should collaborate most closely—targeting those diseases that impose the largest burden on a society and have a highly successful surgical outcome (Table 49-8).111 There are four broad, high-priority areas where surgery has an important role for public health interventions:

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Target</th>
<th>Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Primary</td>
<td>Root causes of disease</td>
<td>Eliminate or reduce risk of developing illness</td>
</tr>
<tr>
<td>2. Secondary</td>
<td>Illness or disease at earliest stages</td>
<td>Limit progression of disease</td>
</tr>
<tr>
<td>3. Tertiary</td>
<td>Disease at later stages</td>
<td>Cure or limit the effect of existing disease</td>
</tr>
</tbody>
</table>


The role of surgery for primary prevention of cancer (Table 49-6)

<table>
<thead>
<tr>
<th>Tertiary Surgical Procedure</th>
<th>Primary Cancer Prevented</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast lumpectomy for ductal carcinoma in situ</td>
<td>Breast</td>
</tr>
<tr>
<td>Colonoscopic polypectomy</td>
<td>Colon</td>
</tr>
<tr>
<td>Colposcopy and excision</td>
<td>Cervical</td>
</tr>
<tr>
<td>Resection of actinic keratosis</td>
<td>Skin</td>
</tr>
<tr>
<td>Resection of leukoplakia and erythroplakia</td>
<td>Oral</td>
</tr>
</tbody>
</table>


Prioritization of surgical conditions (Table 49-8)

<table>
<thead>
<tr>
<th>Priority*</th>
<th>Public Health Burden</th>
<th>Surgical Procedure Successful</th>
<th>Cost-effective and Feasible to Promote Globally</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>High</td>
<td>Highly</td>
<td>Highly</td>
</tr>
<tr>
<td>2</td>
<td>Moderate</td>
<td>Moderately</td>
<td>Moderately</td>
</tr>
<tr>
<td>3</td>
<td>Low</td>
<td>Neither highly nor moderately</td>
<td>Low</td>
</tr>
</tbody>
</table>

*Priority one implies that all three conditions must be met. The priority should be shifted to 2 or 3 if any of the conditions are moderate or low. Data from Mock C, Cherian M, Juillard C, et al: Developing priorities for addressing surgical conditions globally; furthering the link between surgery and public health policy, World J Surg. 2010 Mar;34(3):381-385.

Trauma Care. The Essential Trauma Care Project (EsTC) begun in 2001 is a collaboration effort between the International Association for Trauma Surgery and Intensive Care, an integrated society within the International Society of Surgery-Societe-Internationale Chirurgie (ISS-SIC) and the World Health Organization (WHO), specifically the Violence and Injury Prevention unit. The project culminated in a document that identified 11 core essential trauma care services (“the rights of the injured patient”) that ought to be available at all levels of healthcare facilities (Table 49-9).112 In addition, the document delineated 260 human and physical resources that should be available based on the type of facility (Table 49-11).

The role of surgery for public health strategies (Table 49-9)

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Prevention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trauma care</td>
<td>Prevention of death and chronic disability by the provision of timely, expert, and complete surgical care</td>
</tr>
<tr>
<td>Obstetrical emergencies</td>
<td>Timely surgical intervention in obstructed labor, in pre- and post-partum hemorrhage, and other obstetrical complications</td>
</tr>
<tr>
<td>Acute surgical emergencies</td>
<td>Provision of competent surgery to treat a wide range of emergency abdominal and nonabdominal conditions</td>
</tr>
<tr>
<td>Nonacute surgical conditions</td>
<td>Surgical care for several elective conditions that have a significant effect on the quality of life such as cataract, otitis media, clubfoot, and hernias</td>
</tr>
</tbody>
</table>

The EsTC recommendations provide a cost-effective framework for LMICs to improve their trauma care. These recommendations have been used as a planning guide and as an advocacy statement. To catalyze strengthening trauma and emergency care in low- and middle-income countries, in 2007, the WHA adopted a resolution on emergency care systems (resolution WHA 60.22).113,114 This first-ever WHA resolution dedicated specifically to trauma care highlights the importance accorded by world governments in caring for their injured.

Quality improvement programs provide inexpensive tools to strengthen trauma systems. National trauma registries, integral for trauma research, can be used to monitor and improve patient outcomes.115,116 Yet very few trauma registries exist in LMICs.117,118 The World Bank stated, “It is critical for LMICs to create or strengthen existing trauma systems to improve outcomes.”119,120 Trauma systems exist in varying states of development in different countries, and even within countries. Initiatives to strengthen trauma systems target the full spectrum of services: prevention, prehospital and definitive hospital care, rehabilitation, and process improvement and patient safety initiatives. Nearly 2 million lives could be saved each year if LMICs could design and implement simple trauma care initiatives that reduced the case fatality rates among seriously injured patients to equal those in HICs (Fig. 49-22).119,120

In one Canadian province, introducing simple prehospital

---

**Table 49-10**

<table>
<thead>
<tr>
<th>Essential trauma care services</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Obstructed airway appropriately maintained</td>
</tr>
<tr>
<td>2. Impaired breathing supported</td>
</tr>
<tr>
<td>3. Pneumothorax and hemothorax promptly diagnosed and treated</td>
</tr>
<tr>
<td>4. Bleeding promptly stopped (internal or external)</td>
</tr>
<tr>
<td>5. Shock recognized and treated appropriately (I.V. fluids)</td>
</tr>
<tr>
<td>6. Timely decompression of space occupying lesions to prevent secondary brain injury</td>
</tr>
<tr>
<td>7. Abdominal injuries diagnosed and promptly repaired (intestinal injuries and others)</td>
</tr>
<tr>
<td>8. Disabling extremity injuries corrected</td>
</tr>
<tr>
<td>9. Potentially unstable spine injuries identified and managed (early immobilization)</td>
</tr>
<tr>
<td>10. Minimize consequences of injuries by appropriate rehabilitative services</td>
</tr>
<tr>
<td>11. Medication to provide above services and relieve pain readily available</td>
</tr>
</tbody>
</table>


---

**Table 49-11**

<table>
<thead>
<tr>
<th>Knowledge and skills</th>
<th>Facility level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment of airway compromise</td>
<td>Basic</td>
</tr>
<tr>
<td>Manual maneuvers (chin lift, jaw thrust)</td>
<td>E</td>
</tr>
<tr>
<td>Insertion of oral or nasal airway</td>
<td>D</td>
</tr>
<tr>
<td>Endotracheal Intubation</td>
<td>D</td>
</tr>
</tbody>
</table>

**Equipment and supplies**

| Oral or nasal airway | D | E | E | E |
| Laryngoscope | D | D | E | E |
| Endotracheal tube | D | D | E | E |
| Capnography | I | D | D | D |

E: essential; D: desirable; I: irrelevant (not usually to be considered as the level in question).

*EsTC: Essential Trauma Care.

interventions and improving the focus on trauma at the policy level demonstrated consistent improvement in trauma-related morbidity and mortality, over a relatively short time period (Fig. 49-23).121

**Obstetrical and Other Acute Surgical Emergencies.** Reduction of maternal deaths and long-term disability are high priorities for the international community.122 Despite the 44% reduction in maternal deaths from 1990 to 2015, approximately 830 women—mostly in LMICs—still die daily from preventable causes related to pregnancy and childbirth. For every maternal death, 30 women are incapacitated by chronic problems that reduce their quality of life and ability to care for their families. High priority surgical procedures to improve maternal health include cesarean section, hysterectomy for postpartum bleeding and uterine rupture, management of ectopic pregnancy, and dilatation and curettage.111 In 2015, the LCGS reported that maternal mortality was closely related to density of surgical, anesthetic, and obstetrical providers (SAOPs). They showed that maternal mortality throughout the world appeared to decrease—by 13.1% on average—for every 10 unit increase in SAOPs per 100,000 persons, a strong argument for addressing maldistribution of providers related to surgical disease.16

In terms of nonobstetrical acute surgical emergencies, about 90% could be addressed by developing the capability to care for the 10 most common acute surgical conditions in any local region. While a few types of disease processes vary by geographical location, there are many that are universal, including appendicitis, strangulated hernia, small bowel obstruction, perforated peptic ulcer, fractures, lacerations, and wounds.

**Nonacute Surgical Conditions.** Even common nonacute conditions can have significant impact on the quality of life. Hernias can prevent otherwise healthy individuals from working, especially in societies where the economy relies heavily on manual labor. Cleft lip and cleft palate deformities interfere with the ability to speak or eat properly and predispose affected individuals to chronic ear infections leading to hearing loss. Many live in isolation because social ostracism prevents them from attending school, marrying, or holding jobs.124 Plastic surgeons who pioneered global outreach for reconstructive procedures for cleft lip and palate opened the door for subsequent outreach by other specialties, including ophthalmology, orthopedics, general surgery, urology, and dentistry.125-127

The most common form of blindness is caused by cataracts. Cataracts decrease the quality of life and the socioeconomic status for both the blind person and his or her family. The fact that 90% of blind people no longer work reveals the extra burdens carried by the family members who care for them.128 The Himalayan Cataract Project (HCP) is a highly successful initiative focusing on cataracts in Asia and Africa. HCP priorities and measurable outcomes illustrate how combining key public health concepts with a comprehensive approach to surgical care creates a model for curing disease, building economies, and delivering hope in resource-poor areas9 (Box: The Himalayan Cataract Project: A Sustainable Public Health Approach for Curing Blindness).
The Himalayan Cataract Project (HCP): A Sustainable Public Health Approach for Curing Blindness

According to the WHO criteria, 285 million people worldwide are visually disabled. Of that population, 39 million are classified as bilaterally blind; 90% live in the developing world where poor water quality, lack of sanitation, malnutrition, and inadequate services cause a higher incidence of eye disease. The most common cause of avoidable blindness in LMICs is cataract (43%). Nepal has one of the highest incidences of cataracts partially due to increased exposure to ultraviolet sunlight encountered at its higher elevations; 62% of total blindness in Nepal is due to cataracts.

In 1995, Sanduk Ruit joined forces with Geoffrey Tabin to establish the Himalayan Cataract Project (HCP). In the early 1990s, difficult geography with inadequate transportation, the high cost of intraocular lenses, and a lack of trained ophthalmologists, assistants, and nurses limited access to cataract surgery for the poor.

HCP developed and defined six priorities, each with an associated public health principle and outcome measurement that provided the basis for assessing success and for implementing change (Fig. 49-24). HCP’s care model targeted the entire population of blind people with cataracts regardless of the ability to pay. Since most of the potential patients lived in remote areas, HCP found it imperative to take cataract surgery to the local communities. The Tilganga Institute of Ophthalmology (TIO) in Katmandu, Nepal, has served as a base from which 493 doctors and over 19,000 ophthalmic personnel of all levels have received training since 1994. Through the TIO and its outreach programs, over 4,657,748 people have been screened, and more than 307,611 eye surgeries have been performed since 1994 (Fig. 49-25).

The TIO developed an ophthalmology residency training program implementing standards set forth by the American Academy of Ophthalmology. In addition to the formal residency program for ophthalmologists, HCP established training programs for community eye care workers in a three-year Ophthalmic Assistant Training Program.

Ruit developed an innovative sutureless technique for cataract surgery yielding equivalent results to those in developed countries but also reproducible in resource-constrained areas. By redesigning the intraocular lens implant and mass producing it locally in Nepal for U.S. $4, Ruit and Tabin provided a low-cost alternative to the higher-priced lens produced in developed countries. A local business—the Fred Hollows Intraocular Lens Factory—mass produces the lenses and supports the local economy by creating a new sustainable business.

HCP also designed a compassion-driven, culturally acceptable method for cost-recovery that involves a sliding scale for payment: 45% of patients pay U.S. $120; 20% pay a smaller amount based on their economic situation; and 35% receive cataract surgery for free.

With the rapidity and scale of success experienced in Nepal, HCP and TIO began expanding their efforts globally. HCP is now actively working to replicate and proliferate their model in countries throughout South Asia and Africa by developing high-quality eye care systems, supporting local institutions, and training local doctors and ophthalmic personnel. Since 2005, HCP has trained over 300 ophthalmic personnel from 19 countries.

---

**Figure 49-24.** Himalayan cataract project priorities, public health principles, and outcome measurements. (Redrawn from Himalayan Cataract Project and Tilganga Eye Center, Cureblindness.org, 129-131, by permission. Illustration reproduced with permission from Intermountain Healthcare.)
Cancer Initiatives
Surgery for cancer in public health plays a role not only for curative surgery, but also for early diagnosis, prevention, and palliation. Solid tumors, in their early stages, present insidiously as a nonacute surgical problem. Due to cancer’s recent recognition as a leading cause of death, cancer has been identified as a health priority in LMICs. Most solid tumors are incurable without surgery and at a minimum require surgical excision of the primary lesion.

It is often not appreciated that surgeons provide a significant amount of primary care and are the principle providers involved in endoscopic screening and treatment of gastrointestinal tumors in LMICs. In countries without specialized services, low-cost and effective treatment options combining early prevention and treatment with off-patent drug use have led to coverage of cancer treatment in several middle-income countries’ national health insurance plans. Cancer care provides significant opportunity for including surgery in community-wide public health programs as a high priority according to the prioritization model (see Table 49-6); cancer has a high public health burden, is treated with highly successful procedures, and can be cost-effective and feasible globally. In 2009, a coalition of leaders in cancer care and public health organized the Global Task Force on Expanded Access to Cancer Care and Control in Developing Countries (GTFCCC). GTFCCC’s mission is to expand access to cancer prevention, detection, and care in LMICs. Successful partnerships have already been entered into Haiti, Rwanda, Mexico, Malawi, and Jordan.

Cost-Effectiveness of Surgical Care. Funders in healthcare look for measurable return on their investments. While comparison of outcomes and objective measures would be ideal, reality demonstrates that healthcare budgets more commonly are dictated by politics rather than actual need. Nevertheless, in a world of limited resources and tightening budgets for healthcare, cost-effective analyses of various options for intervention are critical for policy makers. Comparing various options that have different outcomes is an approach called cost-utility analysis (CUA).

Surgical interventions can be evaluated by specific diseases or conditions, or by systems or services required to support the delivery of surgical care. In 1990, the World Bank defined the Disability Adjusted Life Year (DALY) as the sum of Years of Life Lost (YLL) due to premature mortality in the population and the Years Lost due to Disability (YLD) for people living with the health condition or its consequences (DALY = YLL + YLD). Evaluating the cost per DALY averted is one approach for comparing the cost-utility between medical and surgical interventions. Recent surgical cost/DALY studies identifying the cost-effectiveness of various types of surgical care have allowed surgical initiatives to be considered when prioritizing public health initiatives.

The World Bank arbitrarily defined U.S. $100 per DALY averted per day in low-income countries as highly cost-effective. Compared to other public health initiatives, developing basic and emergency surgical care at the district level hospital is as cost-effective as, or more so, than typical public health programs such as retroviral treatments for HIV/AIDS or immunization for measles (Fig. 49-26). Using the WHO’s cost-effectiveness standards, investing in emergency obstetrical systems, including timely cesarean delivery, can also be considered “highly cost-effective” for 48 of 49 countries in which there are currently inadequate numbers of cesarean deliveries. The median cost per DALY averted by cesarean-section was $304. In addition, the cost-benefit ratio in 46 of 49 countries was >1, suggesting that investment in cesarean delivery is a viable economic proposition.

Inguinal hernia repair is one of the most common operations performed worldwide. Tension-free inguinal hernia repairs performed with mosquito netting or polypropylene mesh were cost-effective in Western Ecuador and Western Ghana ($78.18 per DALY and $12.88 per DALY averted, respectively). Using mosquito netting in India was 3700 times cheaper than using traditional polypropylene mesh.

Using “value of lost output” (VLO) data representing 90% of the world’s population, it is estimated that U.S. $20.7 trillion would be lost between 2015 and 2030 due to unmet surgical needs and their inherent morbidity and mortality. Projected...
economic losses of such magnitude have underlined the importance of prioritizing surgical infrastructure and the cost-effective nature of many interventions within the surgical realm.

**Factors Affecting Utilization and Outcome for Surgical Care.** There are three major factors that severely limit utilization of surgical services:

1. Socioeconomic and cultural factors
2. Accessibility of facilities
3. Quality of care (Fig. 49-27)

The decision to seek timely care is affected by the costs associated with time off from work and inability to support the family during the absence, transportation and lodging, and the surgical services themselves. Cultural and religious traditions may define acceptability of various treatment options. For example, many people in Mongolia refuse to have surgery on Tuesdays as this is viewed as a “bad luck” day. Understanding local customs and cultural concerns can improve utilization of surgical services.

At the intersection of cost and culture are “willingness to pay” (WTP) models, which predict how a society’s perceived costs of obtaining care versus tolerating a medical condition will lead to or prevent them from seeking care. Such calculations can inform which policies are most likely to yield improved health for a country or region, and they rely heavily on per capita gross domestic product (GDP) and DALYs averted. It is vital to understand that these models, and the policies they inform, are context-dependent. What is perceived as socially valuable in Tanzania may be seen as overpriced or unnecessary in Haiti. As global surgical advocates work with public health experts to strengthen surgical systems, it will be important to remember that context, culture, and cost are indivisible from one another.

Austere environments, difficult terrain, and long distances from health care facilities significantly delay or prevent access to surgical care. Triage and transfer guidelines along with telemedicine have the potential to mitigate the limitations of geography. However, without adequately trained care providers and support staff, the risk for poor outcomes is increased.

Recognizing these three important factors for increasing utilization and outcomes, Mongolia initiated a public health approach for the management of gallbladder disease incorporating minimally invasive surgery (Box: The Public Health Approach to Management of Gallbladder Disease in Mongolia).

---

**Figure 49-26.** Cost-effectiveness of surgical interventions, compared to two key medical interventions. Note: DALY = disability-adjusted life year. (Reproduced with permission from Intermountain Healthcare.)

<table>
<thead>
<tr>
<th>Orthopedic surgery trip</th>
<th>Trauma center</th>
<th>Cesarean delivery</th>
<th>Hydrocephalus repair</th>
<th>Trachoma surgery</th>
<th>Cleft lip and palate repair</th>
<th>Hernia repair</th>
<th>Surgical hospital</th>
<th>Cataract surgery</th>
<th>Obstetric hospital</th>
<th>Measles vaccination</th>
<th>Antiviral therapy for HIV</th>
</tr>
</thead>
<tbody>
<tr>
<td>$100/DALY</td>
<td>$100/DALY</td>
<td>$100/DALY</td>
<td>$1000/DALY</td>
<td>$1000/DALY</td>
<td>$1000/DALY</td>
<td>$1000/DALY</td>
<td>$1000/DALY</td>
<td>$1000/DALY</td>
<td>$1000/DALY</td>
<td>$1000/DALY</td>
<td>$1000/DALY</td>
</tr>
</tbody>
</table>

**Figure 49-27.** Factors affecting utilization and outcome of surgical care. (Adapted with permission from UNFPA United Nations Population Fund (UNFPA); Setting standards for emergency obstetric and newborn care; Available from: https://www.unfpa.org/resources/setting-standards-emergency-obstetric-and-newborn-care. Illustration reproduced with permission from Intermountain Healthcare.)
The Public Health Approach to Management of Gallbladder Disease in Mongolia

Mongolia, the most sparsely populated country in the world, covers a large geographic area nestled between China and Siberia.\textsuperscript{147} The austere environment with extremes of weather, dry deserts, and high mountains present significant obstacles for road building limiting transportation for patients in the vast rural areas (Fig. 49-28). Significant deficiencies in infrastructure, supplies, equipment, and human resources at primary healthcare facilities exist: sporadic electricity, no fully qualified surgeons or anesthesiologists, and less than half the facilities with running water.\textsuperscript{3} In 2006, Healthcare expenditures reached only U.S. $23.2 per capita.\textsuperscript{148,149}

The second most common cause of inpatient morbidity in Mongolia has transitioned to gastrointestinal diseases with liver disease, appendicitis, and gallbladder disease the top three causes.\textsuperscript{150} While laparoscopic cholecystectomy was introduced in Mongolia in 1994, by 2005 only 2\% of gallbladders were removed laparoscopically, and then, only in the capital city.\textsuperscript{151} A cohort study in 2005 comparing open with laparoscopic cholecystectomy by Dr. Sergelen, the chief of surgery at the Health Sciences University of Mongolia (HSUM), found the wound infection rate to be significantly lower, hospital stays shorter, and hospital expenditures 50\% less with laparoscopy compared to open cholecystectomy.\textsuperscript{152}

Dr. Sergelen formulated a plan to expand access to laparoscopic surgery throughout Mongolia. This plan targeted the three main areas affecting utilization and outcome:

1. **Quality of Care:**
   a. Develop a laparoscopic training didactic and practical course to train surgical teams and transfer skills safely.
   b. Improve the surgical infrastructure for each facility.
   c. Expand the surgical residency to include laparoscopic training.

2. **Accessibility of Quality Care:**
   a. Begin training surgical teams in the capital city, but then expand them to four carefully selected regional diagnostic treatment and referral centers (RDTRCs) in all four quadrants of the country.
   b. Invite industry to offer cost-affordable supplies and replacement parts to sustain the laparoscopic equipment in Mongolia.

3. **Socioeconomic/Cultural Factors:**
   a. Educate the public on the increased benefits of laparoscopic surgery so they would initiate lobbying efforts demanding the government increase funding for these services.
   b. Educate government leaders about the need and benefit of laparoscopic cholecystectomy for the Mongolian people.

The strategic initiative that began by expanding laparoscopic cholecystectomy within the capital city and then to the four key Regional Diagnostic and Treatment Referral Centers (RDTRCs) created the foundation for countrywide access to high-quality modern surgery for a regionally prevalent disease (Fig. 49-29).\textsuperscript{147,153}

In 2011, through a multinational partnership (HSUM, the Dr. WC Swanson Family Foundation (SFF), the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES), and the University of Utah Department of Surgery), Mongolia experienced a
dramatic transformation from open to laparoscopic cholecystectomy; laparoscopic replaced open cholecystectomy and became the primary method to remove the gallbladder in Mongolia (Fig. 49-30).154

As people began to see their neighbors return to functional ability faster with the laparoscopic approach, the Mongolian people developed increased trust in their healthcare providers and the quality of care they could receive. The Ministry of Health committed increased funding for laparoscopic surgery and changed existing laws making it easier for hospitals to purchase the needed equipment and supplies solidifying the needed financial and business models to support laparoscopic surgery in Mongolia.

By 2016, with the introduction of laparoscopic training into the surgical residency program and development of laparoscopic fellowship training for surgical teams from outlying provinces, 17 of 21 provinces now provide laparoscopic cholecystectomy allowing patients the benefits of less pain, smaller incisions, fewer wound infections, and more rapid return to work (Fig. 49-31).

---

**Figure 49-29.** The public health approach to expanding laparoscopy to the regional diagnostic treatment and referral centers of Mongolia (RDTRCs). (Reproduced with permission from Intermountain Healthcare.)

---

**Figure 49-30.** Transition from open to laparoscopic cholecystectomy in Mongolia. (Reproduced with permission from Wells KM, Lee YJ, Erdene S, et al: Building operative care capacity in a resource limited setting: The Mongolian model of the expansion of sustainable laparoscopic cholecystectomy, Surgery. 2016 Aug;160(2):509-517.)
Integrating Value into Global Surgery

Another topic closely related to quality is the concept of value in healthcare, often described in terms of the value equation. The value equation states that value equals quality plus service, divided by cost. In this model, systems that generate high-quality care, at lower costs, produce greater value for stakeholders in that system (i.e., patients, physicians, insurers, and hospitals). This discussion is most pertinent in high-income countries such as the United States, where rising healthcare expenditures are cause for worry. It is also relevant in LMICs, where proposed interventions to improve quality must also control cost. A bidirectional exchange of information related to surgical systems between HICs and LMICs may provide one strategy for reduction of cost in HICs and maximization of quality in LMICs.

Advanced Surgical Care for Resource-Poor Areas

Limited financial, physical, and human resources, political and social conflicts, and austere environments cause many to believe that advanced surgical care is inappropriate in resource-poor countries. Misconception of the needs and abilities of people in LMICs cause some policymakers to discount the desire of people worldwide for advanced surgical care. Developing these capabilities in resource-poor countries has the potential to decrease overall cost and actually develop the infrastructure necessary to entice physicians and other healthcare workers to remain in their own countries. Establishing advanced surgical care requires expertise and services that symbiotically support and improve general medical care. Therefore, many developing countries are actively building capacity and capability to provide the full spectrum of modern surgical care locally.

As economies improve and the benefits of laparoscopic surgery for resource-poor areas become better delineated, patients and doctors, surgical societies, ministries of health, and industries are demanding the benefits of minimally invasive surgery for patients and communities. The economic impact of laparoscopy may be even greater in LMICs than in developed countries. Worldwide, surgeons have identified laparoscopic training as one of their greatest needs. In a 2010 survey, developing laparoscopic and endoscopic skills were identified as the most important skills desired by surgeons from the West Africa College of Surgeons (WACS) (Fig. 49-32).

Transplantation is another area of great interest to people in poor countries partly because of the high prevalence of kidney failure and because chronic dialysis facilities are limited. Hepatoma and liver failure are very common in countries with a strong prevalence of hepatitis B and C. Transplantation has become the treatment of choice for end-stage kidney disease in developed countries as it dramatically improves the quality of life and increases survival rates compared to medical management. Yet, transplantation eludes most of the developing world. Initial attempts to transport critically ill patients from LMICs to developed countries for kidney transplantation were cost-prohibitive. With the alarming increase in the rate at which young people have been presenting with kidney disease in developing countries, the increased utilization placed on the few dialysis machines has been overwhelming. Dialysis units which previously were utilized three times a week, now operate 24 hours a day, 7 days a week, and cannot begin to provide the needed services to the multitudes needing treatment. Even programs to develop peritoneal dialysis cannot fully ease the demand.
The majority of kidney transplants in developing countries are from living related donation because of cultural and legal prohibitions precluding cadaveric transplantation. Laparoscopic living related donation has the potential to increase the voluntary donor pool as patients have less postoperative pain, return to work and activities quicker, and have much better cosmesis than open surgery. Ethical concerns exist for nonrelated donations, however, because of concern for coercion in some countries. Adapting to the limited resources, surgeons have described various cost-saving techniques to facilitate the laparoscopic approach in resource poor areas, such as using endoclips instead of staplers for vascular control, modifications to the surgical approach, and suprapubic extraction of the kidney rather than endocatch removal.

**Academic Global Surgery**

There has been a paradigm shift from traditional reliance on intermittent short-term volunteerism toward a strengthening of the education and research pillars for surgical healthcare in developing regions, a role ideally suited for academic surgery. Global surgery is emerging as a new academic field of endeavor (Table 49-12). Academic institutions have historically pioneered discovery in disease causation and treatment. As globalization expands, academic surgical programs are beginning to respond by broadening their vision and mission. This vision and mission includes interdisciplinary and collaborative approaches to designing innovative, affordable surgical care that is accessible to all through research, education, development, and advocacy.

Responding to the challenges of disparities, new generations of students, faculty, philanthropists, private industry leaders, and policymakers have demonstrated a growing passion to address global surgery as part of global health. Prior to 1984, only 0.32% of physicians and 0.12% of nurses were involved in international health (either paid or volunteer). Recently, interest in global health has exploded among medical students, residents, and faculty in the United States.

**Table 49-12**

**Examples of academic global surgery programs**

<table>
<thead>
<tr>
<th>INSTITUTION</th>
<th>NAME OF GLOBAL SURGERY CENTER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brigham and Women’s Hospital</td>
<td>Center for Surgery and Public Health</td>
</tr>
<tr>
<td>Emory University School of Medicine</td>
<td>Global Surgery Program</td>
</tr>
<tr>
<td>Harvard Medical School</td>
<td>Program in Global Surgery and Social Change</td>
</tr>
<tr>
<td>King’s College London</td>
<td>King’s Center for Global Health and Health Partnerships</td>
</tr>
<tr>
<td>McGill University</td>
<td>Centre of Global Surgery</td>
</tr>
<tr>
<td>Oregon Health and Sciences University</td>
<td>Global Health Advocacy Program in Surgery</td>
</tr>
<tr>
<td>University of British Columbia</td>
<td>Branch for International Surgical Care</td>
</tr>
<tr>
<td>University of California San Francisco</td>
<td>Center for Global Surgical Studies</td>
</tr>
<tr>
<td>University of Chicago</td>
<td>Global Surgery Program</td>
</tr>
<tr>
<td>University of Utah</td>
<td>Center for Global Surgery</td>
</tr>
</tbody>
</table>
Academic involvement in global surgery provides training for the next generation of surgical leaders where they can learn the necessary skills to develop systems for quality and affordable surgical care, both locally and internationally. Leaders for the 21st century will need to know how to provide outstanding cost-effective clinical care for all environments.

Global Surgery programs use a variety of methods to engage faculty, residents, and students to foster an environment of creativity and innovation necessary to generate new ideas for solving some of the most vexing problems in global health. For example, some academic collaborative programs introduce medical, engineering, and business students to the fascinating world of medical device innovation focusing on a variety of communities with very limited resources. Other academic endeavors include global surgery resident rotations, fellowships, journal clubs, classes (locally and online), certificates, Masters of Public Health, bilateral exchange programs, and a variety of educational, research, and entrepreneurial meetings (local, national, and international). Many of these activities provide opportunities for collaborative scholarly work (journal publications, books, book chapters, films, grants, research trials) that present new avenues for academic advancement supporting both colleagues from abroad as well as the home institution. More importantly, Ministries of Health are responding to results of these scholarly works by designing policies that include surgical care in their countrywide health plans.

Partnering academic programs from high-income countries with LMICs and/or with NGOs provides opportunities for collaboration (Box: Academic Global Surgery Partnerships). Global surgery engagements exist along a continuum from simple associations, to alliances, collaborations, or formal partnerships (Fig. 49-33). A true partnership usually involves specified and joint rights and responsibilities. The other engagements vary depending on the amount of integration between the institutions and organizations. Successful sustainable progress for global surgery can occur within any of the different levels of engagement. Many partnerships begin with a simple association or alliance before growing into a formal partnership.

Academic Global Surgery helped progress the foundational knowledge in defining the burden of surgical disease, clarifying the cost-effectiveness of surgical care, and establishing baseline values of surgical capacity. Further work is necessary to move beyond data collection and to use this foundational knowledge to now develop interventional strategies and stimulate sustainable solutions for accessible, affordable, appropriate surgical care for all.

### Academic Global Surgery Partnerships

A. Rwanda Human Resources for Health (HRH) Program

The Rwanda Human Resources for Health (HRH) program is an ambitious 7-year long program of the Ministry of Health (MOH) of Rwanda, funded by the U.S. Government and the Global Fund to Fight AIDS, Tuberculosis, and Malaria. The HRH Program seeks to greatly expand and improve Rwanda’s health care workforce by strengthening national training programs of specialized physicians, nurses, oral health providers, and health managers. The HRH Program is also designed to strengthen the capacity of academic institutions in Rwanda to sustain the training programs initiated and supported by the HRH Program through (a) recruitment and retention of Rwandan faculty, (b) transfer of knowledge and skills to Rwandan faculty, (c) establishment of additional academic partnerships and collaborations between Rwandan academic institutions and U.S. academic institutions. Currently, 22 U.S. academic medical centers and universities are participating in this program, in collaboration with the University of Rwanda—College of Medicine and Health Sciences (UR-CMHS) training faculty. Since the launch of the program in 2012, U.S. institutions have deployed about 100 faculty members per year across these four health-related professions. The recruited U.S. faculty are twinned with UR faculty and senior trainees, paired along common goals and interests, and together they engage in a diversity of activities—including teaching, training, research, clinical care, and care delivery improvement projects.

The program is currently in its fifth year (August 2016–July 2017). Focusing our discussion on the surgical disciplines, the annual intake of postgraduate students (residents) has dramatically increased. The anesthesiology residency, whose annual matriculation ranged from zero to three residents, now routinely admits 10 new residents yearly. The surgery residency has divided into the four specialties of general surgery, orthopedics, urology, and neurosurgery. Annual matriculation across all for programs now ranges from 15 to 20 compared to 3 to 6, prior to the HRH program’s support of the surgery department. In July 2016, UR graduated eight new general surgeons and one new urologist, the largest output to date. Similar training output is noted across the other disciplines and specialties as well. As a result, Rwanda is now on track to achieve most of its targets for the health workforce. Most graduates are deployed across provincial hospitals to provide specialty level care in a decentralized fashion, while a portion are maintained at the teaching hospitals to be recruited as new faculty. The HRH Program also aims to strengthen the quality of the training programs through competency-based training and pedagogic innovation, improvements in infrastructure and equipment within the schools at the CMHS and the teaching hospitals, and stronger administration of the training programs. As the HRH program comes into its final years, efforts are underway towards faculty professional development program that will both ensure that the UR-CMHS is able to
continue making the aforementioned human resource gains independent of this large foreign aid grant, and that the relationships and collaborations forged between academic institutions may continue to grow and find new avenues for productive work together.\textsuperscript{186}

—Robert Riviello, MD, MPH, FACS

### B. Coordinating Nongovernmental Organizations (NGO) and Academic Organizations: IVUmed

Nonprofit organizations (NGO) have filled a niche in establishing surgical care in countries where training centers and healthcare systems are historically nonexistent or understaffed. More recently, professional organizations have developed a focus on specific diseases or patient groups and have become a resource for education and training in poor countries.

For more than 20 years, the IVUmed NGO has focused on urological education and hands-on training in Africa, Asia, and Latin America. IVUmed evolved from a need identified by plastic surgeons that had seen many children with hypospadias and other urological anomalies, such as exstrophy, when providing care for children with cleft lip and palate. Adult surgeons were not trained in the delicate reconstruction of pediatric genitourinary anomalies, and pediatric surgeons were not trained in endoscopic or reconstructive urological surgery. The program has expanded to support training in all aspects of urological care, including adult reconstruction, oncology, and endoscopic management of stones and prostatic disease.

As a nonprofit organization, IVUmed is a partnership between surgeons, anesthesiologists and nurses, academic medical centers, urological professional associations, industry, and the public with urologic surgery training in more than 20 countries. It also provides North American trainees scholarships to travel to low-resource countries to learn and to share knowledge gained in their own programs. Many former scholars become mentors for other residents when they complete their training. The sites with the longest collaborations have developed their own educational programs in general urology or subspecialty areas and are now providing advanced training and care locally (see Fig. 49-33).

### C. Cancer Disparities Consortium in West Africa

Noncommunicable diseases, such as cancer, are a major public health problem in low- and middle-income countries (LMIC). In many LMIC, surgeons, due to the lack of medical oncologists, treat all stages of noninfectious related cancers, such as breast and colorectal cancer (CRC). In 2011, to address the disparity in outcomes for patients with cancer in West Africa compared to the United States, a research and training collaboration was formed between the Obafemi Awolowo University Teaching Hospital in Nigeria and Memorial Sloan Kettering Cancer Center (MSK) in New York. This relationship has now grown to become a consortium of five Nigerian hospitals and the Global Cancer Disparity Initiative Team at MSK.

The consortium began by focusing on important questions regarding CRC: how can early stage patients be identified; what are the demographics of CRC patients in Nigeria; and is the biology of CRC different in Nigeria compared to the USA? These questions are being answered, with the support of two NIH grants, by creating a robust prospective database with a matching biobank. With over 250 patients, the consortium can now describe the metastatic patterns, stages of presentation, and risk factors for CRC in West Africa. Given that over 65% of patients present with stage IV disease, the development of a risk model to identify patients with early stage disease is a priority. This is being accomplished with a 400-patient prospective trial of colonoscopy in patients over 45 years of age with rectal bleeding in three Nigerian cities. Future projects include studying new technologies for CRC and breast cancer screening.

—Peter Kingham, MD, FACS

### Ethics

The ethics involved in working outside one’s own country are complex. While a practitioner’s scope of practice is usually constrained by regulation in America and Europe, in many countries the limits of what one can do are neither regulated nor enforced. Guidelines for what should be done, where, and under what circumstances are beyond the expertise of some ministries of health. Some problems are so episodic that they are not anticipated, and few guidelines exist. For example, in natural disasters and emergencies, should any willing provider from any country be granted permission to provide care? Should specific disaster-related training be encouraged or required?\textsuperscript{187,188} In the nonacute setting, should practitioners not licensed or credentialed in their home environments be allowed to perform volunteer surgery in other countries? What entity should oversee the flow of volunteer practitioners? Can a standard set of guidelines meet the needs of most countries? Currently, there is little cross-national agreement between state entities, like ministries of health and independent organizations and individuals. While many countries require at least temporary licensure, some do not. In many cases enforcement is inconsistent.

With respect to research, the poor historically have not received benefit from research performed on them. In international studies, even local collaborators have been left out of study design and publication.\textsuperscript{189} As internet communications have improved, these lapses are no longer tolerated.\textsuperscript{190} Informed consent for surgical procedures, in the appropriate language and respectful of local customs, is becoming the norm. Few hospitals outside academic medical centers have institutional review boards (IRBs) to oversee the implementation and review of clinical research. In recent years, peer reviewed journals have become more mindful of attribution of credit, and authors are strongly encouraged to design and report studies with local input at all levels.

With regard to transplantation, many countries have laws against cadaveric transplants because of the very real concern for illegal marketing of organs. Even living-donor
transplantation has seen effects of coercion in some regions and for some populations such as prisoners. Nevertheless, the need and popular desire for transplantation is accelerating acquisition of skills and technology to make transplantation available worldwide.191

Finally, what is considered ethical in one country or community might be considered highly unethical in another. Consent for surgery may in one setting rest with the patient, but in another, with the community or family. Values about privacy vary markedly from region to region. Health information in many cultures is considered to be a community concern, not the personal property of an individual patient.

**Innovation in Global Surgery**

The pressing need for surgical care at all levels and the shortage of fully trained surgeons, anesthesiologists, and support personnel as well as equipment and supplies means that opportunities abound for innovation. Innovations in education, including simulation, can shorten the time necessary for learning technical skills. Gaming technology can teach algorithms for interpretation of X-rays and ultrasounds. Telemedicine/telehealth is transforming education through combinations of clinical case-based learning and massively open online courses (MOOC) (Box: Telemedicine). The potential for education in surgery beyond the apprenticeship system championed by Halsted in 1904 is vast.

**Telementoring in Global Surgery**

One excellent example of successful telementoring in surgery is a program started by Allan Okrainec, MD, a minimally invasive surgeon at the University of Toronto, and Georges Azzie, MD, a pediatric surgeon at Toronto’s Hospital for Sick Children. In the mid-2000s, the two imagined utilizing laparoscopic box trainers and videoconferencing technology via Skype to teach minimally invasive techniques to surgeons in LMICs.192 As part of the “Go Global” Initiative of the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES), Dr. Okrainec’s team traveled to Botswana in 2007 to teach a 3-day Fundamentals of Laparoscopic Surgery (FLS) course in person. Although only two of twenty surgeons achieved certification, average posttest scores showed promising improvement.

Realizing that continued mentorship promotes success, Dr. Okrainec’s team reconfigured their teaching model to include telesimulation. In 2009, they carried out an 8-week course in FLS, with weekly meetings via videoconference and real-time simulation demonstration and feedback between Botswana and Toronto. This time, 100% of participants in the tele-simulation group attained certification.193 Subsequently, the team expanded its program to Colombia, with similar success in skill acquisition.194 Since 2009, the program has expanded to 15 countries, with satellite telesimulation sites in Colombia and Ukraine. They have trained more than 300 individuals in FLS skills around the world.

Innovation that radically changes the way we do things and that changes a paradigm of a service or system is called “disruptive”; it abruptly changes an older and more expensive system in favor of a less expensive, more widely available technology or process. The ability for disruptive innovations to transform products and services into affordable realities requires three main factors: a sophisticated technology that simplifies, a low-cost business model, and an economically coherent value network (Fig. 49-34).195

Regulations and standards that vary between countries and locales can facilitate or impede disruptive change. While disruptions often are not qualitatively superior to the status quo, they make the process both less expensive and more accessible, and through multiple iterations, ultimately improve quality as they cycle through the transformative process.

Decentralizing education, laboratory testing, and medical records have been made possible through free and open-source software, apps, and devices such as smart phones, tablets, and laptop computers. Monitoring and imaging devices and laparoscopic instruments designed for low resource environments have the potential to not only improve accessibility in poor countries but also to radically reduce surgical costs in wealthy ones.196

**THE FUTURE FOR GLOBAL SURGERY**

Surgeons of the future will need to educate themselves in areas that have not historically been taught in surgical curricula. Beyond the technical aspects of surgical practice, there is a complex ecosystem that supports surgical care. Surgeons must become more aware of the complexities of cost in order to be able to shape the environment in which they work. They must understand better what patients are seeking from the surgical experience, rather than focusing primarily on a narrow view of what surgery might have to offer. Surgeons must engage in policy development and advocate for affordable and accessible surgical care without sacrificing quality. Thoughtful technology design can focus on improving quality and on decreasing cost, both in poor and wealthy countries. Building surgical capacity through a health systems-focused approach, with robust data collection, and establishment of global surgery centers of
excellence will stimulate improvements in the provision of surgical and anesthesia services. Further integration of surgical care into national health plans of governments should encourage increased investments and political will necessary to create capacity, leading to timely, quality surgical care to all without risk of financial ruin. Our colleagues in public health and the World Bank, Paul Farmer and Jim Kim, have challenged us: “We need our surgical colleagues to speak fluently about rebuilding infrastructure, training, personnel, and delivering high-quality care to the very poorest.”

REFERENCES

Entries highlighted in bright blue are key references.


34. Humphries N, McAleese S, Matthews A, Brugha R. “Emigration is a matter of self-preservation. The working conditions...are killing us slowly”: qualitative insights into health professional emigration from Ireland. *Hum Resour Health*. 2015;13:35.


ENHANCED RECOVERY AFTER SURGERY

History and Overview of Enhanced Recovery

As anesthetic techniques, antibiotics, and minimally invasive surgery have improved surgical care over the centuries, further strategies to continue to improve patient outcomes have emerged. A novel perioperative regimen for patients following colon surgery using early oral nutrition, early mobilization, and epidural analgesia was first described in 1995 by Professor Henrik Kehlet from Copenhagen, Denmark in a small group of nine patients. He subsequently outlined a more detailed multimodal approach to perioperative care in 1997, reducing length of stay to a median of 2 days following sigmoid resection, and thus, has been described as the founder of enhanced recovery after surgery (ERAS). Soon thereafter, several others duplicated the ideas proposed by Professor Kehlet. As there was a great discrepancy between actual practices and evidence-based best practices, as well as geographical and practitioner variations in care, the group desired to create a consensus on best practices with guidelines that could be employed in the clinical arena. The primary goal of ERAS is to treat the surgical patient in a multidisciplinary team approach throughout the perioperative course with the unified goal of accelerating functional recovery and optimizing patient outcomes based on evidence-based medicine (Fig. 50-1).

In order to develop the key tenets of ERAS, the details of preoperative care, intraoperative surgical and anesthetic technique, and postoperative care were scrutinized and standards for each facet of care were developed (Fig. 50-2).

ERAS and its associated principles truly represented a paradigm shift in perioperative care, breaking from the classical teaching of prolonged fasting postoperatively until signs of bowel function, excessive fluid administration, and delayed postoperative mobilization. Instead, mitigating the metabolic and stress responses to surgery through multimodal care and utilizing evidenced based medicine allows for “fast-track” recovery and improved outcomes.

Each facet of perioperative care may have modest benefit to the patient when applied alone, but when an integrated, multimodal enhanced recovery pathway (ERP) is used, the beneficial effects become synergistic. Patients return to presurgical functional states faster, spend less time in the hospital, and experience less morbidity. Furthermore, resource utilization is reduced, healthcare system cost is less, and societal cost is lower with faster return to work and reduced homecare needs. Though initial studies of ERP were primarily performed for colorectal surgery, ERP can be applied to a wide variety of specialties in the inpatient and outpatient setting including urology, orthopedics, and gynecology.

Preoperative Optimization

First proposed in 1949 by the anesthesiologist, Dr. J. Albert Lee, a preanesthetic and presurgical evaluation by an anesthesiologist is associated with improved outcomes for the efficiency of the operating room, the hospital, and most importantly, the patient. The use of a preoperative evaluation results in identifying patients at elevated respiratory risk, a 55% decrease in preoperative testing, an 88% reduction in case cancellations, reduction in day of surgery delays, reduced total length of stay, a positive impact on hospital finances with cost reduction, and lower in-hospital mortality. Therefore, while it is very important for the surgeon to see the patient prior to a surgery, it is also...
Key Points

1. Enhanced recovery after surgery (ERAS) is a paradigm shift in the surgical care of patients. As a multimodal, integrated, evidence-based care pathway, ERAS optimizes patient care in the preoperative, intraoperative, and postoperative setting in order to achieve best patient outcomes. Patients recover faster, experience less physiological stress, enjoy shorter stays in the hospital, and have fewer complications.

2. Setting appropriate expectations, optimizing nutritional and physical status through prehabilitation, and treating medical comorbidities optimizes patients before surgery.

3. Achieving normovolemia both intraoperatively and postoperatively is important in order to maintain perfusion without volume overload, as hypervolemia and hypovolemia are both associated with significant complications. Goal-directed therapy approaches maintain normovolemia with zero fluid balance.

4. Intravenous normal saline administration results in hyperchloremia, which has been associated with increased mortality and morbidity.

5. As pain is a subjective response and cannot therefore be experienced while unconscious, the use of opioids intraoperatively should generally be avoided in order to minimize the multiorgan system side effects of these medications. Additionally, avoiding intraoperative opioids actually improves postoperative pain scores and reduces the need for postoperative opioids.

6. Multimodal analgesia, which includes oral or IV nonopioid analgesia and regional analgesic techniques, can reduce postoperative physiological stress and decrease complications associated with surgery as part of a pain management regimen. Multimodal analgesia has been shown to reduce the number of opioids required for analgesia.

7. The strategies for avoiding postoperative nausea and vomiting include the avoidance of general anesthesia, the use of totally intravenous anesthesia, avoidance of nitrous oxide and volatile agents, minimizing intraoperative and postoperative opioids, and adequate hydration.

8. Enhanced recovery after surgery care pathways can be applied to numerous types of surgery including colorectal, liver, pancreas, bariatric, gynecologic, and urologic surgery with success.

9. Traditional Chinese medicine has been practiced for thousands of years and serves as a distinct cultural heritage of China. Its unique theories and methods are still applied widely in the practice of modern medicine, including disease prevention, disease treatment, and perioperative management.

10. Acupuncture and transcutaneous electroacupuncture can reduce the number of opioids utilized in the perioperative setting. Additionally, acupuncture, transcutaneous electroacupuncture, and some Chinese herb decoctions are effective in the prevention and treatment of postoperative nausea and vomiting.

Setting Expectations and Patient Education. Setting expectations in the preoperative clinic helps to orient patients regarding the entire surgical experience, from what they are expected to do at home before the surgery to the entire length of the recovery both in the hospital and at home. Information on the procedure and typical recovery should be clear, well defined, and consistently reinforced from all healthcare personnel that interact with the patient. Clear expectations of goals prior to surgery, in the hospital, and after discharge should be communicated long before the surgery. Expected length of stay and disposition should also be clearly communicated in order to optimize timely discharge. The preoperative clinic helps in this role and to establish the patient as the leader in his or her own care. The patient must understand that his or her active participation throughout the perioperative experience will facilitate the recovery. A surgeon can do an operation. An anesthesiologist can keep a patient alive while the patient asleep. However, in reality, it is up to the patient to make his or her own care a priority. If individuals train for a race, should they not also prepare for their procedure so that the recovery is swift? The optimization for success at surgery begins preoperatively with smoking cessation, exercise, and nutrition, but it also continues in the hospital and after discharge with pain control, physical activity, discharge planning, and returning to daily activities. Clear expectations at each point in the perioperative continuum, which are communicated to the patient, will improve the perioperative experience for the patient and the provider.

According to Costa, “Evidence shows that patients suffer needlessly due to inadequate preoperative preparation and lack of information regarding their postoperative course as indicated by reports of unexpected pain, fatigue, and the inability to care for oneself.”14 Patients enter physician offices and procedures with a great deal of fear and anxiety related to the identification of a disease, the consequences of treatment of this disease, and/or the fear of death. Eliciting the patient’s concerns and providing optimal communication and education can allay much of this fear and anxiety. There is a substantial perioperative culture change that is brought about by the elements of ERAS, and many patients have personally had or have had a close relative that has experienced surgery that likely did not include many...
components of ERAS. Therefore, this culture change must be clearly disseminated to the patients and include new fasting guidelines, analgesic management, and patient participation in preoperative optimization. Patients can then anticipate and plan for certain events and sensations, such as what and when to eat and drink, how to exercise, what medications will be used, what tubes or lines will be present, and what criteria are used for discharge and return to daily activities. This kind of teaching has been defined as a therapeutic communication to help the patient face and cope with the surgical procedure in a calm manner.  

Approximately 90 million people have difficulty understanding and subsequently acting upon health information distributed to them. Further, the ability of patients to process and understand basic information to make appropriate health decisions is directly related to socioeconomic status and to Caucasian race; furthermore, the readability of patient-directed healthcare material may be too advanced for comprehension by much of the surgical population. Thus, it is appropriate for patient information material to be at no higher than a sixth grade reading level, friendly, clear, concise, and simply designed. The employment of audiovisual aids may assist some patients as well.

**Nutrition.** Surgery results in a significant catabolic stress response on the body, triggering inflammation and nutrient depletion. This stress response results in downstream effects on numerous organ systems and can lead to a higher risk of postoperative complications. Ensuring preoperative adequate nutrition is imperative before a large surgical procedure in order to mitigate adverse outcomes.  

While enteral or parenteral nutritional supplementation can be considered for the most nutritionally compromised patient, the enteral route is always preferred if clinically appropriate and can be adequately achieved in a timely fashion. Two main approaches to preoperative enteral nutrition include standard oral nutrition supplements and immunonutrition supplements, each providing extra protein and calories to supplement the diet. There is no statistical difference in infections, complications, and length of hospital stay between patients given standard oral versus immunonutrition supplements.  

Immunonutrition supplements usually similarly contain high protein, vitamins, and minerals, but they also have the addition of arginine to improve immunity and tissue repair and omega-3 fatty acids to mediate the inflammatory response. The exact dosage of arginine and omega-3 fatty acids that contribute to improved outcomes is not known. Standard oral nutrition products may contain no or lesser quantities of arginine and omega-3 fatty acids when compared to immunonutrition supplements. The exact duration and frequency of supplementation have not been established; however, it is clear that the use of standard oral nutrition products has a positive impact on surgical outcomes by reducing postsurgical complications.

**Exercise and Prehabilitation.** Prehabilitation is defined as “the process of enhancing the functional capacity of the individual to enable him or her to withstand a stressful event.” Both exercise and prehabilitation, which have, heretofore, focused on cardiopulmonary rehabilitation prior to surgery, are very important to optimization of patient outcomes. Perioperative cardiopulmonary exercise testing and prehabilitation in relation to ERAS programs around the world have been analyzed, and it is noted that a reduction in fitness prior to surgery is associated with increased mortality and morbidity in the postoperative arena. Patients who actively exercise even when suffering from documented coronary artery disease, heart failure, hypertension, diabetes, chronic obstructive pulmonary disease, depression, dementia, cancer, and stroke have better outcomes. Furthermore, other literature supports the significant merits of exercise therapy and cardiopulmonary exercise therapy before and after major surgery, with the ability to reduce infection, hospital-associated complications, length of stay, and postoperative mortality.

There is a significant amount of evidence indicating that exercise training is feasible and safe in patients with a spectrum of severe cardiac and pulmonary diseases as many of these patients require surgery to manage other disease processes. A randomized controlled trial involving 246 low-risk patients undergoing cardiac surgery reported a 1-day reduction in ICU stay and a reduced hospital length of stay in the intervention group. Cardiopulmonary fitness was found to be a strong independent predictor of survival after lung surgery, especially for non–small cell lung cancer. Preliminary nonrandomized

---

**Table: ERAS Components**

<table>
<thead>
<tr>
<th>Preoperative phase</th>
<th>Intraoperative phase</th>
<th>Postoperative phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient education &amp; set expectations</td>
<td>Pain blocks</td>
<td>Goal-directed fluid therapy</td>
</tr>
<tr>
<td>Smoking &amp; alcohol cessation</td>
<td>Minimally invasive surgery</td>
<td>Prevention of PONV</td>
</tr>
<tr>
<td>Prehabilitation with diet &amp; exercise</td>
<td>Goal-directed fluid therapy</td>
<td>Early ambulation</td>
</tr>
<tr>
<td>Shortened fasting</td>
<td>Multimodal pain relief</td>
<td>Early feeding</td>
</tr>
<tr>
<td>Antimicrobial prophylaxis</td>
<td>Prevention of PONV</td>
<td>Multimodal pain relief</td>
</tr>
<tr>
<td>VTE prophylaxis</td>
<td>Avoidance of tubes, drains, lines</td>
<td>Early urinary catheter removal</td>
</tr>
<tr>
<td>Minimize bowel preparation</td>
<td>Normothermia</td>
<td>Defined discharge criteria</td>
</tr>
</tbody>
</table>

**Figure 50-2.** Phases and components of an enhanced recovery after surgery pathway. VTE = venous thromboembolism; PONV = postoperative nausea and vomiting.
data from patients undergoing elective rectal cancer surgery within an ERAS program have shown the feasibility of providing a cardiopulmonary exercise interval training program that is delivered three times per week for 6 weeks in a hospital setting after neoadjuvant chemoradiotherapy and before surgery. The interval following neoadjuvant therapy offers a unique window of time to improve the fitness and nutrition of many oncologic patients, which is especially important as these patients can often be some of the most debilitated.

The addition of physical fitness and activity to a preoperative regimen for elderly patients undergoing major abdominal surgery significantly improved mortality, discharge to home versus a care facility, and length of stay. While this study showed that patients benefited from 60-minute sessions 2 to 4 weeks prior to surgery, even brief therapy before surgery, ranging from 1 day to 7 days before major abdominal surgery, have been associated with a significant decrease in postoperative complications. Therefore, some oncologic or semiurgent patients may actually benefit from preoperative cardiopulmonary exercise programs.

Smoking Cessation. Clearly tobacco use, especially smoking, has been well documented across all surgical specialties to increase postoperative mortality, as well as increase postoperative complications including prolonged ventilation, pneumonia, deep venous thrombosis, wound infection, delayed wound healing, and reduced bone fusion. Physiologically, the carbon monoxide and nicotine from tobacco products increase heart rate and blood pressure and the body’s demand for oxygen. Nicotine also causes vasoconstriction, reducing perfusion to many tissue beds. Obviously, it is advantageous for patients to cease smoking preoperatively. There is debate about the duration of the nicotine free days needed preoperatively to offer best outcomes, though the literature suggests that the longer duration of smoking cessation prior to surgery portends better outcomes. This is likely to allow for bronchiolar and collagen remodeling and the several weeks following last exposure that are needed to achieve blood free of nicotine and its derivatives.

Metabolic Stress Response to Surgery

Multiple organ systems interact in numerous metabolic and inflammatory cascades following the stress response to surgery leading to insulin resistance and protein catabolism (Fig. 50-3). Neuroendocrine responses, stress hormones cascades, activation of cytokine, and immune reactions all occur, leading to a catabolic state. Central to this metabolic and inflammatory cascade is the development of insulin resistance, whereby a normal insulin concentration results in a subnormal biologic response. As insulin is the main anabolic hormone involved in glucose control, fat metabolism, and protein balance, insulin resistance disrupts many metabolic pathways.

Hyperglycemia from insulin resistance results from an increase in glucose production and a decrease in glucose uptake by the periphery. In a fed state, insulin levels surge to 6 to 8 times basal levels, which stops glucose production and increases peripheral glucose uptake three- to fourfold. When fasting, insulin levels remain in a relative steady state with minimal effects on glucose and protein metabolism. Even when insulin levels increase to three times basal levels, there is no increase in peripheral glucose uptake. Therefore, in the postoperative fasting state, without the assistance of exogenous insulin, peripheral..
glucose uptake is reduced. The resulting hyperglycemia can be corrected, however, with the use of exogenous insulin, and when normoglycemia is achieved in the perioperative period, the main components of metabolism also normalize.51

The preoperative and postoperative fasting state triggers insulin resistance resulting in a catabolic state with gluconeogenesis and protein breakdown. Following prolonged fasting states with stress such as following surgery, protein catabolism can be increased several fold over baseline. As less glycogen is stored in the muscle and loss of lean body mass occurs, there is less muscle function and therefore less capacity to mobilize.

In addition to the typical metabolic effects of surgery, pain has been demonstrated to increase insulin resistance. In healthy male volunteers undergoing painful stimulation, glucose uptake was reduced as a direct result of decreased insulin sensitivity.52 In addition, serum cortisol, epinephrine, and free fatty acids were all increased following painful stimulation.

Elective surgery results in a state of insulin resistance, with the magnitude of surgery corresponding to a decrease in insulin sensitivity.53 For example, the difference in insulin sensitivity following laparoscopic cholecystectomy versus open cholecystectomy is 2.5-fold. More complex abdominal surgery such as an open colorectal resection results in a 3.5-fold increase in insulin resistance over laparoscopic cholecystectomy. As levels of insulin resistance increase, complications also increase.54 In addition to an association with complications, insulin resistance has been shown to be an independent predictor of length of stay.55

The increased postoperative complications associated with insulin resistance may be not only from the direct metabolic effects of insulin on glucose but also from free radical formation. In peripheral tissues that are independent of insulin metabolism, and therefore do not store glycogen, the increased plasma glucose levels result in greater glycolysis and oxygen free radical formation. This leads to alterations in gene expression, which in turn propagates a cycle of increased inflammation causing even more insulin resistance.56 Elective surgery has been implicated in increased inflammatory gene pathways and changes in insulin signaling genes in both adipose and skeletal muscle tissues.57-59

Components of an ERP may offset the metabolic and stress responses of surgery. Preoperative carbohydrate supplementation has been shown to counter the catabolic effects of the fasted state by stimulating glucose uptake and transitioning metabolism to a more anabolic state with improved insulin sensitivity.58 Preoperative carbohydrate supplementation also reduces protein loss and improves muscle strength postoperatively.59-61 When preoperative carbohydrate supplementation is added to epidural analgesia, there is even greater improvement in insulin resistance.55

Preoperative Fasting and Preoperative Carbohydrate Loading

Traditionally, patients have been instructed to fast for 6 to 12 hours before surgery to reduce the risk of aspiration of gastric contents during the induction of anesthesia. This fasting state results in a prolonged period without nutrients or hydration prior to and during surgery, and it can lead to insulin resistance, hyperglycemia, failure to achieve a postsurgical anabolic state, and sometimes, the need for insulin. Both European and American Societies of Anesthesiology guidelines have supported the use of clear liquid oral intake up to 2 hours prior to surgery with the exception of patients with gastroparesis, intestinal obstruction, or dysphagia.62-64 Carbohydrate oral intake up to 2 hours prior to surgery does not increase aspiration in healthy adults undergoing elective surgery and in fact reduces preoperative hunger, thirst, anxiety, and nausea.65,66 In addition, a fasting time of 2 to 4 hours versus more than 4 hours actually results in smaller gastric volume and a higher gastric pH value.67-75 Nevertheless, prior to the introduction of ERAS, the dogma of prolonged nothing-by-mouth status widely adopted many decades ago had little advocacy to change, despite evidence supporting more liberal fasting parameters. Current guidelines support fasting from clear liquids for 2 hours and solid food for 6 hours.

Preoperative carbohydrate loading prior to surgery in the form of a carbohydrate rich clear liquid improves patient nausea and discomfort over preoperative water hydration or a prolonged fasting state.76 It may also have further benefits over low carbohydrate clear liquid beverages or fasting by changing the overnight fasting state to a fed state and thus altering glucose, protein, and fat metabolism by increasing postoperative insulin sensitivity.77-79 Furthermore, in patients whose expected length of stay is greater than 2 days, there is a significant length of stay reduction in patients that receive preoperative carbohydrate loading.80 The best carbohydrate loading drink is unclear as studies are heterogeneous and the carbohydrate content is variable. Nevertheless, the carbohydrate drink should be hypotonic for faster gastric emptying, result in a fed state with full glycogen stores, and reduce postoperative insulin resistance. The most commonly studied carbohydrate loading drink includes 100 g of carbohydrate the evening prior and 50 g of carbohydrate 2 to 3 hours prior to surgery.

Intraoperative Considerations

Surgical Considerations.

Prevention of surgical site infection consists of the use of mechanical, chemical, and/or antimicrobial modalities. Mechanical and chemical methods include the use of patient bathing preoperatively and skin preparation with betadine, chlorhexidine, or similar chemical in order to limit the microbial content of the skin. Additionally, the appropriate use of antimicrobial prophylaxis should be employed and follow guidelines specific to the type of surgery and for duration of antibiotic prophylactic administration.

Minimally invasive surgical approaches should be considered as minimally invasive techniques have demonstrated improved outcomes across surgical specialties, including reductions in length of stay and postoperative complications. Additionally, the use of catheters or drains should be limited unless necessary, as these hinder the patient’s perceived ability for ambulation.81-84

Hypothermia Prevention.

Hypothermia is a common perioperative problem. Up to 90% of all patients undergoing elective surgery suffer from inadvertent postoperative hypothermia.85,86 Those at highest risk include patients over the age of 60 years, and/or patients that have malnourishment, preexisting hypothermia, preexisting medical comorbidities that impair body temperature regulation (including advanced diabetes with nephropathy and hypothyroidism), who are undergoing general anesthesia, and who are undergoing a major long surgery. Further, in patients who experience hypothermia, surgical complications are increased, including impaired wound healing, wound infection, pressure ulcers, cardiac disorders including arrhythmia and infarction, as well as increased bleeding requiring blood transfusion (Table 50-1).
The reasons for hypothermia are multifactorial. Radiation, the transfer of heat by electromagnetic waves through space without a medium, accounts for 50% to 70% of heat loss. Convection, the loss of heat through ambient air stream, accounts for 15% to 25% of heat loss. Evaporation accounts for 5% to 20%, and conduction accounts for 3% to 5%. Temperature reduction can also be accelerated by cold intravenous fluids, low operating room temperatures, and a decreased thermoregulatory threshold, which occurs during the administration of general anesthesia. Further, the ability to compensate for reduction in body temperature is also compromised by muscle relaxation and anesthesia in general, as these processes impair shivering and thermoregulatory vasoconstriction. There are steps to take to prevent this hypothermia including active, convective heating using clean, filtered, forced-air warming blankets in patients in the preoperative area (prewarming) and also during anesthesia; thermal insulation; warmer ambient operating room temperatures, warmed irrigation solutions during surgery; and warmed infusions and blood products.

**Venous Thromboembolism Prophylaxis.** Venous thromboembolism (VTE), which includes deep venous thrombosis (DVT) and pulmonary embolism (PE), is the number one cause of potentially preventable death in common but preventable causes of morbidity and mortality in the perioperative patient. Several national quality improvement organizations have cited VTE prophylaxis for patients at risk as a priority for individual physicians and for hospitals because this intervention reduces adverse patient outcomes and hospital costs.

Surgical patients have increased risk for VTE due to advanced age, multiple medical comorbidities, prolonged procedure times, the inflammatory and hypercoagulable state of surgery, and immobility. Specific risk factors include major general, vascular, or orthopedic surgery; lower extremity paralysis due to spinal cord injury; fracture of the pelvis, hip, or long bones; multiple trauma; cancer; prior VTE; age 40 years and higher; obesity; immobility; oral contraceptive use; hypercoagulability syndromes; and severe cardiopulmonary disease (prior myocardial infarction, congestive heart failure, chronic obstructive pulmonary disease).

Postoperative DVT is usually asymptomatic, and fatal PE can often be the first sign of VTE. DVT occurs after approximately 25% of all major surgeries without prophylaxis, and PE occurs after 7%. Since screening modalities (such as venous duplex imaging) in asymptomatic patients have low sensitivity to detect clot, the best approach is to systematically apply prevention strategies to all patients undergoing surgery, with treatment choices based on patient-related and procedure-related risks.

VTE prophylaxis is therefore an important component in optimal perioperative care and current surgical practice and should be included in all practice guidelines. Appropriate VTE prophylaxis should be given preoperatively, intraoperatively, and postoperatively based upon current guidelines for the surgery type. Examples of nonpharmacologic methods include early ambulation, graduated compression stockings, and intermittent pneumatic compression devices. Pharmacologic methods include the use of low dose unfractionated heparin, low molecular weight heparin, and in some case, factor Xa inhibitors.

**Perioperative Fluid Management.** Current and traditional fluid management strategies, which are based on a fixed fluid requirement per patient per case, have failed to improve outcomes. More modern goal-directed therapy (GDT) intravenous fluid approaches rely on the use of advanced medical devices, including esophageal Doppler monitors and other noninvasive cardiac output or bioimpedence models to determine whether or not patients are “fluid responsive” during surgery. In the setting of a normal ejection fraction, fluid is only administered when the expectation is that cardiac output will increase, and vasopressors are utilized if the aforementioned devices show fluid will not increase cardiac output. Excess fluid in certain general surgical cases can cause ileus and bowel edema, and in cardiac cases, it can cause hemodilution. Patients randomized to restricted and liberal fluid resuscitation strategies found a clear linear relationship between total fluids administered (and weight gain) and complications following colorectal surgery including pulmonary edema and tissue-healing complications. Further multiple studies exist demonstrating fewer complications with normovolemia than with liberal strategies of fluid resuscitation.

It must be understood that goal-directed therapy does, in no way, mean reduction in fluid administration. For some procedures, it may be necessary to administer more than anticipated fluid volumes (orthopedics), while for others, the opposite may be true (abdominal). Normovolemia is important to maintain perfusion without volume overload. Thus, the idea behind goal-directed therapy is to maintain zero fluid balance coupled with minimal weight gain or loss. Hypovolemia is associated with reduced circulating blood volume, decreased renal perfusion, altered coagulation, microcirculation compromise, and endothelial dysfunction, among other processes. Hypervolemia is associated with splanchic edema, decreased pulmonary gas exchange secondary to pulmonary edema, impaired wound healing, anastomotic dehiscence, decreased mobility, altered coagulation, and endothelial dysfunction, amidst others processes (Fig. 50-4).

Esophageal Doppler is a mode by which ultrasound is used to monitor and guide intraoperative fluid management has been used quite frequently. The use of this device for fluid optimization has been studied in several randomized controlled trials or meta-analyses, all of which showed a significant reduction in length of stay of up to 4 days. Reductions in length of stay have been seen in gastrointestinal surgery, trauma surgery, urologic surgery, and also the orthopedic population. Alternative devices such as arterial waveform analyzers and pulse oximeter waveform analyzers have been studied and may be promising with the added advantage of lower cost over esophageal Doppler.

Postoperatively, once the patient is adequately tolerating at least a liquid diet and maintaining adequate hydration,
supplemental intravenous fluids should be minimized or terminated. The use of Dopplers or other volume status wave form analyzers have not been studied in the nonventilated postoperative patient and therefore cannot be used to reliably assess volume status. Clinical judgment based on patient factors, surgery type, and the clinical findings should be considered in the decision to continue intravenous fluids. However, once the patient is able to maintain adequate hydration, supplemental fluids should be used judiciously to limit fluid overload, tissue and lower extremity edema, and the constraints that the intravenous medication pole has upon patient-initiated ambulation.

It is not enough to have normovolemia, but one must also consider the type of fluid that should be used for resuscitation. From a recent Cochrane review, there is no evidence that colloids are superior to crystalloid for resuscitation in patients. Therefore, crystalloid fluids should generally be the primary intravenous fluid during the perioperative course. In cardiac surgery, the utilization of 0.9% normal saline solution was associated with hyperchloremia and poor postoperative outcomes, including higher length of stay and increased mortality. Further, a more balanced crystalloid, such as Plasma-Lyte, was associated with improved outcomes in 22,851 surgical patients. In this study, there was a 2.05 odds ratio predictor of mortality with normal saline. Other complications such as acute kidney injury, gastrointestinal complications, major hemorrhage, and major infection were also increased in the group of patients that were hyperchloremic after normal saline administration. Based on such evidence, it would seem prudent to proceed with a more balanced solution, such as Plasma-Lyte, to reduce complications.

**Perioperative Pain Management**

According to the International Association for the Study of Pain (IASP) Taxonomy, the definition of pain is described as “an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage.” There are two important implications of this: pain is completely subjective in that it is whatever the patient says it is, and patients cannot experience pain while unconscious. This has some effects on how the anesthesiologist treats sympathetic stimulation in the operating room. Heretofore, elevations in heart rate and blood pressure were treated with opioid medications, as these sympathetic markers were considered surrogates for patients experiencing pain while under general anesthesia. However, other medications can be used to treat the sympathetic response to surgery such as β-blockers and deepening the anesthesia without administering opioid medications while the patient is unable to experience the pain.

The mainstay of alleviating pain has historically relied almost exclusively on opioids, especially with the usage of patient controlled analgesia devices (PCAs). However, limiting opioids in the perioperative setting is of substantial benefit. Opioids, in fact, reduce pain immediately after administration. However, they also worsen pain scores after they wear off, increase postoperative opioid requirements, increase nausea and vomiting, cause respiratory depression, reduce gastrointestinal motility, worsen urinary retention, induce endocrine dysfunction, and suppress the immune system. There have been a number of randomized controlled trials that have shown that as opioid administration increases, pain scores and postoperative nausea and vomiting increase. Exposure to any fentanyl or opioid in the operating room worsens postoperative pain scores and should therefore be limited or omitted. Opioid containing PCAs have been the standard for opioid administration because of their safety and efficacy in patients to control the administration of opioids. However, because a PCA only offers opioid medication, there is the possibility that the desired analgesic effect will be associated with the aforementioned complications of opioids.

Despite their disadvantages, opioids are still quite useful in the treatment of pain. However, the ERAS protocols focus on opioids as a single component of a comprehensive pain relief strategy, not as the mainstay for treatment. Instead, multimodal
analgesia should be emphasized by utilizing multiple medications to limit postoperative pain and therefore opioid use. Multimodal analgesia mitigates the side effects of opioids by opioid reduction and enhances pain management. Preoperative and postoperative administration of acetaminophen and celecoxib or other nonsteroidal anti-inflammatory drugs, as well as gabapentin have been shown to be efficacious.\textsuperscript{130-135} Intraoperatively, the utilization of ketamine, lidocaine, and magnesium, act as adjunctive measures to limit pain and have been utilized to reduce the utilization of opioids in the postoperative period.\textsuperscript{136-140} Administration of lidocaine and ketamine can also be continued in the postoperative setting (Table 50-2).

Neuraxial opioid analgesia, the administration of opioids through either the intrathecal or epidural route, can be accomplished by either a single shot (both spinal and epidural) or catheter-based therapy (epidural). The use of opioids by this route was shown to have improved pain relief when compared to preoperative oral, IV, or intramuscular morphine.\textsuperscript{144} Further, neuraxial opioid analgesia is associated with lower postoperative pain scores in adults and children who undergo surgery.\textsuperscript{145} Neuraxial analgesia can also be performed with local anesthetic only. Finally, the American Pain Society (APS) recommends the utilization of such postoperative analgesic methods in patients who undergo major surgeries, including thoracic and abdominal

<table>
<thead>
<tr>
<th>Table 50-2</th>
<th>ADVANTAGES</th>
<th>DISADVANTAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analgesic medications</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Opioids</strong>\textsuperscript{141}</td>
<td>Relieve pain immediately after administration</td>
<td>Worsen pain scores after opioids wear off</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Increase postoperative opioid requirements</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Central Nervous System:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Euphoria/dysphoria</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sedation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Respiratory depression</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nausea/vomiting</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cough suppression</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cardiovascular:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bradycardia/tachycardia</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arterial/venous dilation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Renal:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Antidiuretic</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Increased sodium resorption</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Urinary retention</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gastrointestinal:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Decreased motility</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Constipation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Biliary colic</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Endocrine:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stimulates release of antidiuretic hormone</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stimulates release of prolactin</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stimulates release of somatotropin</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Decreases luteinizing hormone</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Immunologic:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Decreased immune system function</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cancer growth</td>
</tr>
<tr>
<td><strong>NSAIDs</strong>\textsuperscript{142}</td>
<td>Reduce inflammation</td>
<td>Renal insufficiency</td>
</tr>
<tr>
<td></td>
<td>Synergistic effect with opioids</td>
<td>Increased bleeding</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Delay healing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adverse cardiovascular risk</td>
</tr>
<tr>
<td><strong>Local anesthetics</strong>\textsuperscript{140}</td>
<td>Opioid-sparing effect</td>
<td>Cardiac toxicity</td>
</tr>
<tr>
<td></td>
<td>Decrease PONV</td>
<td>Central nervous system toxicity</td>
</tr>
<tr>
<td></td>
<td>Reduce ileus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Possible anticancer effect</td>
<td></td>
</tr>
<tr>
<td><strong>Ketamine</strong>\textsuperscript{136,137,143}</td>
<td>Opioid-sparing effect</td>
<td>Dysphoria</td>
</tr>
<tr>
<td></td>
<td>May prevent opioid-induced hyperalgesia and chronic pain syndromes</td>
<td>Hallucinations</td>
</tr>
<tr>
<td><strong>Gabapentinoids</strong>\textsuperscript{130-135}</td>
<td>Opioid-sparing effect</td>
<td>Visual disturbances</td>
</tr>
<tr>
<td></td>
<td>Reduce opioid side effects</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reduce postoperative pain</td>
<td></td>
</tr>
</tbody>
</table>
procedures, cesarean sections, and hip and lower-extremity surgeries; this is especially recommended for patients at risk for cardiopulmonary complications or prolonged ileus.\textsuperscript{145}

Multimodal analgesia can also include regional analgesic techniques such as peripheral nerve blocks, paravertebral blocks, plexus blocks, and local infiltration, which can reduce postoperative physiological stress and decrease complications associated with surgery as part of a regimen.\textsuperscript{146} These techniques have been shown to reduce the amount of opioids required for analgesia and also have been shown to reduce the adverse events seen with epidural local anesthetics (such as urinary retention and hypotension) and/or opioid-containing PCAs.

### Postoperative Nausea and Vomiting Prevention

Postoperative nausea and vomiting (PONV) is very common and can cause significant distress to patients, with the incidence of vomiting at approximately 30\%, nausea at 50\%, and the combination of PONV as high as 80\%. All result in poor patient satisfaction, increased recovery room length of stay, and higher costs to the healthcare system.\textsuperscript{147-152} This could further increase the time to first feeding, which in turn may prolong ileus and/or hospital stay.

Dr. Gan and colleagues developed a consensus guideline for the management of nausea and vomiting and details the risk and possible choices for the treatment of PONV.\textsuperscript{153} Risk factors include female sex, history of PONV or motion sickness, nonsmoking, younger age, general versus regional anesthesia, use of volatile anesthetics and nitrous oxide, postoperative opioids, duration of anesthesia, and the type of surgery including cholecystectomy, laparoscopy, gynecological, and strabismus. The strategies for avoiding PONV include the avoidance of general anesthesia, the use of totally intravenous anesthesia, avoidance of nitrous oxide and volatile agents, minimizing intraoperative and postoperative opioids, and adequate hydration.\textsuperscript{150,152,154-159} The medications to prevent, abort, and reduce PONV include perphenazine, aprepitant, dexamethasone, scopolamine, dolasetron, granisetron, and ondansetron, among others.\textsuperscript{160-164} PONV should be targeted before it occurs for optimal prevention (Fig. 50-5).

#### Early Nutrition and Postoperative Ileus Prevention

Postoperative ileus is the most common cause of prolonged hospital stay and readmissions following surgery on the digestive tract, occurring in up to 19\% of cases.\textsuperscript{165} Not only is this adverse to the individual patient clinically, this also results in doubling of the total cost of the index hospital stay and thus carries a tremendous socioeconomic impact globally.\textsuperscript{166} Numerous risk factors contribute to postoperative ileus and include open surgery, increased surgery length of time, blood transfusion, fasting, fluid overload, opioids, postoperative nausea and vomiting, and other pharmacological agents. While some risk factors are unavoidable in certain patients, others are modifiable, and therefore minimization of the risk of postoperative ileus is achievable.

Nasogastric tubes (NGTs) were previously used prophylactically to prevent ileus, limit distension on the gastrointestinal anastomosis, as well as to prevent pulmonary complications. However, NGT use actually delays return of gastrointestinal activity and increases pulmonary complications without preventing anastomotic leaks in numerous types of surgery, including gastroduodenal, biliary, trauma, and esophageal.\textsuperscript{167-169} Therefore, the routine use of NGTs for prophylaxis should be avoided.

Addressing the numerous risk factors for postoperative ileus has a benefit on the reduction of the incidence of postoperative ileus. For example, mitigating the surgical trauma through

---

**Figure 50-5.** Risk factors of and treatment options for postoperative nausea and vomiting. PONV = postoperative nausea and vomiting.
minimally invasive surgery and meticulous surgery with minimal blood loss reduces postoperative ileus, either directly by limiting the inflammatory response with smaller incisions or indirectly through reduced opioid use.\textsuperscript{170-174} Anesthetic technique can also aid in prevention of postoperative nausea and vomiting, which in turn reduces the need for parenteral opioids, a known risk factor for postoperative ileus. Multimodal pain strategies and neuraxial blocks reduce opioid use and therefore minimize nausea, improve early enteral nutrition, limit intravenous fluid administration, and improve ambulation. Maintenance of normovolemia in the perioperative setting should be achieved as fluid overload and dehydration both negatively affect return of bowel function, length of stay, and complications.\textsuperscript{105,106,175} Clearly, each facet of the perioperative care process is intricately intertwined to the next and has implications in total body homeostasis.

Other measures may also assist in the prevention of postoperative ileus; however, the role of each in the setting of an ERAS pathway is unclear. Chewing gum is hypothesized to reduce postoperative ileus by stimulating the cephalovagal reflex and is considered a form of sham feeding. Prior to the introduction of ERAS, the use of chewing gum in multiple abdominal surgeries demonstrated faster intestinal recovery with variable impact on length of stay.\textsuperscript{176-178} Following the use of ERAS and associated early enteral feeding, the benefit of chewing gum is less clear.\textsuperscript{179,180} Alvimopan is a mu opioid receptor antagonist that is administered prior to surgery and twice daily postoperatively. Pooled analysis of phase III trials demonstrated a reduction in postoperative NGT use, faster return of bowel function, and earlier discharge by 0.7 days.\textsuperscript{181} However, following ERAS implementation, small studies demonstrate a reduction in ileus and length of stay with the use of alvimopan in open surgeries without a benefit in laparoscopic surgery.\textsuperscript{182-184}

**Mobilization**

Though no metric of mobilization has been clearly defined, early mobilization following surgery is an important component of ERAS that accelerates the return to baseline functional status. Prolonged postoperative bedrest leads to deconditioning, increased deep venous thrombosis risk, and loss of muscle mass. Deterioration of mobility and activities of daily living can be seen in older patients after only 2 days of hospitalization.\textsuperscript{185} Therefore, preoperative encouragement of an exercise program and perioperative mobilization can have dramatic impacts on not only the elderly but all patients undergoing surgery. Patients that begin a preoperative exercise program are more active postoperatively and have a faster return to baseline exercise capacity when compared to patients undergoing a postoperative exercise program.\textsuperscript{186} As patients in an ERAS program are directed to spend time out of bed and to ambulate, early mobilization is therefore encouraged. Many other facets of ERAS will assist in this early mobilization: postoperative nausea prevention, limiting drain use, and improved pain control. Setting preoperative expectations of mobility through patient education in the clinic setting and postoperative nursing unit engagement in promoting mobility provide the proper setting for improved patient compliance with early and frequent mobilization. However, compliance with this is highly variable and difficult to track and may be hindered if pain in inadequately controlled or if the patient is tethered to devices such as drains, catheters, and IVs. As with other components of ERAS, engagement of all providers of patient care from outpatient nursing, to inpatient nursing, physicians, and physical therapists, will improve compliance not only with mobility but often with other components as well. Reinforcement of expectations can be achieved with preoperative educational pamphlets, postoperative daily task lists, mobility logs, and pedometers.

### ERAS in CRS

As the initial development of ERAS pathways occurred in colon surgery patients, the largest preponderance of data exists in this surgical specialty with the first ERAS guidelines developed in 2012 after many studies demonstrated positive outcomes applying the basic tenets of ERAS surgery.\textsuperscript{84} In 1997, Henrik Kehlet published the initial series of patients, applying novel perioperative care strategies to colon resection patients.\textsuperscript{23} Following this, several groups also applied these principles to colon resection patients. These studies demonstrated that the principles of early mobilization, early feeding, and optimized intravenous fluid administration resulted in patients tolerating a diet sooner, faster return of bowel function, and earlier discharge.\textsuperscript{187-189} Furthermore, these studies elucidated reduction in complications such as urinary tract infections, ileus, and cardiopulmonary complications. Nygren et al also demonstrated that muscle strength and lung function were less reduced after colon resection within an enhanced recovery protocol compared to traditional perioperative care.\textsuperscript{187} Additionally, as ERAS is expected to diminish the metabolic and hormonal stress response to surgery, attenuation of TNF-α, IL-1β, IL-6, and IFN-γ occurred after ERAS cases when compared to traditional perioperative care, and cortisol levels were not found to increase immediately postoperatively in ERAS, while those with traditional perioperative care experienced immediate and sustained cortisol elevation.\textsuperscript{190}

Larger series of patients followed with Delaney et al publishing a single institutional experience of 1000 consecutive laparoscopic colectomy patients with short length of stay and low readmission and mortality rates.\textsuperscript{191} The protocol was further applied to rectal surgery with success.\textsuperscript{187,192} Meta-analyses and systematic reviews demonstrate less opioid use, shorter length of stay, decreased morbidity, and no increase in readmission rates for laparoscopic or open colon or rectal resections when an ERAS protocol is utilized.\textsuperscript{193-197} Several groups have even discharged patients in as little as 24 hours following colon resection, with Gignoux et al even discharging patients on the same day.\textsuperscript{198-200} ERAS can also be applied to octogenarian patients with compliance to the protocols and with no increased readmission or mortality rates.\textsuperscript{201} Also, patients with diverting stomas can benefit from ERAS protocols, though diverting ileostomy may slightly delay discharge over patients with no ileostomy.\textsuperscript{202}

Adherence to the numerous tenets of ERAS is inversely related to length of stay postoperatively in colorectal surgery.\textsuperscript{203,204} When compliance with ERAS measures is lower, length of stay is longer. The strongest predictors for shorter duration of stay include preoperative carbohydrate loading, no nasogastric tube, early mobilization, early oral nutrition, totally intravenous anesthesia, early removal of urinary catheter, and the use of nonopioid analgesia. Predictors for deviation from an ERAS program and thus resultant longer length of stay include pathologic diagnosis, intraoperative complications, high blood loss, surgery length, lack of mobilization, emesis, persistent use of intravenous fluids, reinsertion of urinary catheter, and poor pain control.\textsuperscript{205-208} Though readmissions are no higher than traditional perioperative care, several factors have been implicated in readmission, including poor ERAS compliance and...
preoperative neoadjuvant chemoradiation.

Predictive tools using artificial neural networks may assist in clinical decision-making. The most common reasons for readmission include bowel obstruction and skin and soft tissue infection. While patient and quality outcomes are clearly important, the added benefit of cost savings has been demonstrated. Surgery using the tenets of ERAS is both less costly for direct and indirect costs to the hospital and effective. Furthermore, patients in an ERAS program returned to work faster and had less caregiver burden, resulting in huge indirect savings as well.

**ERAS in Hepatopancreatico-biliary Surgery**

An initial experience applying an ERAS protocol in 61 consecutive patients undergoing liver resection demonstrated 92% of patients tolerating a diet on postoperative day 1, a reduction in length of stay from 8 to 6 days, and no increased readmissions or morbidity. A subsequent randomized trial in patients undergoing open major liver resection demonstrated a length of stay reduction from 7 to 4 days with a decrease in medically related postoperative complications and no increase in surgical complications or readmission; ERAS patients also reported improved quality of life over controls. Several meta-analyses have similarly supported the use of ERAS protocols in liver surgery, citing reduced morbidity, hospital stays, cost, and time to recovery of bowel function without increasing mortality or readmission rates. When adherence to all elements of an ERAS protocol for liver resection was less, hospital length of stay was longer.

In 2012, the ERAS Society published recommendations for patients undergoing pancreaticoduodenectomy after several studies published early outcomes in this surgical population. Patients undergoing pancreaticoduodenectomy often have high rates of delayed gastric emptying; the use of ERAS has reduced the incidence of delayed gastric emptying by nearly half, thus allowing earlier feeding in this complex patient population. Additionally, multiple prospective cohort or retrospective studies have shown that ERAS protocols offer significant benefits to patients undergoing both distal pancreaticoduodenectomy and pancreaticoduodenectomy with reduced hospital stay and complications. Several meta-analyses or systematic reviews have confirmed shorter length of stays, decreased complications, and lower cost with ERAS protocols. Furthermore, the use of ERAS protocols in elderly patients undergoing pancreaticoduodenectomy continue to show improved outcomes in length of stay and morbidity, showing that ERAS protocols allow this more fragile patient population to recover faster.

**ERAS in Gastrectomy and Esophagectomy**

Patients undergoing foregut surgery have notoriously been subjected to prolonged periods of nasogastric tube decompression and resultant starvation while surgical dogma dictated this fasting time diminished the risk of complications from anastomotic leak. Nevertheless, more liberal removal of the nasogastric tube and limited fasting, as components of ERAS protocols, have demonstrated improved recovery and outcomes. Randomized controlled trials demonstrate that removal of the nasogastric tube in the operating room and early feeding, as components in an ERAS program, result in shorter length of stay, fewer grade III or higher postoperative complications, and faster return to baseline weight and functional status. Multiple nonrandomized studies and meta-analyses verify reduction in length of stay and no increase in complications. In 2014, consensus guidelines for ERAS after gastrectomy were published, and these include no routine use of nasogastric decompression, early feeding within the first postoperative day, and early consideration for nutritional support if the patient is malnourished or unable to maintain at least 60% of caloric requirements.

Esophagectomy surgery is notoriously complicated and fraught with complications secondary to multiple factors including surgical complexity and medical comorbidity. Postoperative management is governed by the idiosyncrasies of the operating surgeon more so than many other specialties and therefore heterogeneous. While studies of ERAS in gastrectomy suggest no routine use of nasogastric tubes and include early feeding, most ERAS programs for esophagectomy encompass all components not related to feeding, but prolonged nasogastric decompression remains. Many of these patients, however, do receive early enteral nutrition through the use of jejunostomy tubes commonly placed at the time of resection. When defined protocols are followed in this cohort of patients, length of stay is reduced, and complications and readmissions are, at a minimum, not increased. Systematic reviews demonstrate a reduction in length of stay, anastomotic leak, and pulmonary complications without increased mortality or readmission.

**ERAS in Bariatric Surgery**

Bariatric surgeons have applied clinical pathways to both the preoperative and postoperative periods for many years, which have resulted in improved outcomes. The adoption of ERAS protocols in these clinical pathways has offered further success for these patients. A randomized trial for laparoscopic sleeve gastrectomy demonstrated a reduced length of stay to 1 day postoperatively in ERAS patients, and others have discharged patients on postoperative day 1 following Roux-en-Y gastric bypass. Furthermore, earlier discharge of patients on postoperative day 1 has not been demonstrated to increase resource utilization, with no increase in patient phone calls, emergency department visits, or readmissions. A meta-analysis confirms success of ERAS in bariatric surgery with reduction of length of stay without increase in complication or complication severity, while a second meta-analysis demonstrates an increase in minor complications without increasing patient morbidity. Following a thorough review of the literature supporting its use, the ERAS Society published guidelines for ERAS bariatric protocols in 2016.

**ERAS in Other Surgical Specialties**

Though ERAS has been applied more broadly to complex abdominal surgery, there is surprising little data in its use in large ventral hernia repair and other abdominal wall reconstructive techniques. Three studies report experience in open large ventral hernia repair with varying techniques of abdominal wall reconstruction including myofascial release. Each study cites faster return of gastrointestinal function and reduction in length of stay by up to 2.5 days. Furthermore, there were no increases in readmission, postoperative complications, or reoperation. Though no long-term follow-up, there is no report that early feeding results in intestinal compromise from “tight” closure or early hernia recurrence. In fact, as ERAS reduces the incidence of postoperative vomiting and ileus, it is likely very beneficial for this patient population.

ERAS has been introduced to non–general surgery subspecialties as well. Complex urological procedures such as radical cystectomy have trialed ERAS over the last decade with favorable results. In this patient population, length of stay was reduced, and complications were similar to or reduced when...
Setting Up an ERAS Program

The successful implementation of an enhanced recovery program depends heavily upon cultural change and excellent organizational behavior. As the ERAS program encompasses so many facets of patient care, the implementation team should not only include surgeons and anesthesiologists but also inpatient and outpatient nurses, pharmacists, information technology specialists, compliance officers, and hospital administration.

Initial strategy for ERAS implementation should define the scope of practice change by identifying the current state and the goal state of care. Protocol content can be discussed in a small group of engaged stakeholders. Once the protocol elements are defined, all stakeholders should review and discuss the protocol in detail in order to identify barriers to implementation, identify solutions to these barriers, and finalize the protocol. This allows for all stakeholders to remain engaged and have ownership in the protocol. Appropriate informational resources for hospital and office staff education are created in order to have a thorough and successful educational campaign. Patient informational resources should also be developed in order to set clear expectations throughout the perioperative process. Additionally, and importantly, standardized order sets are also developed to ensure that all components of an ERAS program have little variation in order to improve compliance. A final implementation date is defined after coordinating that all stakeholders are indeed ready for launch. Appropriate educational campaigns are performed in a timely fashion prior to final implementation to outpatient office staff, inpatient units, preoperative and postoperative care units, operating room personnel and physicians, and midlevel providers and trainees (Fig. 50-6).

Compliance and auditing should be done with relative frequency in the beginning, from weekly to biweekly and then monthly. Team meetings with all stakeholders present allow the team to address any issues in timely fashion with a multidisciplinary approach and thus upholds accountability. The frequency of checkpoint meetings can be reduced over time. Sharing data of successes and failures keeps the team engaged. It is also recommended that a financial team be employed. The financial savings after adoption of an ERAS protocol can be substantial from reduction in length of stay, medication use, and resource utilization. Ideally, a portion of the cost savings should be funneled back into the ERAS program in order to ensure program maintenance and compliance and also to assist with expansion to other service lines or patient care improvement projects.

TRADITIONAL CHINESE MEDICINE IN SURGICAL PATIENTS

History of Traditional Chinese Medicine

Traditional Chinese medicine is one of China’s outstanding national cultural heritages and the quintessence of China, sharing a deep history and common homology with the Chinese culture. It is derived from the rich experience and theoretical knowledge that Chinese people have used to combat disease for thousands of years. The achievements of traditional Chinese medicine caught the attention of the world because it examines each function of the human organism and adjusts those functions to achieve ultimate balance.

Traditional Chinese medicine is the oldest medicine in China. It is a medical system with unique theory, style, diagnoses, and treatments, which were gradually formed throughout the historical medical practice of the Chinese nation. Its development not only depended on the practice but also resulted from the systemic mode of thinking and Chinese philosophy of protecting life shape. Yet, it is a traditional subject that still stands in the modern world of science.

Traditional Chinese medicine theory mainly resulted from the summary of practice and was continuously enriched and developed in practice. As early as 2000 years ago, Huang Di Nei Jing wrote the earliest existing theory of traditional Chinese medicine in China, and it summarized the treatment experience and medical theory to that time. Combining the achievements of other natural science branches and simple materialism and dialectical thinking of the Chinese culture, Huang Di Nei Jing comprehensively expounded the knowledge of human anatomy, physiology, and pathology and also stated the diagnosis, treatment, and prevention of disease, thus founding the preliminary theoretical basis of traditional Chinese medicine.

Based on herbal remedies of primitive people, Shen Nong Ben Cao was the earliest existing monograph on herbal pharmacology in China. It summed up the study of herbology to date, including 365 kinds of drugs until the Han Dynasty. Long-term clinical practice and modern scientific research show that the effects of the drugs described in the book are mostly correct.

In the third century, the famous physician Zhongjing Zhang of the Eastern Han Dynasty delved into the classical medical books such as Su Wen, Zhen Jing, and Nan Jing and extensively collected the effective prescriptions, combining this with his own clinical experience. He published a famous book about typhoid fever, Shang Han Bing Za Lun. This book established the theoretical system and treatment principles that applied dialectical therapy of Chinese medicine and influenced the future of traditional Chinese medicine.

In 610 A.D., Yuanfang Chao et al wrote Treatise on the Etiology of Various Diseases, which is the earliest existing monograph on causes of symptomatology in China. This book also recorded the intestinal anastomosis, abortion, tooth extraction, and other operations, indicating the breadth of surgery practiced at that time. Subsequently, the ancient Chinese government in 659 A.D. issued Tang Xiu Ben Cao; it was not only the first pharmacopoeia of ancient China but also the first national pharmacopoeia of the world. It was published 883 years earlier than the Nuremberg Pharmacopoeia, which was issued by the European Nuremberg government in 1542 A.D.

The Tang Dynasty physician Simiao Sun authored Bei Ji Qian Jin Yao Fang and Qian Jin Yi Fang. In these two books, clinical subjects, acupuncture, dietary therapy, disease prevention, and life preservation were discussed. These were an outstanding achievement of the time, especially in the prevention and treatment of nutritional deficiency diseases.

Between the 12th and 14th centuries, i.e., the Jin and Yuan eras of China, several new Chinese medicine theories emerged. There were four representative scholars. Wansu Liu (1120–1200 A.D.) thought the symptoms of shanghan (exogenous febrile disease) were related to “excessive internal heat,” so the herbal
Figure 50-6. Implementation process of an enhanced recovery after surgery pathway. ERAS = enhanced recovery after surgery; CRNA = certified registered nurse anesthetist; NP = nurse practitioner; PA = physician assistant; IT = information technology; PACU = postanesthesia care unit; APP = advanced practice provider; LOS = length of stay; PCA = patient-controlled analgesia.
characteristics of cold and cool were used in treatment, and this was known as the “cold and cool” treatment style. Congzheng Zhang (approximately 1156–1228 A.D.) thought the cause of disease was “exogenous evil” (exogenous pathogenic factor) invading the human body; therefore, the treatment focused on “eliminating evil,” through the diaphoresis, emetic, and/or purgative methods, which were known as the “offensive precipitation” style. Dongyuan Li (1180–1251 A.D.) proposed that “internal injury of viscera causes various diseases,” and pyretic tonification of the spleen and stomach was emphasized in the treatment, known as the “invigorating the spleen” style. Zhenheng Zhu (1281–1358 A.D.) thought “yang was always excessive, while yin was always insufficient” in the body, so the treatment was focused on nourishing yin and cutting down heat, which was known as the “nourishing yin” style.

In approximately the 11th century, the Chinese began to use human pox vesicle exposure as vaccination to prevent smallpox and thus pioneered the field of medical immunology and vaccination. In the 17th to 19th centuries, due to the epidemics of infectious diseases, the seasonal febrile disease theory developed. This theory broke the traditional Chinese medicine long-term conception that the pathogens invaded from the surface to the inside of the body. In the mid-17th century when bacteriology had not yet appeared, this was undoubtedly a great pioneering work and laid the groundwork for the epidemiology of the spread of disease.

In the period from the Opium War (1838–1842 A.D.) to the founding of the People’s Republic of China in 1949 A.D., Western medicine was introduced to the continent of China. However, Western medicine and Chinese medicine theories were unique to each other. There was almost no interchange between them, either in theory or in practice; yet together, they formed a unique and relatively mature theoretical system. In the 1920s to 1930s, Western medicine proposed abolishing Chinese medicine. At the same time, Chinese medicine did not deny the merits of Western medicine but thought that Chinese medicine was superior to Western medicine. However, there were several advocates to combine the theories who published such works as “Chinese Medicine for Main, Western Medicine for Use,” “Using Their Respective Strengths, Reaching the Same Goal by Different Means,” and “Chinese Medicine Treat Internal Disease, Western Medicine Treat Surgical Disease.”

In short, the pervasive opinion was to learn from the other’s strong points to make up for one’s deficiencies. This forged the trail for establishing modern Chinese integrative medicine.

In the past decade, a series of significant progresses and breakthroughs have been made in the modern study of traditional Chinese medicine theory and its application to clinical practice. For example, acupuncture anesthesia can be used for small splint fixation, and the treatment of acute abdomen integrates both traditional Chinese and Western medicine. Most lately, Tu Yo Yo’s team found that artemisinin, derived from the wormwood plant, treats malaria, and the team was awarded the Nobel Prize in 2015.

Clearly, integration of traditional Chinese medicine with modern science and technology will advance knowledge and treatment. Modern diagnostic instruments and techniques have become auxiliary methods of clinical diagnosis and treatment of traditional Chinese medicine, making up for the deficiency of the traditional “four ways of diagnosis” methods and improving the accuracy and efficiency of traditional Chinese medicine diagnosis and treatment. Thus, modern science and technology’s infiltration, transformation, and integration in all aspects of Chinese medicine will be one of the distinctive characteristics of the future development of Chinese medicine.

In summary, traditional Chinese medicine and pharmacy are an important part of the splendid culture of the Chinese nation, making outstanding contributions over thousands of years because of its systemic theory, distinctive treatment methods, significant efficacy, and abundant historical documentation. Modern Chinese medicine includes traditional Chinese medicine and thus offers integrated and superior health service. This integration of Eastern and Western medical ideas and philosophies are important for the future of this modern medical era.

Preoperative Nutritional Optimization

Many surgical abdominal diseases have a long incubation period before clinical presentation, during which time the patients may develop malnutrition including specific nutrient deficiencies and hypoproteinemia. These comorbidities directly influence the surgical treatment and postoperative effect of the patient undergoing an operation. In this setting, some experts advocate the use of parenteral nutrition and other adjunctive nutritional measures. These measures can often improve the patient’s nutritional status, but they are difficult to popularize because of the high cost of treatment and common complications with parenteral nutrition.

In traditional Chinese medicine theory, it is thought that there are a variety of “asthenic symptoms” in patients who need surgical treatment and that applying the treatment principle of “treating deficiency with tonification” improves conditions throughout the body. On the basis of the traditional Chinese medicine theory of the “concept of holism” and “treatment according to syndrome differentiation,” the all nourishing decoction (Shi quan Dabu decoction) and Buxong Yiqi decoction are used in patients with the “deficiency of vital energy and blood syndrome” before surgery and have achieved good results. Similar success has been shown with the Shenmai injection and Astragalus injection. For example, when traditional Chinese medicine is used effectively to treat patients with breast cancer before an operation, it enhances the general body status, improves the patient’s energy, and regulates the liver and kidney functions, which ultimately promotes the success of the operation and controls progression of the tumor.

Bowel Preparation for Surgery

In gastrointestinal surgery, the method of “purgation and offensive precipitation” is used in preparation for the operation. Either Large Chengqi decoction or Seasoning Chengqi decoction significantly increases the gastrointestinal motility and washes the gastrointestinal stagnation to prepare the bowel for surgery. These decoctions also improve visceral blood flow and peritoneal absorption, promote early recovery of postoperative bowel function, and prevent superimposed infection of the intestine and the effect of endotoxin.

Preoperative Optimization During Sepsis and Infection

Due to infection, endotoxemia, blood loss, and other factors, many critically ill patients will deteriorate or progress to shock. Initial surgery during profound shock may be counterproductive until adequate resuscitation is achieved. In addition to resuscitation, blood transfusion, and antimicrobial treatment if indicated for sepsis, the traditional Chinese medicine treatment based on dialectics can help to create a favorable condition for surgery, if indicated.
For traumatic shock and anaphylactic shock patients, the use of flavored pure ginseng decoction supplemented by blood transfusion and other comprehensive antishock measures, can rapidly raise blood pressure and provide a more optimal setting for surgical treatment.\textsuperscript{282} Qingdan decoction and Qingyi decoction, which are composed of herbs to clear away heat and toxins, remove stasis, purge the bowel, and are used to treat acute obstructive suppurative cholangitis and hemorrhagic and/or necrotizing pancreatitis.\textsuperscript{283-285}

**Perioperative Pain Management**

The application of traditional Chinese medicine in perioperative pain management is has become more prevalent in recent years. Research of acupuncture for analgesia began in the 1950s in China and has developed into a combined acupuncture and medicine anesthesia that is currently recognized by the medical field.\textsuperscript{286} During surgery, the combined anesthesia is composed of acupuncture and opioid drugs. Experimental data show that the combined acupuncture anesthesia could reduce the dose of opioid by 50\%.\textsuperscript{287} Acupuncture also helps the management of postoperative pain by enhancing the level of the endogenous opioid, encephalin. Encephalin restrains the pain signal from being transmitted to the central nervous system, blocks the body’s reaction to pain, and increases the pain threshold accordingly.\textsuperscript{288} In several clinical trials, the needed doses of opioid at 8, 24, and 72 hours postoperatively were reduced in acupuncture groups compared to control groups.\textsuperscript{289} In addition, electroacupuncture has been demonstrated to be effective at alleviating postoperative pain and assisting in recovery. One randomized controlled trial shows electroacupuncture significantly reduced the dose of fentanyl used, improved the quality of recovery and decreased the incidence of anesthesia related side effects for patients undergoing surgery.\textsuperscript{290}

**Postoperative Nausea and Vomiting Prevention**

PONV is a common complication after surgery. Acupuncture and herbs can be applied to prevent and treat PONV. Acupuncture or transcutaneous electroacupuncture improves stress-induced impairment in gastric motility functions, significantly inhibits the frequency of transient lower esophageal sphincter relaxations in response to gastric distention, and suppresses retrograde peristaltic contraction.\textsuperscript{291-293} In clinical trials, intraoperative P6 acupuncture point (Neiguan acupuncture) stimulation during surgery significantly reduces the incidence of PONV over 24 hours, and the efficacy of P6 stimulation is similar to that of commonly used antiemetic drugs in the prevention of PONV.\textsuperscript{294} Furthermore, electroacupuncture restrains the release of gastrointestinal peptides and consequently relieves PONV with an efficacy comparable to ondansetron.\textsuperscript{295} Moreover, a randomized, prospective, double-blinded clinical trial shows that auricular acupuncture within 24 hours postoperatively has a similar effect to prevent PONV.\textsuperscript{296}

Additionally, some Chinese herb decoctions, such as Liu Jun Zi decoction and Cheng Qi decoction, administered perioperatively, show possible effectiveness to reduce the severity of PONV and to relieve abdominal distension.\textsuperscript{297,298}

**Early Nutrition and Postoperative Ileus Prevention**

In traditional Chinese medicine, acupuncture and decoctions have a role in the prevention of postoperative ileus. For example, early acupuncture on Zusanli, Shangjuxu, and Xijuxu acupoints, combined with early enteral nutrition, can effectively improve gastrointestinal function and shorten the length of stay after surgery.\textsuperscript{299} In addition, Dachengqi decoction applied to patients after laparotomy improves gastric dysrhythmia, promotes intestinal peristalsis, and enhances gastrointestinal motility.\textsuperscript{300,301} Furthermore, in one randomized trial, the combination of Simo decoction and acupuncture reduces the incidence of postoperative ileus and shortens hospital stay for patient undergoing abdominal surgery when compared to the perioperative use of chewing gum.\textsuperscript{302} Other traditional Chinese medicine methods such as electroacupuncture combined with Evodia hot compress, confers benefit in postoperative recovery of gastrointestinal function of patients who have undergone abdominal surgery.\textsuperscript{303}

**Traditional Chinese Medicine in Common Surgical Conditions**

**Colon Surgery.** Several traditional Chinese medicine decoctions can assist in bowel preparation prior to surgery. For example, during the bowel cleansing before surgery, Dachengqi decoction can be used to promote bowel peristalsis and evacuation, thus preventing contamination during surgery and reducing the risk of postoperative complications of infection. Additionally, after colon surgery, traditional Chinese medicine therapies such as acupuncture and decoctions prevent postoperative ileus, reduce the incidence of PONV, and promote the recovery of colon function. As a result, traditional Chinese medicine can shorten the hospital stay after colon surgery.

**Appendicitis.** In general, patients suffering from acute appendicitis will undergo appendectomy. Exceptionally, when a case of acute simple appendicitis or a periappendiceal abscess is encountered, Chinese herbs together with antibiotics can serve as an alternative treatment. In early acute simple appendicitis, oral Qinghua decoction can help the appendix infection resolve without surgical management.\textsuperscript{304} This nonsurgical treatment also can be applied to patients with periappendiceal abscesses that is unsuitable for surgery. These Chinese medicine herbs activate blood flow, dissolve stasis, clear heat, and remove toxicity.

**Biliary Disease.** Cholelithiasis is a common disease of the biliary tract that can result in cholecystitis and the possible need for cholecystectomy. Several traditional Chinese medicine herbs can relieve the symptoms of acute cholecystitis and delay the progression of the disease, possibly preventing the need for cholecystectomy. These Chinese medicine herbs are composed of herbs for clearing heat and secreting bile (Qing Re Li Shi), herbs for promoting circulation of Qi and relieving pain (Xing Qi Zhi Tong), and herbs for clearing heat and promoting diuresis (Qing Re Li Shi). According to different Chinese medicine therapies, the use of these herbs can also be applied to patients postoperatively from abdominal or biliary surgery to adjust biliary excretion and/or prevent cholestasis.

Unfortunately, severe cholecystitis or cholangitis may progress to liver abscess. When a liver abscess develops, percutaneous catheter drainage and Chinese medicine herbs are applied in Chinese medicine practice.\textsuperscript{305} Generally, in the early stage of abscess, Chinese medicine herbs are used for clearing heat and removing toxicity (Qing Re Jie Du) and for promoting blood circulation and removing blood stasis (Huo Xue Hua Yu). When abscess is evident, Chinese medicine herbs for clearing heat and cooling blood are added.\textsuperscript{306} The main function of these Chinese medicine herbs is to relieve infection, reduce inflammation, and activate intestinal motility in order to evacuate the
toxicity, reduce the inflammatory response, and thus hasten recovery.

**Pancreas Surgery.** In recent years, the early use of traditional Chinese medicine and enteral nutrition treatment in patients after pancreatectoduodenectomy, can help to hasten recovery in gastrointestinal function, improve nutritional status and immune function, and also reduce postoperative complications. The combined use of total parenteral nutrition and *Astragalus* injection can improve the nutrition status of patients with obstructive jaundice and improve the immune function of these patients. All Nourishing (Shiquan Dabu) decoction has the role of increasing the level of plasma albumin and hemoglobin, which can be used as a recipe in surgical nutrition therapy.

**Intestinal Obstruction.** Adhesive ileus is the most common type of small intestinal obstruction and also is the kind to which traditional Chinese medicine therapies apply widely. The methods of traditional Chinese medicine treatment include acupuncture, Chinese herbal enema, and gastrointestinal intubation. Acupuncture or transcutaneous electroacupuncture applied to acupoints such as Zusanli point, Neiguan point, Zhongwan point, and Tianshu point has remarkable regulatory effects on gastrointestinal function and can promote relief of obstruction. Dachengqi decoction combined with the Chinese medicine herbs for clearing heat and removing toxicity (Qing Re Jie Du), and for promoting blood circulation and removing blood stasis (Huo Xue Hua Yu), can enhance gastrointestinal motility, improve blood circulation of the intestine, reduce intestinal capillary permeability, protect the barrier function of the intestinal mucosa, and help inflammatory edema to resolve. Additionally, acupuncture, Chinese herbal enema, and gastrointestinal intubation can shorten the time of the obstructive event and reduce the length of stay in patients suffering from adhesive intestinal obstruction. These methods of traditional Chinese medicine treatment can also be applied to the treatment of postoperative ileus.

**REFERENCES**

Entries highlighted in bright blue are key references.


73. Read MS, Vaughan RS. Allowing pre-operative patients to drink: effects on patients’ safety and comfort of unlimited oral water until 2 hours before anaesthesia. Acta Anaesthesiol Scand. 1991;35:591-595.


86. Moola S, Lockwood C. Effectiveness of strategies for the management and/or prevention of hypothermia within the adult perioperative environment. Int J Evid Based Healthc. 2011;9:337-345.


107. Mc Ardle GT, McAuley DF, McKinley A, et al. Preliminary results of a prospective randomized trial of restrictive versus...


266. Zhou B, Sun R. About the cultural support of the development of traditional Chinese medicine. The 12th National Symposium on Chinese Medicine Culture. 74-79.


Understanding, Evaluating, and Using Evidence for Surgical Practice
Andrew J. Benjamin, Andrew B. Schneider, Jeffrey B. Matthews, and Gary An

INTRODUCTION

The singular importance of this chapter rests on the following chain of reasoning:

1. The state of surgical science and knowledge is constantly changing.
2. The education of a surgeon is a continuous process.
3. Surgeons need to know how to evaluate new surgical knowledge to maintain their education in order to best serve their patients.
4. This chapter provides guidance as to how surgeons might navigate, interpret and apply this new knowledge.

Notably, this reasoning also applies to the process of acquisition of new knowledge itself, which explains why this inaugural chapter on evidence-based medicine is occurring in the 11th edition of this book. Recognizing the impermanence and fluidity of knowledge is a critical insight for the responsible surgeon, but so too is realizing that “good practice” cannot occur without reference points as to what should be done given the current imperfect state of knowledge. These dual recognitions inform the organization of this chapter, which introduces and describes the currently accepted approach to evidence-based medicine, and then follows by noting a series of current issues that anticipate the likelihood that what is meant by “evidence-based medicine” will evolve over the coming years. To a great degree, the goal of the Scientific Method, through its emphasis on skepticism and falsifiability, is predicated upon overriding observational/experiential bias by the application of rigorous methodology statistical analysis. The dangers of bias were recognized at the dawn of the Scientific Era, and continue to manifest today (Box: The History and Sources of Bias in Biomedical Literature).

WHAT IS EVIDENCE-BASED MEDICINE?

For centuries, the practice of medicine was guided primarily by anecdotal experience, often based on rationales that did not arise from a rigorous scientific process and sustained by the fundamental barriers associated with being able to learn from one’s experience (e.g., cognitive bias). For example, treatments such as bloodletting and purging were based on ostensible principles of bodily humors originating from the Ancient Greeks, and persisted well into the 18th century despite repeated disastrous outcomes. To a great degree, the goal of the Scientific Method, through its emphasis on skepticism and falsifiability, is predicated upon overriding observational/experiential bias by the application of rigorous methodology statistical analysis. The dangers of bias were recognized at the dawn of the Scientific Era, and continue to manifest today (Box: The History and Sources of Bias in Biomedical Literature).

The History and Sources of Bias in Biomedical Literature

In Sir Francis Bacon’s initial description of the scientific method in his Novum Organum he notes what he terms the “idols of the mind,” in essence recognizing and classifying the sources of cognitive bias that limit the reliability of subjective observation and interpretation. These “idols” are:

- Idols of the tribe (Idola tribus): A humans’ tendency to perceive more order and regularity in systems than truly exists, and arises from their preconceived ideas about things
- Idols of the cave (Idola specus): Arising from an individual’s personal limitations in reasoning due to particular personalities, subjective likes and dislikes
Key Points

1. Cognitive bias is inescapable, and limits the ability of both individual practitioners and the surgical field in general, to advance and improve in a scientific fashion. Evidence Based Medicine is an attempt to codify the process of interpreting experience, assessing the literature and translating it into practice.

2. Dealing with and interpreting the vast amount of surgical literature available on the Internet can be daunting, and this task can be aided by the application of identified formats for executing online search. The PICO (Patient/Population, Intervention, Comparison and Outcome) format is a commonly used method for codifying online search.

3. Not all literature or evidence is created equal. There exist various approaches, such as the Oxford Center for Evidence Based Medicine (CEBM) Levels of Evidence or the GRADE (Grading and Recommendations, Assessment, Development and Evaluation) system, that have been developed to provide guidance in assessing and reifying scientific literature.

4. The conversion of evidence into clinical practice often manifests in the creation of clinical guidelines. As with all things related to evidence based medicine, not all guidelines are created equal, and therefore there are certain characteristics that can be used to evaluate the quality of a particular clinical guideline.

5. There are specific challenges in the application of evidence based medicine to surgery, not least of which is the difficulty in performing a truly randomized clinical trial. The CONSORT (Consolidated Standard of Reporting Clinical Trials) guidelines were developed to serve as minimal recommendations for reporting randomized clinical trials.

6. The well-known saying “There are lies, damn lies and then statistics” points to the recognition that statistical tools can be prone to misuse, and emphasizes the need to understand the appropriate application, limits of and interpretation of reported statistics.

7. Evidence based medicine has not thus far been held to its own standards of evidence. Recognizing that available “evidence” is a constantly shifting landscape should warn one against the dangers of epistemic certainty, and further emphasizes the fact that surgical education is an ongoing and perpetual process.

In the medical field, the transition from accumulated anecdote to true statistical analysis can be seen in the emergence of clinical epidemiology as a field in 1938, which began to shift the focus from descriptions of individual patients to trends affecting entire populations. This shift, however, was accompanied by new challenges, as different means of turning anecdotal experience into statistics (e.g., case series, observational studies, retrospective studies, prospective studies) meant that now practitioners needed to be able to compare these “scientific” presentations against each other in order to best establish their practices. The processes and methods of aggregating, comparing, and translating these different types of data from the medical literature into clinical practice were explicitly established in the latter part of the 20th century, particularly arising from efforts at McMaster University, which eventually led to a fundamental framework for literature-informed medical decision-making known as evidence-based medicine (EBM).

EBM is defined as the “conscientious, explicit and judicious use of current best evidence in making decisions about treating individual patients.” This term was coined by Gordon Guyatt in 1991, focusing on assessing the credibility of the medical literature, understanding the presented results, and applying the information to individual practice. EBM is defined by three epistemological principles:

- Principle 1: Not all evidence is created equal, and the practice of medicine should be based on the best available evidence
- Principle 2: The pursuit of truth is best accomplished by evaluation of the totality of the evidence, and not selecting evidence that favors a particular claim
- Principle 3: Clinical decision-making requires consideration of patients’ values and preference

The adoption of EBM in the discipline of surgery has lagged compared to nonsurgical specialties. To a great extent, this is due to the challenges of achieving the highest level of evidence noted in principle 1: definitive conclusions from a randomized controlled trial (RCT). A literature analysis of MEDLINE from 1966 to 2000 demonstrated that only 15.1% of the 134,689 RCTs evaluated a surgical topic. In the early days of EBM during the 1990s, surgical RCTs accounted for only 7% of published articles in surgical journals; most of the articles were retrospective studies and case series, which are essentially aggregated anecdotes. Over the next decade, the relative frequency of RCTs in surgery further decreased, accounting for 3.4% of all publications in 2003. As a result, most of the available evidence to guide surgical practice today remains based on retrospective reviews, nonrandomized trials, and expert opinion. The barriers to performing prospective RCTs in surgery remain substantial: standardization of clinical presentation and, of course, accounting for variations in operative technique and the ability to blind studies to reduce experimental bias. The relative paucity in RCTs in surgery make it even more
important that surgeons understand the best-practice methods to critically appraise available evidence, while recognizing the limitations and potential pitfalls of those methods, in order to optimize their practice and decision-making regarding patient care when high quality evidence may not be available. Herein we present the steps of such a workflow, starting with an initial search for information, identification of the classes of information that such a search can return, and then guidelines by which that information is evaluated, compared, and aggregated.

Searching for Information: Patient/Population, Intervention, Comparison, and Outcome

Technology has substantively changed how information can be sought and retrieved. Online search engines such as MEDLINE via PubMed, which contains over 26 million citations, have dramatically enhanced the ability to access biomedical literature. However, there is a very real potential for such access to become overwhelming. Effective and efficient use of search engines can be enhanced by framing the clinical question in a format designed to improve the relevancy of search results. PICO is one such format, where the acronym stands for Patient/Population, Intervention, Comparison, and Outcome.

- Patient or population is the specific group of individuals for which the questions is being asked.
- Intervention is the treatment or technique of interest for the defined patient or population. Intervention might be a procedure, such as “laparoscopic appendectomy” or be defined as an exposure of interest, such as “smoking.”
- Comparison is the alternative treatment or technique to which you are comparing the intervention. Terms might include, for example, “open appendectomy” or “observation.”
- Outcome of interest is the final step of the PICO format. Examples include “mortality,” “operative time,” and “wound infection.”

As with all online search strategies, there is a trade-off between the specificity of the search and the breadth of the returned items. When using PICO to inform clinical decision-making, it is generally advantageous to be as precise and specific as possible when initiating a search: this increases the likelihood the search will return information most germane to the particular clinical scenario. This is accomplished by the use of “AND” in the framing of the search to encompass the set of questions of interest. For example, one could construct a query consisting of a particular procedure, with a particular method, with a particular outcome metric, such as “(distal pancreatectomy) AND splenectomy AND (splenic preservation) AND morbidity” to frame a PICO question.

Types of Studies

Principle 1 of EBM states that not all evidence is created equal; therefore, evaluating the evidentiary quality of the results of an online search requires classifying the returned search items by type of study. As noted earlier, acknowledging that the “gold standard” level of evidence, RCTs, are rare in the surgical literature, the application of EBM to surgery requires increased familiarity with the types of alternative studies available, with their relative strengths and weaknesses. These types are listed below:

- **Meta-analysis**: A meta-analysis is a technique to combine similarly published data in order to increase the overall statistical power compared to each study individually. The amount of interstudy heterogeneity (methods, study population, endpoints, etc.) should be limited to allow for the generation of informative conclusions. The pooling of similar studies enables researchers to generate a new statistical conclusion based on a substantially larger sample size. These approaches, though useful, have their limitations: the inclusion of inappropriate studies and the mislabeling of a meta-analysis leading to inaccurate conclusions. Attention should be directed toward this type of evidence when clinical guidelines do not exist.

- **Systematic Review**: Like meta-analyses, systematic reviews use standardized methods to search for and appraise studies in order to attempt to reduce bias. However, systematic reviews do not utilize quantitative methods to summarize the results. For this reason, systematic reviews are often not considered to provide the same strength of evidence as a meta-analysis.

- **Cross-Sectional Studies**: In a cross-sectional study, exposures and outcomes are measured at a single point in time. The prevalence of the outcome is then compared in patients who did and did not have the exposure. Multiple exposures and outcomes can be measured at the same time, which is an advantage; however, there are important limitations. One significant limitation is that a temporal relationship cannot be determined between exposure and outcome because they are measured simultaneously. These studies will often form the foundation for more definitive studies.

- **Case Control Study**: In a case-control study, cohorts are determined by the presence or absence of a particular outcome of interest. This is in contrast to a cross-sectional study where samples are determined by the presence or absence of an exposure. Once the samples have been identified based upon outcome, then possible prior exposures are identified, and the odds of those exposures are compared between cohorts.

- **Case Series**: A Case Series involves a report of a small group of patients that share specified clinical features; this generally does not include description of a control group. Case series are prevalent in the field of surgery, and some of the most famous eponymous procedures originated from case series, including the Whipple procedure and Nissen fundoplication. This type of study provides weak evidence due to issues with patient selection, biases, and confounding factors. However, the findings from a case series can be used to generate hypotheses for a randomized control trial.

- **Expert Opinion**: Expert opinions represent the lowest level of evidence and is representative of a clinician’s individual experience and anecdotes. Prior to evidence-based medicine, expert opinion was the primary means of teaching medicine and shaping the field. However, the opinions of clinicians can vary substantially leading to a wide range of potential unproven treatments for a medical issue. Thus, expert opinion should only be solicited in the complete absence of evidence in the literature.

It should also be noted that irrespective of the type of study or recommendation, there are additional factors that can contribute to bias in publication. To a great degree these are extrapolations of the sources of individual cognitive bias, but writ large across an entire community (see **Box: The History and Sources of Bias in Biomedical Literature**).
Hierarchies of Evidence

The original architects of EBM codified the notion that certain types of evidence are superior to others based on characteristics of study design, depicting this concept as a “pyramid,” with expert opinion comprising the base of the pyramid and randomized controlled trials at the peak (Fig. 51-1). Although conceptually appealing, this initial attempt to “rank” the evidence was relatively simplistic and rested on unproven assumptions that RCT were inherently superior to observational studies. While RCTs theoretically provide higher quality evidence compared to observational studies, RCTs can also have significant limitations and biases (see later section, “The Challenges of Applying EBM to Surgery”). Furthermore, translating the results from well-crafted RCTs can be challenging, where the specific restrictive criteria for executing a high-quality RCT can inherently limit its applicability to clinical scenarios not specifically noted or tested in the RCT. Therefore, one could find oneself in the situation of trying to compare an RCT on a related but clearly distinct use case with a well performed observational study that more closely approximated the clinical scenario in question. This led to the subsequent development of more refined frameworks to assess the quality of evidence in order to try and address these issues, although there is currently no consensus on a single framework. The current situation is that while many newer systems have devised ways in which studies can move up and down the pyramid, for well-designed studies, the pyramid largely remains intact.

The initial hierarchies of evidence were limited because they entangled the method of evidence collection with underlying study design. They failed to recognize principle 2 of EBM: “the pursuit of truth is best accomplished by evaluation of the totality of evidence” and the principle that “health claims be based upon systematic reviews which summarize the best available evidence.”1 The earliest hierarchies positioned systematic reviews at the top of the pyramid followed by RCTs; however, this classification failed to acknowledge that systematic reviews can summarize any type of evidence. Cohort studies, case-control studies, and even case reports can be the subject of systematic review. The importance of systematic review in EBM cannot be understated: systematic reviews are the most cited type of study, and these studies are essential for the development of clinical guidelines and influencing the direction of future studies.2 When applied in a timely manner, systematic reviews have resulted in major practice changes, for example, encouraging early postoperative enteral feeding compared to parenteral nutrition to prevent sepsis.3

Tools to Evaluate a Body of Evidence

By 2002, over 100 unique evidence rating systems existed,2 and the differences among them may be nontrivial. Depending upon the specific criteria used, the “strength” of evidence might differ widely from system to system. For example, the American Association of Orthopedic Surgeons (AAOS) published guidelines in 2009 for prevention of venous thromboembolism (VTE) in patients undergoing hip or knee surgery that conflicted with the widely used American College of Chest Physician (ACCP) guidelines, despite having access to the same data. While the ACCP considered VTE prophylaxis to be a grade 1 recommendation with level A evidence, the AAOS recommendation varied based upon risk of pulmonary embolism and bleeding, with no recommendation being greater than B and all recommendations being based upon level III evidence.4 In the following section we present a few of the most widely accepted tools for assessing the quality of evidence.

CEBM Levels of Evidence. One of the most widely adopted systems for grading evidence is the Oxford Center for Evidence Based Medicine (CEBM) Levels of Evidence. The original CEBM system was released in 2000 and was subsequently updated in 2011. Earlier systems of evidence ranking were criticized because they categorically placed randomized trials above observational studies, although observational studies and even anecdotes can occasionally give the “best” evidence in certain clinical situations. CEBM was therefore developed to not only improve the ranking of evidence but also to aid clinicians in quickly searching for the best evidence available for a given clinical question (Table 51-1). It is designed as both a tool for traditional critical appraisal as well as a pragmatic system that clinicians can use to answer clinical questions in real time. It can be used as a heuristic that clinicians and patients can utilize to answer clinical questions quickly and without resorting to preappraised sources.14 The CEBM Levels of Evidence system begins with choosing a clinical question from the first column of the table provided by the creators (see Table 51-1) (for example, “How common is the problem?”, “Does this intervention help?”, or “Is this test worthwhile?”). Therefore, each row of the CEBM Levels of Evidence represents a series of steps one should follow to find the best evidence for the question chosen. Strong evidence is likely to be found in columns to the left of the table, while weak evidence will be found in columns to the right. After completing a clinical query using the table, a final “level” of evidence is assigned on a scale from 1 to 5 based upon the types of studies found to answer the initial question (1 = highest rated evidence; 5 = lowest rated evidence). However, the levels are not intended to provide one with a definitive judgment regarding the quality of evidence. There may be cases where “lower level” evidence—for example, an observational study with a large treatment effect—provides stronger evidence than a “higher level” study, such as a systematic review with an inconclusive result.

CEBM should be thought of as a hierarchy of the likely best evidence. An advantage of CEBM is that it allows the potential of resorting to individual studies for the best evidence, while other systems generally assume that there is a systematic
### Table 51-1

<table>
<thead>
<tr>
<th>QUESTION</th>
<th>STEP 1 (LEVEL 1*)</th>
<th>STEP 2 (LEVEL 2*)</th>
<th>STEP 3 (LEVEL 3*)</th>
<th>STEP 4 (LEVEL 4*)</th>
<th>STEP 5 (LEVEL 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>How common is the problem?</td>
<td>Local and current random sample surveys (or censuses)</td>
<td>Systematic review of surveys that allow matching to local circumstances**</td>
<td>Local non-random sample**</td>
<td>Case-series**</td>
<td>n/a</td>
</tr>
<tr>
<td>Is this diagnostic or monitoring test accurate? (Diagnosis)</td>
<td>Systematic review of cross-sectional studies with consistently applied reference standard and blinding</td>
<td>Individual cross-sectional studies with consistently applied reference standard and blinding</td>
<td>Non-consecutive studies, or studies without consistently applied reference standards**</td>
<td>Case-control studies, or &quot;poor or non-independent reference standard&quot;**</td>
<td>Mechanism-based reasoning</td>
</tr>
<tr>
<td>What will happen if we do not add a therapy? (Prognosis)</td>
<td>Systematic review of inception cohort studies</td>
<td>Inception cohort studies</td>
<td>Cohort study or control arm of randomized trial*</td>
<td>Case-series or case-control studies, or poor quality prognostic cohort study**</td>
<td>n/a</td>
</tr>
<tr>
<td>Does this intervention help? (Treatment Benefits)</td>
<td>Systematic review of randomized trials or n-of-1 trials</td>
<td>Randomized trial or observational study with dramatic effect</td>
<td>Non-randomized controlled cohort/follow-up study**</td>
<td>Case-series, case-control studies, or historically controlled studies**</td>
<td>Mechanism-based reasoning</td>
</tr>
<tr>
<td>What are the COMMON harms? (Treatment Harms)</td>
<td>Systematic review of randomized trials, systematic review of nested case-control studies n-of-1 trial with the patient you are raising the question about, or observational study with dramatic effect</td>
<td>Individual randomized trial or (exceptionally) observational study with dramatic effect</td>
<td>Non-randomized controlled cohort/follow-up study (post-marketing surveillance) provided there are sufficient numbers to rule out a common harm. (For long-term harms the duration of follow-up must be sufficient.)**</td>
<td>Case-series, case-control, or historically controlled studies**</td>
<td>Mechanism-based reasoning</td>
</tr>
<tr>
<td>What are the RARE harms? (Treatment Harms)</td>
<td>Systematic review of randomized trials or n-of-1 trial</td>
<td>Randomized trial or (exceptionally) observational study with dramatic effect</td>
<td>Non-randomized controlled cohort/follow-up study**</td>
<td>Case-series, case-control, or historically controlled studies**</td>
<td>Mechanism-based reasoning</td>
</tr>
<tr>
<td>Is this (early detection) test worthwhile? (Screening)</td>
<td>Systematic review of randomized trials</td>
<td>Randomized trial</td>
<td>Non-randomized controlled cohort/follow-up study**</td>
<td>Case-series, case-control, or historically controlled studies**</td>
<td>Mechanism-based reasoning</td>
</tr>
</tbody>
</table>

*Level may be graded down on the basis of study quality, Imprecision, Indirectness (study PICO does not match questions PICO), because of inconsistency between studies, or because the absolute effect size is very small; Level may be graded up if there is a large or very large effect size.

**As always, a systematic review is generally better than an individual study.

How to cite the Levels of Evidence Table


*OCEBM Table of Evidence Working Group = Jeremy Howick, Iain Chalmers (James Lind Library), Paul Glasziou, Trish Greenhalgh, Carl Heneghan, Alessandro Liberati, Ivan Moschetti, Bob Phillips, Hazel Thornton, Olive Goddard, and Mary Hodgkinsan
Definitions of GRADE Evidence Quality

High quality – Further research is very unlikely to change confidence in the estimate of effect.

Moderate quality – Further research is likely to have an important impact confidence in the estimate of effect and may change the estimate.

Low quality – Further research is very likely to have an important impact on confidence in the estimate of effect and is likely to change the estimate.

Very low quality – Any estimate of effect is very uncertain.

review available. Additionally, other systems are built around considering the strength of evidence for therapeutic effects and harms, while CEBM allows appraisal of evidence for prevalence of disease, accuracy of diagnostic tests, prognosis, therapeutic effects, rare harms, common harms, and usefulness of screening.

Grading and Recommendations, Assessment, Development, and Evaluation. Alternatively, the Grading and Recommendations, Assessment, Development and Evaluation (GRADE) system classifies the quality of evidence into one of four levels: high, moderate, low, and very low15 (Box: Definitions of GRADE Evidence Quality). Evidence quality in the GRADE system is not assigned solely on study design. For example, a randomized controlled trial begins at “high quality,” but may be demoted due to one or more of the following: study limitations, inconsistent results, indirectness of evidence, imprecision, or reporting bias. Alternatively, observational studies (cohort or case-control studies) start as “low quality” but may be upgraded if there is a large magnitude of the treatment effect, evidence of a dose-response relationship, or if all plausible biases would decrease the magnitude of a treatment effect. Thus, the GRADE system of evaluating the quality of evidence provides more granularity than the traditional hierarchy system, which assigns quality based upon study design alone. Although the GRADE system has significant advantages, it is more complex and has a steeper learning curve than traditional systems. Finally, GRADE is intended for appraising a body of evidence, such as in a systematic review.

In addition to providing a transparent approach to grading evidence quality, the GRADE system outlines an approach to the development and assignment of strength to clinical recommendations. GRADE’s sophisticated hierarchy of evidence allows the system to protect against both superficial assessment and unwarranted confidence in all classes of study design. Since its development, the increasing use of GRADE has resulted in higher quality and rigor of systematic reviews due to standards outlined by the system.15 In creating a recommendation regarding a body of evidence, GRADE allows experts to account for limitations in bodies of evidence comprising of RCTs, while also allowing for the rating of observational studies as high quality in cases where RCTs are not feasible (i.e., an RCT cannot ethically be performed). GRADE therefore potentially allows for observational studies to provide definitive evidence of causal association (e.g., alcohol causing cirrhosis or asbestos causing mesothelioma) where RCTs may not be ethical or necessary.

One of the major advantages of GRADE is that it specifically addresses the process of moving from evidence to recommendations. The process begins with the creation of a summary of findings table. A summary of findings table consists of a presentation not only of evidence quality but also estimates of the relative and absolute effects of patient-centered outcomes (Fig. 51-2). The summary of findings format was created to minimize framing effects, where different raters may come to varied conclusions based upon identical information due to the information having a contrasting presentation in terms of gain versus loss.16 GRADE and similar EBM systems specifically takes into consideration judgement of risk versus benefit, resource use, feasibility, and equity to attempt to make decision-making as consistent as possible across a range of reviewers.2 Despite all of the aforementioned considerations when constructing a guideline, it is important to realize that patient values or preferences may immediately invalidate any recommendation. Evidence is often constructed based upon measurement of outcomes such as morbidity, mortality, or survival; however, patients may be more concerned with quality of life or avoiding invasive interventions. GRADE attempts to acknowledges this intrinsic variability within its system of grading.

In terms of the overall strength of a recommendation that GRADE can assign, two grades are possible: “strong” and “weak.” A strong recommendation is one where positive effects of an intervention clearly outweigh the negative effects or vice versa. A weak recommendation is one where the association is less clear, either because of low quality evidence or because the evidence clearly suggests that the positive and negative effects are similar. However, quality of evidence is not the only factor that affects the strength of a recommendation (Table 51-2). Factors such as uncertainty of patient values or whether an intervention is an appropriate use of resources can play a role in the strength of a recommendation as well. Therefore, it is important to note that a “strong” or “weak” recommendation may be given regardless of the classification of the evidence. For example, there is a strong recommendation that patients with Zollinger-Ellison syndrome be treated with PPI. This recommendation is made despite weak evidence to support this practice because the potential benefits far outweigh the potential risks.17

Although the systems for grading evidence are well developed, it is important to remember that the studies used for evidence are judged based on their internal validity, or the extent to which a causal conclusion is warranted based upon application of the results to the study population. This means that care must be exercised when applying a recommendation to a given patient, as the external validity of a recommendation, or generalizability of a causal conclusion to populations outside of the scope of the original studies, may not be appropriate. Therefore, all evidence must be applied within the context of the patient in front of you.

Synthesis of Evidence—Clinical Guidelines

The Institute of Medicine defines a clinical guideline as “statements that include recommendations, intended to optimize patient care, that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options.”18 Clinical guidelines may reflect previous published studies of varying design and quality, as well as expert opinion, and often represent the highest level of applied clinical evidence. Numerous guidelines have been published; however, like individual studies, even guidelines can vary in quality. The highest quality and most clinically useful guidelines tend to have the following qualities:
### UNDERSTANDING, EVALUATING, AND USING EVIDENCE FOR SURGICAL PRACTICE

#### CHAPTER 51

#### Summary of findings:

Compression stockings compared with no compression stockings for people taking long flights

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative comparative risks* (95% CI)</th>
<th>Relative effect (95% CI)</th>
<th>Number of participants (studies)</th>
<th>Quality of the evidence (GRADE)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patients or population:</strong> Anyone taking a long flight (lasting more than 6 hours)</td>
<td><strong>Settings:</strong> International air travel</td>
<td><strong>Intervention:</strong> Compression stockings¹</td>
<td><strong>Comparison:</strong> Without stockings</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Without stockings</strong></td>
<td><strong>With stockings</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Symptomatic deep vein thrombosis (DVT)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>See comment</td>
<td>See comment</td>
<td>Not estimable</td>
<td>2821 (9 studies)</td>
<td>See comment</td>
<td>0 participants developed symptomatic DVT in these studies.</td>
</tr>
<tr>
<td><strong>Symptom-less deep vein thrombosis</strong></td>
<td><strong>Low risk population</strong>²</td>
<td>10 per 1000</td>
<td>1 per 1000 (0 to 3)</td>
<td>RR 0.10 (0.04 to 0.26)</td>
<td>2637 (9 studies)</td>
</tr>
<tr>
<td><strong>High risk population</strong>²</td>
<td>30 per 1000</td>
<td>3 per 1000 (1 to 8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Superficial vein thrombosis</strong></td>
<td>13 per 1000</td>
<td>6 per 1000 (2 to 15)</td>
<td>RR 0.45 (0.18 to 1.13)</td>
<td>1804 (8 studies)</td>
<td>+++O Moderate³</td>
</tr>
<tr>
<td><strong>Oedema</strong> Post-flight values measured on a scale from 0, no oedema, to 10, maximum oedema.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>The mean oedema score ranged across control groups from 6 to 9.</td>
<td>The mean oedema score in the intervention groups was on average 4.7 lower (95% CI –4.9 to –4.5).</td>
<td>1246 (6 studies)</td>
<td>+++O Low⁴</td>
<td></td>
</tr>
<tr>
<td><strong>Pulmonary embolus</strong></td>
<td>See comment</td>
<td>See comment</td>
<td>Not estimable</td>
<td>2821 (9 studies)</td>
<td>See comment</td>
</tr>
<tr>
<td><strong>Death</strong></td>
<td>See comment</td>
<td>See comment</td>
<td>Not estimable</td>
<td>2821 (9 studies)</td>
<td>See comment</td>
</tr>
<tr>
<td><strong>Adverse effects</strong></td>
<td>See comment</td>
<td>See comment</td>
<td>Not estimable</td>
<td>1182 (4 studies)</td>
<td>See comment</td>
</tr>
</tbody>
</table>

*The basis for the **assumed risk** is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the intervention group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio GRADE: GRADE Working Group grades of evidence (see explanations)

¹All the stockings in the 9 trials included in this review were below-knee compression stockings. In four trials the compression strength was 20–30 mmHg at the ankle. It was 10–20 mmHg in the other four trials. Stockings come in different sizes. If a stocking is too tight around the knee it can prevent essential venous return causing the blood to pool around the knee. Compression stockings should be fitted properly. A stocking that is too tight could cut into the skin on a long flight and potentially cause ulceration and increased risk of DVT. Some stockings can be slightly thicker than normal leg covering and can be potentially restrictive with tight foot wear. It is a good idea to wear stockings around the house prior to travel to ensure a good, comfortable fitting. Stockings were put on 2 to 3 hours before the flight in most of the trials. The availability and cost of stockings can vary.

²Two trials recruited high risk participants defined as those with previous episodes of DVT, coagulation disorders, severe obesity, limited mobility due to bone or joint problems, neoplastic disease within the previous two years, large varicose veins or, in one of the studies, participants taller than 190 cm and heavier than 90 kg. The incidence for 7 trials that excluded high risk participants was 1.45% and the incidence for the 2 trials that recruited high-risk participants (with at least one risk factor) was 2.43%. We have rounded these off to 10 and 30 per 1000 respectively.

³The confidence interval crosses no difference and does not rule out a small increase.

⁴The measurement of oedema was not validated or blinded to the intervention. All of these studies were conducted by the same investigators.

⁵If there are very few or no events and the number of participants is large, judgement about the quality of evidence (particularly judgements about precision) may be based on the absolute effect. Here the quality rating may be considered “high” if the outcome was appropriately assessed and the event, in fact, did not occur in 2821 studied participants.

⁶None of the other trials reported adverse effects, apart from 4 cases of superficial vein thrombosis in varicose veins in the knee region that were compressed by the upper edge of the stocking in one trial.
Table 51-2
Factors that affect the strength of a recommendation

<table>
<thead>
<tr>
<th>FACTOR</th>
<th>EXAMPLES OF STRONG RECOMMENDATIONS</th>
<th>EXAMPLES OF WEAK RECOMMENDATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of evidence</td>
<td>Many high quality randomized trials have shown the benefit of inhaled steroids in asthma</td>
<td>Only case series have examined the utility of pleurodesis in pneumothorax</td>
</tr>
<tr>
<td>Uncertainty about the balance between desirable and undesirable effects</td>
<td>Aspirin in myocardial infarction reduces mortality with minimal toxicity, inconvenience, and cost</td>
<td>Warfarin in low risk patients with atrial fibrillation results in small stroke reduction but increased bleeding risk and substantial inconvenience</td>
</tr>
<tr>
<td>Uncertainty or variability in values and preferences</td>
<td>Young patients with lymphoma will invariably place a high value on the life prolonging effects of chemotherapy than on treatment toxicity</td>
<td>Older patients with lymphoma may not place a higher value on the life prolonging effects of chemotherapy than on treatment toxicity</td>
</tr>
<tr>
<td>Uncertainty about whether the intervention represents a wise use of resources</td>
<td>The low cost of aspirin as prophylaxis against stroke in patients with transient ischemic attacks</td>
<td>The high cost of clopidogrel and of combination dipyridamole and aspirin as prophylaxis against stroke in patients with transient ischemic attacks</td>
</tr>
</tbody>
</table>

1. An explicit description of development and funding processes that is publicly available.
2. A transparent process that minimizes bias, distortion, and conflicts of interest.
3. Developed by a multidisciplinary panel composed of: clinicians, methodological experts, and representatives, including a patient or consumer, of populations expected to be affected by the guideline.
4. Utilizes rigorous systematic evidence review and considers quality, quantity, and consistency of the aggregate of available evidence.
5. Summarizes evidence about potential benefits and harms relevant to each recommendation.
6. Explains the parts that values, opinion, theory, and clinical experience play in deriving recommendations.
7. Provides a rating of the level of confidence in the evidence underpinning each recommendation and a rating of the strength of each recommendation.
8. Undergoes extensive external review that includes an open period for public comment.
9. Has a mechanism for revision when new evidence becomes available.

Depending upon the clinical question, such guidelines are often interpreted as the standard of care. However, multiple clinical guidelines may be applicable with respect to various aspects of a given clinical situation and must not be followed blindly without considering specific situational issues through the lens of an experienced clinician. Moreover, guidelines do not (and probably cannot) exist for all clinical situations. Clinicians often must resort to other resources to enrich the context in which decisions are made, and, as with all evidence, care must be taken not to extrapolate the application of a clinical guideline beyond its specific conditions.

THE CHALLENGES OF APPLYING EBM TO SURGERY

As noted earlier, the application of EBM to surgery has lagged behind other fields of medicine, and this has been attributed to the difficulty in establishing a sufficient mass of evidence with the “gold standard” RCT. Here we describe the process of evaluating the quality of a RCT and note the challenges related to the execution of a high-quality RCT in a surgical context.

Analysis of a Surgical Randomized Control Trial

Sufficient knowledge of the trial’s methodological accuracy and results is essential for critical appraisal. However, less than half of journal articles adequately report the study design. This deficiency led to the development of the Consolidated Standards of Reporting Trials (CONSORT) guidelines in 1992, which was subsequently revised in 2010. These guidelines are a minimal set of recommendations for reporting RCTs (blindness, randomization, etc.) to facilitate critical appraisal. Many of the surgical journals now require completion of a CONSORT checklist prior to submission of the RCT manuscript (Fig. 51-3). Establishing this requirement has standardized the way articles are presented and analyzed. The two key aspects to focus on when assessing a RCT are internal and external validity.

Internal Validity

Determining the degree that the results of the RCT are accurate and consistent for the sample patients is called internal validity. Without internal validity, a study cannot be properly appraised, as the study was not constructed properly to answer the hypothesis without avoiding bias or confounding factors. The internal validity of a RCT requires the evaluation of several properties: randomization, blinding, equivalence among groups, completeness of follow-up, and accuracy of analysis. These properties are discussed in the following section.

Randomization. Randomization is the creation of participant groups with similar known and unknown prognostic factors to achieve the goal of eliminating selection bias. For example, if the investigator can decide which treatment the patient receives, he or she may assign a participant to a study arm that is more favorable for that specific patient. On outcomes analysis, certain groups may have an overestimated treatment effect due to patient selection and not necessarily the intervention itself. The methodology
<table>
<thead>
<tr>
<th>Section</th>
<th>Item No</th>
<th>Checklist item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title and Abstract</td>
<td>1a</td>
<td>Identification as a randomized trial in the title</td>
</tr>
<tr>
<td></td>
<td>1b</td>
<td>Structured summary of trial design, methods, results, and conclusions</td>
</tr>
<tr>
<td>Background and Objectives</td>
<td>2a</td>
<td>Scientific background and explanation of rationale</td>
</tr>
<tr>
<td></td>
<td>2b</td>
<td>Specific objectives or hypotheses</td>
</tr>
<tr>
<td>Trial Design</td>
<td>3a</td>
<td>Description of trial design (such as parallel, factorial) including allocation ratio</td>
</tr>
<tr>
<td></td>
<td>3b</td>
<td>Important changes to methods after trial commencement with reasons</td>
</tr>
<tr>
<td>Participants</td>
<td>4a</td>
<td>Eligibility criteria for participants</td>
</tr>
<tr>
<td></td>
<td>4b</td>
<td>Settings and locations where the data were collected</td>
</tr>
<tr>
<td>Interventions</td>
<td>5</td>
<td>The interventions for each group with sufficient details to allow replication, including how and when they were administered</td>
</tr>
<tr>
<td>Outcomes</td>
<td>6a</td>
<td>Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed</td>
</tr>
<tr>
<td></td>
<td>6b</td>
<td>Any changes to trial outcomes after the trial commenced, with reasons</td>
</tr>
<tr>
<td>Sample size</td>
<td>7a</td>
<td>How sample size was determined</td>
</tr>
<tr>
<td></td>
<td>7b</td>
<td>When applicable, explanation of any interim analyses and stopping guidelines</td>
</tr>
<tr>
<td>Randomization:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequence Generation</td>
<td>8a</td>
<td>Method used to generate the random allocation sequence</td>
</tr>
<tr>
<td></td>
<td>8b</td>
<td>Type of randomization; details of any restriction (such as blocking and block size)</td>
</tr>
<tr>
<td>Allocation concealment mechanism</td>
<td>9</td>
<td>Mechanism used to implement the random allocation sequence</td>
</tr>
<tr>
<td>Implementation</td>
<td>10</td>
<td>Who generated the random allocation sequence, who enrolled participants, and who assigned interventions</td>
</tr>
<tr>
<td>Blinding</td>
<td>11a</td>
<td>If done, who was blinded after assignment to interventions and how</td>
</tr>
<tr>
<td></td>
<td>11b</td>
<td>If relevant, description of the similarity of interventions</td>
</tr>
<tr>
<td>Results</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant flow</td>
<td>13a</td>
<td>For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analyzed for the primary outcome</td>
</tr>
<tr>
<td></td>
<td>13b</td>
<td>For each group, losses and exclusions after randomization, together with reasons</td>
</tr>
<tr>
<td>Recruitment</td>
<td>14a</td>
<td>Dates defining the periods of recruitment and follow-up</td>
</tr>
<tr>
<td></td>
<td>14b</td>
<td>Why the trial ended or was stopped</td>
</tr>
<tr>
<td>Baseline data</td>
<td>15</td>
<td>A table showing baseline demographic and clinical characteristics for each group</td>
</tr>
<tr>
<td>Numbers analyzed</td>
<td>16</td>
<td>For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups</td>
</tr>
<tr>
<td>Outcomes and estimation</td>
<td>17a</td>
<td>For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)</td>
</tr>
<tr>
<td></td>
<td>17b</td>
<td>For binary outcomes, presentation of both absolute and relative effect sizes is recommended</td>
</tr>
<tr>
<td>Ancillary analyses</td>
<td>18</td>
<td>Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory</td>
</tr>
<tr>
<td>Harms</td>
<td>19</td>
<td>All important harms or unintended effects in each group</td>
</tr>
<tr>
<td>Discussion</td>
<td>20</td>
<td>Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses</td>
</tr>
<tr>
<td>Generalizability</td>
<td>21</td>
<td>Generalizability (external validity, applicability) of the trial findings</td>
</tr>
<tr>
<td>Interpretation</td>
<td>22</td>
<td>Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence</td>
</tr>
<tr>
<td>Other Information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Registration</td>
<td>23</td>
<td>Registration number and name of trial registry</td>
</tr>
<tr>
<td>Protocol</td>
<td>24</td>
<td>Where the full trial protocol can be accessed, if available</td>
</tr>
<tr>
<td>Funding</td>
<td>25</td>
<td>Sources of funding and other support (such as supply of drugs), role of funders</td>
</tr>
</tbody>
</table>

of randomization should always be reported and carefully analyzed by the reader. Certain approaches of randomization called quasi-random allocation (date of birth, day of week, participant number, etc.) are not truly random and cannot be fully concealed from study personnel. Additionally, the concept of randomization eliminating bias is only theoretical. To truly ensure the probability of confounders being equally balanced between groups, a trial must be repeated indefinitely. Understanding this impracticality, we accept that randomization will suffice.

**Blinding.** Blinding aims to reduce certain biases that can affect the outcome of the study. A subject’s knowledge of the group that they were randomized can lead to a performance bias, which can influence subjective outcomes (placebo effect). Importantly, authors should be explicitly clear regarding which groups (subjects, clinicians, assessors) are blinded and avoid using non-specific phrases such as “double-blinded” or “triple-blinded.” Achieving blinding and minimizing bias is a major hurdle in the execution of surgical RCTs, where there are the ethical dilemmas surrounding “sham” or placebo surgery (though for a counter-argument, see reference no. 23). Moreover, blinding is impossible when comparing an operative versus a nonoperative intervention.

**Equivalence Among Groups.** During accrual, randomization helps to ensure that each group in the study shares equivalent baseline demographics and unmeasured biases. However, throughout the study, each group should be treated equally (excluding the actual intervention) with respect to the number of clinical visits, diagnostic tests, etc. Enforcing the same protocol to each study participant further decreases sources of bias and provides increased validity when performing final analysis.

**Completeness of Follow-Up.** Attrition bias is the differences that occur between the groups when participants withdraw from the study. A pattern can usually be identified (the treatment, side effects of treatment, long follow-up time, or other factors) that leads to withdrawal from the study. These events can hinder the ability to interpret the results of the study, and researchers should consider these implications during trial design. Furthermore, the mechanism of attrition may manifest in a bias; patients who elect to remain in a study may in fact select for characteristics that affect or determine efficacy (see the following section).

**Accuracy of Analysis.** Analyzing the results of only participants who completed all follow-up visits throughout the study can lead to skewed and inaccurate conclusions. Thus, most RCTs follow the principle of intention-to-treat (ITT) analysis. ITT analysis includes study participants who underwent initial randomization assignment regardless of events that transpired after randomization; thus ITT analysis is often described as “once randomized, always analyzed.” Removal of noncompliers from statistical analysis may overestimate the effect size of the intervention. Furthermore, in clinical practice, a portion of patients will be noncompliant, and thus ITT analysis will more accurately represent the overall population.

**External Validity**
The goal of an RCT is to show a causative relationship between an intervention and an outcome. However, to change clinical practice, the results of the RCT must be both relevant and generalizable to the clinical population; this assessment is called external validity.

**Number Needed to Treat.** The number needed to treat (NNT) is defined as the number of patients that undergo the intervention before a single patient benefits compared to the control group in the trial. It is computed as the inverse of the risk difference between two groups. The smaller the NNT, the more efficacious a treatment. For example, in an RCT comparing laparoscopic cholecystectomy to observation to prevent recurrent idiopathic acute pancreatitis, the number needed to treat was five patients. The NNT should also be weighed against the adverse effects of the intervention.

**Number Needed to Harm.** While NNT reports the number of patients who undergo the intervention before a single patient benefits, number needed to harm (NNH) describes how many patients undergo the intervention for one person to have an adverse event. The higher the NNH, the safer a treatment is. In general, interventions with a low NNT and high NNH are preferred. However, NNT and NNH should not be used in isolation when determining the appropriateness of intervention as neither number takes into account the degree of benefit to harm.

**Generalizability of Results.** RCTs have specific exclusion and inclusion criteria to recruit a study population that is homogenous with the goal of limiting sources of bias. While this method is appropriate for RCTs, the results may not directly translate to “real-world” situations with greater heterogeneity within the potential target population (see prior comment in “Hierarchies of Evidence”), leading to a potentially significant discrepancy between trial results and their implementation for day-to-day clinical decisions. In addition, RCTs often come to a conclusion that determines the best treatment for the “average” patient enrolled in the trial. However, most patients are not “average,” and therefore the proposed conclusion may not be relevant. Additional studies about the intervention of interest in more heterogeneous populations can help convince physicians to change their clinical practice; these correlate to phase 4 pharmaceutical trials and point to the importance of continued postpractice change data collection and analysis. More importantly, principle 3 of EBM, which states that “clinical decisions should be influenced by patient values and preference,” needs to be accounted for, especially with the implementation of a new practice guideline or pattern.

**Additional Challenges to Conducting a Surgical RCT**
In addition to methodological issues that might limit the reliability of a RCT, there are also considerable logistical barriers to performing a RCT. These are not trivial factors, and they contribute heavily to the number and size of RCTs that can be done, particularly in surgical populations.

**Recruitment.** One of the most challenging aspects of an RCT is recruiting an adequate number of patients to provide a high and sufficient degree of statistical power to demonstrate a measurable difference between interventions. This becomes exponentially more difficult with the prevalence of certain rare diseases. To help overcome low accrual, many trials expand their study to other hospitals and facilities at the expense of increased heterogeneity. While this may decrease internal validity, the benefit is the increase in external validity.

**Learning Curves and Expertise-Based Design.** Pharmaceutical-based RCTs normally have higher internal validity compared to surgical trials because of the effect of surgeon experience and technique affecting patient outcomes; this is especially impactful when new surgical procedures are introduced. While the administration of a drug is a straightforward process without measurable deviation, the same cannot be said
regarding surgery. Novel surgical procedures have defined learning curves even for the most experienced surgeons. During this learning process, surgeon inexperience, either in technical features of the procedure or procedure-related decision-making, can lead to adverse patient outcomes. Thus, neglecting the learning curve can lead to an underestimation of the success of the experimental intervention; conversely, accounting for the learning curve can be necessary in assessing how a new procedure can best be disseminated across the community. Furthermore, beyond the evaluation of new procedures, even with established procedures each individual surgeon is likely to have acquired throughout his/her career unique techniques and habits when operating on patients. This heterogeneity of surgeon experience and technique can limit standardization for a trial intervention.

To help solve the issue of surgeon heterogeneity and inexperience, RCTs can employ "expertise-based design." In this method, patients remain randomized to either the intervention or control, but the operating surgeons are experts in the surgery they are performing. This technique is already followed during cross-specialty RCTs, such as open gastrostomy tube placement versus interventional-radiology (IR) gastrostomy tube placement. However, this does not model everyday clinical practice because not all surgeons are considered experts in the procedure described in a particular trial.

**All-or-One Situation.** Despite continual pressure to prove treatment effect by using a RCT, there are situations when conducting a trial does not make ethical or common sense. A famous example is from the *British Medical Journal* in 2003 that questioned as to why there are no RCTs evaluating the use of parachutes during gravitational free-fall. The authors state that the evidence to support the use of parachutes is purely observational yet it is considered a "gold standard" practice. This demonstrates the concept of an all-or-none situation, where the study population exposed to a risk experiences the outcome and none of the population experiences the outcome with the intervention. Performing an RCT on this type of situation would be dangerous and unethical, and thus purely observational data can provide a high degree of sufficient evidence.

**Noninferiority Trials.** As reviewed earlier, trialing a new therapy compared to a placebo or sham raises serious ethical issues, especially when an effective therapy has already been established. Moreover, a portion of randomized control trials today are evaluating secondary endpoints, such as quality of life, safety, and cost efficiency of a new therapy compared to the existing gold standard. These studies are called noninferiority trials, with the intent to prove efficacy that is not worse than the existing therapy. For example, a 2004 study compared open versus laparoscopic colectomy for colon cancer. The aim was to show similar oncologic endpoints with improved secondary outcomes (improved cosmesis, decreased postoperative pain, decreased hernia incidence). The prevalence of these trials has increased substantially from under 100 in 2005 to nearly 600 in 2015. The most important consideration when evaluating this type of trial is the prespecified margin of noninferiority, a value that is largely arbitrary in the literature.

**USE AND MISUSE OF STATISTICAL SIGNIFICANCE**

The use of statistical methods is central to the scientific process; it is only through statistics that the problem of induction can be addressed. While this chapter is not intended to be a comprehensive description of statistical methods, understanding the appropriate application of statistical tools is critical to being able to assess the conclusions presented in the literature, and therefore we present a summary of those statistical terms that are most germane to being able to interpret a clinical study.

**Type I and Type II Errors**

By necessity, statistical testing requires declaration of a null hypothesis, usually corresponding to the "default" state (i.e., no difference or the patient is healthy). The alternative hypothesis would then negate the stated null hypothesis (i.e., there is a difference or the patient is unhealthy). The result of a statistical significance test may either reject or accept the null hypothesis, and this result can correspond with the true state (a correct decision) or not correspond with the true state (an error). Two types of error are possible (Table 51-3).

**Type I Error.** A type I error occurs when the null hypothesis is rejected but is actually true in the population. This may also be referred to as a false positive. The type I error rate, denoted by the Greek letter α (alpha), is the probability that the null hypothesis is rejected given that it is true. The error rate may also be referred to as the significance level, and often a value of 0.05, or 5%, is frequently used in the literature.

**Type II Error.** A type II error is the failure to reject the null hypothesis when the null hypothesis is false. This error may also be referred to as a false negative. The type II error rate is denoted by the Greek letter β (beta), is the probability that the null hypothesis is rejected given that it is false. The error rate may also be referred to as the significance level, and often a value of 0.05, or 5%, is frequently used in the literature.

**P Values**

The P value was an innovation most closely associated with Sir Ronald Fisher, one of the founders of modern statistics. The definition of a P value is the probability of an observed result given the assumption that the null hypothesis is true. The arbitrary value established for a result having statistical significance rather than “pure chance” is less than 1 in 20 defined as a P value less than 0.05. Put differently, the chance of making a false-positive conclusion is 5% at a P value of 0.05 (type I error). This risk of making a false-positive conclusion is called a "type I error." Importantly, the P value reported in the study is specific for that study’s patient sample and may not be generalizable to the overall population. The probability of a false positive report not actually having an association depends not

---

**Table 51-3**

**Decisions regarding the null hypothesis**

<table>
<thead>
<tr>
<th>TABLE OF ERROR TYPES</th>
<th>NULL HYPOTHESIS (H₀) IS TRUE</th>
<th>NULL HYPOTHESIS (H₀) IS FALSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision about null hypothesis (H₀)</td>
<td>Reject</td>
<td>Type I error (false positive)</td>
</tr>
<tr>
<td>Fail to reject</td>
<td>Correct inference (true negative)</td>
<td>Type II error (false negative)</td>
</tr>
</tbody>
</table>
only on the associated \( P \) value, but also the prior probability that the association is real and the statistical power of the given study.\(^{31,32}\) The basis of this is due to perpetual undersampling of all possible relationships in a given scientific domain. This will inherently lead to type I errors with respect to all clinical possibilities. Recently, statisticians have postulated that utilizing a \( P \) value of 0.05 will lead to wrong conclusions at least 30% of the time and may be even higher with underpowered studies.\(^{33}\)

The use of \( P \) values also categorizes statistical conclusions in a binary format. Should a \( P \) value of 0.049 be significant but a 0.051 not be significant? Furthermore, \( P \) values provide no insight into the effect size being measured. Simply, an intervention may be statistically significant but lack any clinical significance. Purely utilizing a \( P \) value to determine the value of research findings without assessing the effect size, confidence interval, and power of the study can be misleading.

Despite these flaws identified in \( P \) values, the frequency of their appearance in modern literature has continued to increase.\(^{34}\) Each reader should be carefully skeptical of \( P \) values and await replication with similar significance for confirmation. Fisher did not anticipate or endorse the use of the modern \( P \ < 0.05 \) criteria. Rather, he envisioned that experiments would be repeated until the investigator was sure that he or she had learned how to use the experimental intervention to get a predictable result.

**Alternative to \( P \) Values**

One potential alternative to Fisher’s approach and the limitation of \( P \) values is Bayesian statistics. The common element of Bayesian statistics is to provide a probability of a hypothesis being true by using prior knowledge or empirical data to estimate four probabilities:

1. The probability that the hypothesis is true.
2. The probability that the hypothesis is true given the observed data.
3. The probability that the alternative hypothesis is true.
4. The probability that the data would have been observed if the alternative hypothesis is true.

These parameters are used to calculate a Bayes factor, or a ratio of the likelihood probability of two competing hypotheses. One difficulty for many studies is that there can be very little reliable data that can be used to estimate these probability parameters.

It is important to remember that both \( P \) values and Bayes factors are mathematically defined entities, and many of the issues that have arisen with \( P \) values are due to how they are interpreted by scientists and clinicians. A false interpretation of a Bayes factor is just as troublesome as a false interpretation of a \( P \) value.

**External Consistency**

GRADE is one of several EBM systems that aim to evaluate evidence and create recommendations, but it is unknown how it compares with other previously established systems.

The GRADE Working Group attempted to address this question by comparing six different systems (The American College of Chest Physicians Evidence-Based Guidelines, Australian National Health and Medical Research Council Guidelines, Oxford Centre for Evidence-Based Medicine, Scottish Intercollegiate Guidelines Network, U.S. Preventive Services Task Force Recommendations, U.S. Task Force on Community Preventive Services Recommendations) on 12 criteria to assess the overall usefulness of each approach. The authors found that there was poor agreement about the sensibility of the six systems.\(^{35}\) Given that there is no agreed upon or proven gold standard, one may be concerned about the lack of external consistency among different systems. GRADE was constructed to overcome these issues; however, the system’s ability to do so has never been formally assessed.

The example of the Surviving Sepsis Campaign (SSC), an important attempt to produce guidelines to improve the care of patients with sepsis or septic shock, suggests that GRADE has not overcome these problems. The endorsement of the SSC by many influential organizations underscores its importance. Nonetheless, the SSC illustrates some of the important difficulties with grading in general and with the GRADE system (Box: Examples of Inconsistent Use of EBM).

**Examples of Inconsistent Use of EBM**

**Surviving Sepsis Campaign**

- The Surviving Sepsis Campaign recommended rapid use of intravenous antibiotics in their 2004 guidelines, which was given a grade of “E,” corresponding to a recommendation based upon level IV or V evidence, or the lowest levels possible.
- In the 2008 update, the same recommendation was given; however, it was given a grade of 1B/1D (depending on if shock was present), corresponding to a “strong” recommendation.\(^{37}\)
- Between 2004 and 2008, three additional studies were published; however, none were randomized controlled trials or came to conclusions that were different than the numerous studies that were published prior to 2004.\(^{38-40}\)

**Internal Consistency**

In 2005, the GRADE working group published a pilot study of the system which found varied levels of agreement on the quality of evidence for the outcomes in question among 17 assessors (kappa values [Box: The Kappa Coefficient] for agreement beyond chance ranged from 0 to 0.82; mean \( k = 0.27; k < 0 \) for four judgements). The authors concluded that “judgements about evidence and recommendations are complex” and stated that with discussion they could resolve most disagreements.\(^{41}\) No assessment of reliability or proof of usefulness has been presented regarding the GRADE system since these findings.\(^{42}\)

**System Issues**

The GRADE group considers the “strength” of their recommendations to reflect “the degree of confidence that the desirable effects of adherence to a recommendation outweigh the
The Kappa Coefficient

The Kappa coefficient is a statistic that measures inter-rater agreement for qualitative items. It is thought to be a more robust measure than simple percent agreement since \( \kappa \) takes into account the possibility of the agreement occurring by chance. In general, \( \kappa \) values < 0 indicate no agreement, 0 to 0.2 slight agreement, 0.21 to 0.4 fair agreement, 0.41 to 0.60 moderate agreement, 0.61 to 0.80 substantial agreement, and 0.81 to 1 as almost perfect agreement.

undesirable effects.”43 However, at the same time, the GRADE system allows the strength of a given recommendation to exist independent of the quality of evidence that underpins that recommendation. The GRADE Working Group states that “separating the judgements regarding the quality of evidence from judgements about the strength of recommendations is a critical and defining feature of this new grading system.”42 However, such a system allows for “high quality” evidence for small effects while “low quality” evidence with a strong recommendation is highly implausible except for certain obvious observations.

Finally, the touted advantage of the leveling process in determining the quality of evidence requires significant individual adjudication. A given study design begins at a level of quality and can be upgraded or downgraded based on several judgments regarding adequacy of blinding, follow-up, consistency, generalizability, and effect size. Graders are supposed to balance the level of quality using these factors, yet each is fundamentally different and cannot be simply added or subtracted, and it is therefore up to individual judgment as to how to weigh each factor.

Validity

The GRADE system is well described in a series of publications; however, none of the publications provide validation, data, or proof of the usefulness of the system. The only publication with data is mentioned earlier, which showed a low kappa for interobserver agreement.46 Based upon the systematic tenets of EBM and lack of literature-based proof for the effectiveness of GRADE, there would not be a basis for its use in creating recommendations. For example, no RCT assessing the effect of using EBM on patient outcomes has been undertaken. Therefore, EBM does not satisfy its own requirements and is, ironically, a form of systematic expert opinion. There is no data to suggest that systematic EBM approaches are superior to the decision-making capabilities of competent physicians with knowledge of the recent medical literature.

Implications of EBM

The GRADE Working Group suggests that “strong recommendations should require little debate and would be implemented in most circumstances.”42 Although most strong recommendations are likely accurate, definitive recommendations may have unintended consequences. For example, a definitive recommendation may have the effect of limiting debate or further research on a topic where the recommendation is misguided, and there are numerous examples where “strong” recommendations were later retracted. High-level EBM recommendations concluded that antibiotic prophylaxis should be used in necrotizing pancreatitis based upon multiple prospective randomized controlled trials, meta-analyses, and systematic reviews.14,44,45 These recommendations were later reversed, as additional trials showed that there was no benefit to antibiotic use in these patients.46

A valid concern regarding EBM is that established systems may lead to “strong” recommendations that are hard to challenge. This may even lead to situations where life-saving prospective studies are deemed “unethical” due to the presence of high-level, strong recommendations. As such, some groups have even issued warnings about converting practice guidelines into law.47,48

THE ALTERNATIVES TO EBM

EBM is appealing due to its ability to reduce and cope with uncertainty; however, the ability to mitigate uncertainty is not without drawbacks. The various EBM systems that exist are not always consistent in their evaluation of evidence, and even a single system may assign varying grades based on several subjective factors. Finally, the performance of EBM in improving patient care has never been validated. Therefore, while most certainly a useful tool, the limitations of EBM must be recognized to avoid blind adherence to guidelines and oversimplification of the complex clinical decision making that occurs in daily clinical care.

Although striving for certainty is understandable, it is contrary to the reality of medicine in which decisions regarding individual patients are inherently complex. In fact, as science strives for “precision” and “individualized” medicine, EBM’s focus of creating guidelines to care for the “average” patient will exist as a paradox. The best physicians function on a foundation of scientific theory expressed in a setting of practical knowledge gained in a local context, or tacit knowledge. This is how complex physiology and pathology are combined to make a specific decision for an individual patient. Therefore, although it is tempting to think that EBM makes surgery more scientific, one must remember that EBM itself is not founded in scientific principal.

So, what is the alternative? The alternative is a commonsense application of scientific principals and healthy skepticism for the ongoing use of EBM as a guideline for practice. This allows physicians to use published guidelines, applied within the context of their practice, until a grading system has definitively been shown to positively affect patient outcomes or more precise application of patient data is made possible. Recommendations certainly can be useful information; however, clinicians should also understand that there is a nuance with respect to adherence to guidelines and that much lies outside the reaches of EBM. As such, understanding that daily clinical practice involves hundreds of decisions that require varying proportions of explicit and tacit knowledge is important in devising a system where guidelines are flexible and receptive to continual feedback based upon the experiences of practicing physicians.

WHAT CAN RESEARCHERS DO TO IMPROVE THE VALIDITY OF RESEARCH FINDINGS?

Although it is impossible to know the truth with absolute certainty, researchers can take steps to ensure that the posttest probability is maximized. First, researchers can attempt to obtain better-powered evidence. Although even high-powered, low-bias meta analyses are not perfect, they do approach a theoretical “gold standard” of research, and although increasing power is important in arriving at correct conclusions, even high-powered studies can have significant biases. Additionally, obtaining large-scale evidence may not be possible for many research questions.
“You keep using that word. I do not think it means what you think it means.”
—Inigo Montoya from The Princess Bride

Crisis of Reproducibility and Medical Reversal: Implications for EBM

This chapter started by noting that the landscape of scientific knowledge is constantly evolving and that this fact impacts how we use and evaluate evidence as well. This 11th edition of Schwartz’s Principles of Surgery is being produced at a particularly volatile period in biomedical research as basic assumptions as to how scientific literature determines what constitutes “evidence” are being reassessed in a critical fashion. We believe it does a disservice to our readers if we fail to note and describe these trends, as they directly affect the basis of this chapter. The reassessment of biomedical literature and clinical trials can be loosely grouped into two distinct, but related topics: the crisis of reproducibility and the issue of medical reversal.

The Crisis of Reproducibility

Over the past decade it has become increasingly recognized that certain medical studies, held forth as index publications upon which were based either fundamental precepts of practice or to justify entire directions of drug discovery, could not be reproduced independently. This failure strikes at a fundamental assumption of science: that well performed studies with sufficient statistical significance represented generalizable knowledge that could be built upon. However, estimates of irreproducibility range from 75% to 90% based on mathematical inference, and practical investigations have shown as few as 0 in 52 observational study findings being confirmed by randomized controlled trials (RCTs). Methodological errors in study design, patient selection, or research practices have been proposed as major contributing factors in the debate over replication of scientific studies. However, despite the importance of replicating research findings, there is increasing concern that in modern research there is an intrinsic bias towards positive results in publication. Biases in study design, data collection, data analysis, or presentation of findings can lead to research findings when they do not truly exist. As bias increases, the positive predictive value (PPV) of a given finding being true decreases considerably. The overall effect of bias again depends on both the power and prestudy odds of a given study. In some fields, it may in fact be the case that research findings are simply a measure of the prevailing bias. Medical research operates in areas with low pre- and poststudy probability for true findings, meaning it may be quite common that observed effect sizes varying around the null hypothesis (what one would expect from chance alone) are simply measuring the prevailing bias of a given field.

In addition to bias, the globalization of research means that at any given time it is almost a certainty that multiple research teams are investigating the same question or topic. Despite this fact, research findings by single teams are often considered in isolation, and the first to report a finding receives significantly more attention than subsequent studies. Suppose multiple research teams are investigating a given question with the null hypothesis being that there is no difference in treatment two treatment strategies. The probability that at least one of the groups will claim a significant research finding increases, and the positive predictive value decreases as the number of research teams increases. Unfortunately, there is little way to control for this phenomenon other than increasing the power of each individual study.

Due to the combination of the aforementioned factors, the current framework of research means it is quite difficult to end up with a PPV >50%. Based on mathematical principles, even a well-constructed, adequately powered RCT with a pretest probability of 50% will arrive at a true conclusion only about 85% of the time. These findings limit the available literature upon which evidence-based medicine (EBM) relies and place a greater burden on practitioners when they are attempting to analyze and draw conclusions from what they find.

Medical Reversal

A related topic that directly impacts how EBM is carried out is that of medical reversal. This term was introduced by Vinay Prasad and Adam Cifu in 2011 to describe the process and pitfalls by which a previously established practice or drug falls out of favor because it is subsequently identified not to work. As such, the issue of medical reversal is impacted by the decision for a particular therapy to become adopted in the first place (ostensibly based on the principles of EBM) and the barriers to how subsequent evidence (either acquired through studies, or, more importantly, upon a more critical reassessment of the basis of its initial adoption) can reverse a prior recommendation. The set of intersecting issues related to medical reversal are highly complex (interested readers are encouraged to delve into the growing list of reports on this topic), but in terms of EBM, central issues addressed in medical reversal pertain to the use of surrogate endpoints in clinical trials, the presentation/misrepresentation of clinical trial effects, the effect of bias (academic and economic) in trial reporting and dissemination, and the strength and reliability of alternatives to RCTs (for all their flaws). As with the crisis of reproducibility, understanding the factors of medical reversal directly impacts what is appropriately considered “evidence” when executing EBM, placing greater responsibility on the surgical practitioner when determining what is appropriate or optimal care.

It should come as no surprise to the attentive reader that many of the issues related to the crisis of reproducibility and medical reversal refer back to the sources of bias and potentially perverse incentives originally noted by Francis Bacon back in 1620 (Box: The History and Sources of Bias in Biomedical Literature).
Second, as was noted previously, multiple teams often simultaneously address a given research question, and it is not proper to focus on any one study in isolation. Instead, clinicians should focus on the body of evidence in its entirety. A potential solution would be connecting groups through networking of data. This would allow for more accurate analysis and drawing of conclusions, although it would require a significant change in the culture of academic research practices.

Today, clinicians rely on the statistics provided in a scientific study to provide a summary of the results. We place trust and confidence that the paper’s biostatistician accurately and truthfully calculated these statistics without incorporating conscious bias. Each article should completely answer four questions regarding the results of the study:

1. What is the statistical significance of the results?
2. What is the effect size and is this clinically relevant?
3. What is the confidence interval?
4. What is the underlying power of the study to detect a meaningful difference?

Significant progress has been made since the adoption of EBM; however, the current direction of EBM-based guidelines have focused on populations as opposed to the complex, nuanced interactions that occur on a case by case basis. Algorithmic protocols actually serve to steer the focus away from an individual patient, at times leading to a disconnect between patients and physicians when physicians propose treatment based upon guidelines that do not adhere to that patient’s goals and values. So what can surgeons do to combat this, and how should they practice? One must ask: “What is the best course of action for this patient, in these circumstances, at this point in their illness or condition?” Therefore, evidence must be synthesized and then individualized for each patient encounter by interconnecting it with the ethics, personality, and values associated with the case at hand. Tools such as risk calculators are useful in informing discussion, but they should by no means be definitive evidence to recommend for or against a particular treatment. Judgment remains necessary in the practice of medicine, and therefore guidelines should be thought of as “rules of thumb” that require context as opposed to “rules of law.”

REFERENCES

Entries highlighted in bright blue are key references.


Ambulatory Surgery
Marcus Adair, Stephen Markowiak, Hollis Merrick, James R. Macho, Kara Richardson, Moriah Muscaro, Munier Nazzal, and F. Charles Brunicardi

INTRODUCTION

Ambulatory Surgery

Ambulatory Surgery is a multidisciplinary field in which surgical procedures are performed on patients who are not expected to be admitted to the hospital. The field includes procedures performed on patients in the setting of hospital outpatient departments (HOPDs), freestanding ambulatory surgery centers (ASCs), and those performed in doctor’s offices. The word ambulatory comes from the Latin ambulare, which means “to walk,” indicating that the patients arrived at the procedure on their own and departed after the procedure to their home environment.1

Improved anesthesia techniques, the development of minimally invasive procedures, and changes to healthcare policy (particularly healthcare funding) have been the driving factors behind the increase in ambulatory surgery. Prior to these advances, almost all surgery was performed in an inpatient hospital setting. Any outpatient surgeries were minor, performed in physicians’ offices, and paid for by Medicare and insurers as part of the physician’s office visit reimbursement.2

Since the early 1980s, the volume of ambulatory surgery has increased in the United States.2,3 Between 1981 and 2005, the number of outpatient surgeries nationwide grew almost 10-fold to over 32.0 million per year. Outpatient procedures grew from 19% to 60% of all surgical procedures, by volume, in the United States from 1981 to 2011.4,5 Strong financial incentives exist for hospitals to shift some surgeries to an outpatient setting. The number of Medicare-certified ASCs has also increased steadily, from fewer than 300 in the early 1980s2 to 5532 in 2016 (Fig. 52-1).5

By definition, procedures at ASCs and physician’s offices are performed without the full resources of a hospital. The press has sensationalized a few adverse patient outcomes in these settings and made claims about the overall safety based on these isolated events. However, much of the convenience, high patient satisfaction rates, and cost-efficiency of ambulatory surgery is lost when performed in a hospital setting. Thus, the challenge underlying ambulatory surgery is performing safe operations on carefully selected patients in a manner that is patient-family-centric and economical.

A majority of large hospitals also have their own outpatient surgery departments that exist within the hospital. ASCs perform procedures faster and more efficiently than HOPDs and at higher volume than both HOPDs and physicians’ offices. Additionally, the number of certified, freestanding ASCs nationwide has eclipsed the total number of hospitals by more than 1000 centers.6 For these reasons, a plurality of outpatient surgical volume in the United States is now performed in ASCs. It is critical for the modern surgeon to have a grasp of the unique clinical challenges and economic impacts of ambulatory surgery.

Ambulatory Surgery Centers

ASCs are independent healthcare that offer patients the convenience of having surgery performed safely without admission to a hospital. ASCs provide only elective surgical services rather than emergency care. According to the Centers for Medicare & Medicaid Services (CMS), effective May 18, 2009, “ASCs are any distinct entity that operates exclusively for the purpose of providing surgical services to patients not requiring hospitalization and which the expected duration of services would not exceed 24 hours following an admission.”7 Ambulatory surgery centers should not be confused with office-based surgery practices or with other outpatient centers that provide diagnostic services or primary healthcare, such as urgent care centers, community health centers, mobile diagnostic units, or rural health clinics. ASCs are distinguished from these other healthcare facilities by (a) their use of a referral system for accepting patients and (b) their maintenance of a dedicated operating room. The first feature means that any patient who wants to be treated in an ambulatory surgery center must first consult a

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>2153</td>
</tr>
<tr>
<td>Ambulatory Surgery Centers</td>
<td>2153</td>
</tr>
<tr>
<td>Aspects Leading to Increased Utilization</td>
<td>2154</td>
</tr>
<tr>
<td>History of Ambulatory Surgery and Ambulatory Surgery Centers</td>
<td>2154</td>
</tr>
<tr>
<td>Procedures Performed</td>
<td>2156</td>
</tr>
<tr>
<td>Benefits of Ambulatory Surgery Centers</td>
<td>2156</td>
</tr>
<tr>
<td>Factors Contributing to Popularization of ASCs</td>
<td>2157</td>
</tr>
<tr>
<td>Regulation, Costs, and Quality</td>
<td>2158</td>
</tr>
<tr>
<td>Regulation</td>
<td>2158</td>
</tr>
<tr>
<td>Costs</td>
<td>2158</td>
</tr>
<tr>
<td>Quality</td>
<td>2159</td>
</tr>
<tr>
<td>Ownership</td>
<td>2159</td>
</tr>
<tr>
<td>Potential for Conflict of Interest</td>
<td>2159</td>
</tr>
<tr>
<td>Challenges</td>
<td>2159</td>
</tr>
<tr>
<td>Reimbursement</td>
<td>2159</td>
</tr>
<tr>
<td>Patient Selection</td>
<td>2159</td>
</tr>
<tr>
<td>ASCs vs. Hospital Outpatient Departments vs. Office-Based Surgical Suites</td>
<td>2160</td>
</tr>
<tr>
<td>Aging Population</td>
<td>2160</td>
</tr>
<tr>
<td>Conclusion</td>
<td>2160</td>
</tr>
</tbody>
</table>
Key Points

1. Define ambulatory surgery and the unique aspects of ambulatory surgery centers (ASCs).
2. Understand the history of ambulatory surgery and ASCs in the United States.
3. Review the most common procedures performed in ASCs compared to those performed in a hospital setting.
4. Discuss the financial benefits for patients and physicians within the structure of ASCs as they pertain to physician ownership, healthcare systems partnerships, and insurance reimbursement.
5. Understand the regulatory and accreditation processes that affect ASCs.
6. Predict how anticipated medical advances, technological development, and an aging population will affect the field of ambulatory surgery.

Aspects Leading to Increased Utilization

Technological Developments. Improvements in anesthesia have facilitated the safe practice of outpatient surgery by the use of new medications, improved techniques in regional anesthesia, and better management of postoperative pain. The development of minimally invasive surgical techniques such as fiberoptic endoscopy, arthroscopy, ophthalmologic procedures, and laparoscopic and robotic surgery have made it possible for patients to be discharged the on the same day as the surgery.

Reduced Cost for Patients Without Compromise in Quality of Care. Ambulatory surgery facilities are highly specialized centers originating from a service model rather than the tradition hospital model. This approach allows for streamlined processes and reduced costs. Staffing is the largest cost for most healthcare facilities including ASCs, thus same-day surgery eliminates the need for overnight nursing and support staff.

ASCs Offer Reduced Cost for Healthcare Systems. A review of commercial medical claims data found that annual U.S. healthcare costs are reduced by approximately $3.8 billion due to the availability of ASCs. Patients save more than $1.5 billion due to lower deductibles and coinsurance payments. Over the next decade, ASCs are expected to save the U.S. healthcare system between $32.5 and $57.6 billion. This cost reduction is driven by the fact that, in general, ASC prices are significantly lower than HOPD prices for the same procedure in all markets, regardless of payer. Table 52-1 displays the cost savings compared to hospital outpatient departments for the most commonly performed procedures at ASCs nationwide. While most hospitals offer outpatient surgery, ambulatory surgery centers are regarded as a superior choice for certain procedures because of facility efficiencies and price regulation under the outpatient prospective payment system.

HISTORY OF AMBULATORY SURGERY AND AMBULATORY SURGERY CENTERS

Ambulatory surgical practice traces its history from the early work of itinerant dental surgeons who traveled their circuits by horseback and trains. They frequently operated in hotel rooms and then moved on. In 1909, James Nicoll (Fig. 52-2), a pediatric surgeon in Scotland, wrote of his experiences with ambulatory anesthesia and surgery on nearly 9000 children as outpatients during a 10-year interval at Glasgow Royal Hospital for Sick Children. Operations included cleft lip and palate repair, correction of pyloric stenosis, mastoidectomy, repair of inguinal and umbilical hernias, and management of spina bifida and depressed skull fractures. Nicoll pleaded with his fellow surgeons to perform more pediatric operations on an outpatient basis, stating that “a large number of the cases at present treated in-door constitutes a waste of the resources of a children’s hospital . . . . The results obtained in the out-patient department at a tithe [small part] of the cost are equally good.”

Ralph Waters (Fig. 52-3) was a pioneer in the field of ambulatory surgery. He developed an office-based practice in Sioux City, Iowa in 1919. Waters used nitrous oxide, morphine, and scopolamine. He believed medical conditions had to be well controlled prior to surgery and that certain medical conditions precluded outpatient care. In these ways, Waters’ clinic became the prototype for the modern free-standing ASC. Waters subsequently went on to establish the first academic residency program for training anesthesiologists at the University of Wisconsin.

Prior to the advent of freestanding ASCs, the concept of ambulatory surgery first needed to gain acceptance in the form of HOPDs. In 1959, Eric Webb and Horace Graves advocated outpatient surgery because of a shortage of hospital beds in Vancouver. The first HOPD in the United States was established in 1962 at the University of California, Los Angeles by David Cohen and John Dillon, who also sought to address a shortage of hospital beds. These efforts proved to be safe and cost-effective.
Table 52-1

Comparison of top 10 procedures performed at ASCs vs. hospitals nationwide

<table>
<thead>
<tr>
<th>TOP 10 PROCEDURES PERFORMED AT AMBULATORY SURGICAL CENTERS BY VOLUME AND CPT CODE</th>
<th>NUMBER PERFORMED</th>
<th>AVG PAY PER CLAIM</th>
<th>SAVINGS AT ASC</th>
<th>TOP 10 PROCEDURES PERFORMED AT HOSPITAL OUTPATIENT DEPARTMENTS (HOPD) BY VOLUME AND CPT CODE</th>
<th>NUMBER PERFORMED</th>
<th>AVG PAY PER CLAIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cataract surgery with intraocular lens (66984)</td>
<td>1155, 283</td>
<td>$959</td>
<td>$219</td>
<td>Subcutaneous tissue debridement (11042)</td>
<td>841,517</td>
<td>$1319</td>
</tr>
<tr>
<td>Esophagogastroduodenoscopy with biopsy (43239)</td>
<td>524,082</td>
<td>$301</td>
<td>$110</td>
<td>Esophagogastroduodenoscopy with biopsy (43239)</td>
<td>628,900</td>
<td>$541</td>
</tr>
<tr>
<td>Colonoscopy and biopsy (45380)</td>
<td>416,218</td>
<td>$352</td>
<td>$172</td>
<td>Aspiration/injection of joint (20610)</td>
<td>578,407</td>
<td>$141</td>
</tr>
<tr>
<td>Colonoscopy with lesion removal (45385)</td>
<td>331,565</td>
<td>$401</td>
<td>$20</td>
<td>Cataract surgery with IOL implant (66984)</td>
<td>512,191</td>
<td>$1,178</td>
</tr>
<tr>
<td>Spine epidural injection foraminal (64483)</td>
<td>282,962</td>
<td>$335</td>
<td></td>
<td>Colonoscopy and biopsy (45380)</td>
<td>472,886</td>
<td>$524</td>
</tr>
<tr>
<td>Postlaser cataract surgery capsulotomy (66821)</td>
<td>275,760</td>
<td>$227</td>
<td></td>
<td>Colonoscopy with lesion removal (45385)</td>
<td>350,001</td>
<td>$421</td>
</tr>
<tr>
<td>Spine epidural injection lumbar, sacral (62311)</td>
<td>210,159</td>
<td>$358</td>
<td>$120</td>
<td>Spine epidural injection lumbar, sacral (62311)</td>
<td>326,956</td>
<td>$478</td>
</tr>
<tr>
<td>Injection, paravertebral facet joint (64493)</td>
<td>174,450</td>
<td>$306</td>
<td></td>
<td>Insertion of temporary bladder catheter (51702)</td>
<td>308,614</td>
<td>$69</td>
</tr>
<tr>
<td>Diagnostic colonoscopy (45378)</td>
<td>157,951</td>
<td>$401</td>
<td>$100</td>
<td>Appl. of multilayer compression system (29581)</td>
<td>303,026</td>
<td>$97</td>
</tr>
<tr>
<td>Colorectal screening, high-risk individual (G0105)</td>
<td>128,181</td>
<td>$333</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The first truly freestanding ASC, “SurgiCenter,” was opened in Phoenix, Arizona by Wallace A. Reed and John L. Ford in 1970 (Fig. 52-4). Reed and Ford were committed to providing timely, convenient, and comfortable surgical services to patients in their community, and therefore, avoiding more impersonal venues like regular hospitals. Prior to opening SurgiCenter, Reed and Ford were frustrated with having patients wait 6 weeks or more to get elective surgery, and in some cases, found surgeries canceled because the rooms booked were needed for emergencies.\(^9\)
In its first 6 months, Surgicenter performed more than 1200 operations by 153 surgeons. No adverse cardiac events occurred during the procedures, and only one patient was hospitalized following surgery (due to poorly controlled diabetes, not because of the operation itself). Surgicenter’s incredible success resulted in more than 400 visitors touring the facility in the first year in order to learn about the new model for patient care. Reed and Ford realized that a tremendous need existed for freestanding, independent ASCs. According to Reed, this propelled the formation of what eventually became the Ambulatory Surgery Center Association (ASCA), a major credentialing body within the field of ambulatory surgery.

The Orkand Report of 1976, a U.S. government-sponsored study of outpatient surgery, concluded that ambulatory surgical facilities can significantly reduce costs while maintaining the same high quality of surgical and anesthetic care achieved in hospitals.

The Society for Ambulatory Anesthesia (SAMBA) was established in 1984 to further the development of ambulatory anesthesiology as a subspecialty. The field continued to advance with the publication of Wetchler’s Anesthesia for Ambulatory Surgery in 1985, the introduction of the journal Ambulatory Surgery in 1993, and with the first state requiring accreditation for all outpatient facilities (California in 1996). Many other states have since adopted these high standards and require accreditation of ASCs. CMS now requires certification for all ASCs (Fig. 52-5 and 52-6).

**PROCEDURES PERFORMED**

By 1982, CMS had approved payments to ASCs for more than 200 procedures. Steady growth in the number of ASCs (Fig. 52-7) and the number of surgical procedures performed in the outpatient setting, including HOPDs, has continued since. Each year physicians perform more than 23 million procedures in ASCs. This shift toward outpatient procedures has increased due to advancements in medical practice and technology that have reduced the need for overnight hospital stays. Most patients, except those with complicated health conditions, can be served in the outpatient setting. Common ASC procedures include colonoscopies, cataract surgeries, tonsillectomies, and arthroscopic orthopedic surgeries. CMS currently approves and reimburses more than 3500 procedures in the ASC setting. New developments continue to expand the scope of ASCs.

ASCs may perform surgeries in several specialties or dedicate their services to one specialty, such as eye care or sports medicine. The procedure must not pose a significant safety risk and not require an overnight stay when performed in an ASC. The types of surgical procedures performed in ASCs have undergone significant changes in recent years. Many of the early ASCs were outpatient centers for plastic surgery. Advances in minimally invasive surgical techniques in other specialties, however, led to the establishment of ASCs for orthopedic, dental, and ophthalmologic procedures. See Fig. 52-8 for a recent analysis of specialty services provided by ASCs nationwide.

**BENEFITS OF AMBULATORY SURGERY CENTERS**

Since their founding over 40 years ago, ASCs have grown exponentially. These distinct entities have provided physicians an avenue to provide specialized, efficient, and quality
Figure 52-7. As of June 2017, California has 794 ASCs, making it the leading state in terms of number of ASCs. It is followed by Florida with 417 ASCs and Texas with 366 ASCs. Vermont and the U.S. Virgin Islands have the lowest number of ASCs with one each.13

Care to patients who need surgical procedures. Patient satisfaction with same-day surgery has remained relatively high since ASCs started in 1970. It is important to recognize that patients undergoing generally nonemergent surgery that does not require a hospital stay are relatively satisfied overall. Historically, the field of ambulatory surgery has been associated with very high patient satisfaction.14-15 In the future, the CMS Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey will report on nationwide patient satisfaction with ASCs and HOPDs.16

Factors Contributing to Popularization of ASCs

Cost. In many cases an outpatient procedure performed in an ASC is between one-half to one-third the cost as the same procedure performed in a hospital. In large part, ASCs affect cost savings by eliminating overnight hospitalizations and emergency procedures. ASCs perform fewer extensive diagnostic tests and dispense fewer medications. These facilities are not staffed around the clock and are not encumbered by the need for expensive and highly specialized equipment as are hospitals. For example, the Medicare Payment Advisory Commission found that a cataract operation cost only $942 at an ambulatory surgery center in 2001 as opposed to $1334 at a hospital. Figures for an endoscopy and biopsy of the upper digestive tract were $429 and $359; for a diagnostic colonoscopy, $429 and $401; and for epidural anesthesia, $320 and $183, respectively5,7 (see Table 52-1).

Organization, staffing, and specialization may play a large role in the cost differences between ASCs and HOPDs. On average, patients who were treated in ASCs spend 31.8 fewer minutes undergoing procedures than patients who were treated in HOPDs. ASCs could generate savings of $363 to $1000 per outpatient case.4

Comfort. Whereas most hospitals keep patients recovering from a surgical procedure in separate rooms, in an ASC the patient usually can spend the recovery period after surgery with their loved ones. Limiting the number of delays and disruption of emergent cases allows the surgeon to spend more time with the patient in the preoperative and postoperative areas.17-19

Convenience. Because ASCs usually schedule routine cases lasting no longer than 2 hours (average 30–45 minutes), and handle no emergency cases, scheduling is typically accurate. By avoiding the logjam, ASCs reduce the waiting time for elective procedures. A study by Hair et al reviewing Medicare patients again showed freestanding ASCs performed surgeries in less time than hospital-based ASCs overall and for procedures on various anatomic systems that resulted in reduced total time spent in facility with earlier discharge.19 These results corroborate the notion that freestanding ASCs tend to be more efficient than HOPDs.20 One possible advantage for patients would be that they are able to leave an ASC relatively quickly after their surgery, resulting in less time away from work and family. This may be particularly true for pediatric patients or parents.

Efficiency. This advantage is particularly important to surgeons. It takes much less time to prepare an operating room in a specialized ASC for the next patient than in a standard hospital. Improved efficiency allows the surgeon to treat more
patients in the same amount of time than he or she would be able to do in a hospital; some surgeons maintain that they can do three times the number of procedures in an ASC as they could in a hospital setting. Many doctors prefer working in an ASC because they can set the standards for staffing, safety precautions, and postoperative care, rather than having these things decided for them by a hospital manager.\(^1\) Trentman and coauthors discuss several factors that affect patient flow and could result in differences in preoperative and recovery times for outpatient procedures between ASCs and hospitals.\(^2,3,5,19\) For example, compared to the situation in hospitals, in ASCs surgeons are more likely to be assigned to a single operating room for all cases, which reduces delays; the operating room often is closer to the preoperative and recovery rooms because facilities are smaller; teams of staff have clearer and more consistent roles, with less personnel turnover; and staffing is not done by shifts—that is, staff members go home only after all cases are finished, which creates incentives to work quickly. In addition, hospitals may be more likely to have emergency add-on and bring-back cases for more complex cases that compete with surgeries getting moved or cancelled due to emergencies. ASCs enable patients to go home on the same day, therefore spending less time with staff in postoperative recovery rooms.

**REGULATION, COSTS, AND QUALITY**

**Regulation**

Healthcare facilities in the United States are highly regulated by federal and state entities. ASCs are included in this oversight, with both federal and state laws and regulations governing all aspects of them. Independent observers evaluate the safety and quality of care provided in ASCs through three processes: Medicare certification, state licensure, and voluntary accreditation.

To obtain Medicare certification, ASCs must meet the Medicare certification requirements, known as the Conditions for Coverage. These conditions include specifying standards for administration of anesthesia, quality evaluation, operating and recovery rooms, medical staff, nursing services, and other aspects of care. An ASC must have an inspection conducted by a state official or a representative of an organization that the government has authorized to conduct that inspection. These inspectors visit the ASC to verify that it meets established standards. Each state determines the specific requirements ASCs must meet for licensure. An ASC does not have to be certified by Medicare in order to be accredited by JCAHO; however, most ASCs provide care to Medicare beneficiaries, so it is important to meet their requirements in order to be reimbursed appropriately. Medicare inspection and certification of ambulatory surgery centers is a separate process from professional accreditation.

To obtain state licensure, many states have independent rules and regulations as well as associated fees. These third-party bodies can include Accreditation Association for Ambulatory Healthcare (AAAHC), American Association for Accreditation of Ambulatory Surgery Facilities (AAASFP), and the Joint Commission on Accreditation of Healthcare Organizations (JCAHO). In states for which accreditation is not mandated, ASCs may undergo voluntary accreditation through these same associations. In 37 states, any party looking to open an ASC must demonstrate a need for the ASC to exist.\(^21\) State licensure requirements generally exist for both healthcare facilities and healthcare professionals.\(^14\)

There was previously controversy as to whether accreditation status affects safety outcomes. In multivariate analyses that controlled for facility volume and patient characteristics, patients at Joint Commission–accredited facilities were still significantly less likely to be hospitalized after colonoscopy. Specifically, compared with patients treated in nonaccredited ASCs regulated by the state agency, patients treated at accredited facilities were less likely to be hospitalized within 7 to 30 days after surgery.\(^22\) All accredited ASCs must meet specific standards that are evaluated during on-site inspections. Patients who visit accredited ASCs can be assured that those medical facilities have rigorous checkpoints to ensure high standards.

The Ambulatory Surgery Center Association (ASCA) or Association of Ambulatory Surgery Centers (AAASC)—merged. The ASCA serves as the national membership organization as well as the advocacy group for ASCs.\(^2,21\) The ASCA works with legislative and regulatory bodies, liaises with other organizations to improve access, reduce the costs of healthcare, encouraging insurance coverage of outpatient procedures, and works to establish standards for ASCs. ASCA requires all of its ASC Association has an integral role to ensure top-quality healthcare from the nation’s ASCs. The ASCA was established when the two leading national ASC associations—Federated Ambulatory Surgery Association (FASA) and the American Association for Accreditation of Ambulatory Surgery Centers (AAASC)—merged. The ASCA was instrumental in forming the accrediting body that is now the largest accreditor of ambulatory surgery centers in the country.\(^22\)

**Costs**

Today, more than 5300 Medicare-certified ASCs offer similar services compared to those performed at hospitals, and so at a more efficient rate with lower costs. ASCs are able to accomplish this by decreasing administrative and overhead expenses.\(^5\) Expenses for an ASC include staff wages, insurance, utilities, rent, janitorial services, as well as resources necessary to handle patient records, including technology systems. ASCs are able to schedule procedures without the risk of surgeries getting moved or cancelled due to emergencies. Additionally, ASCs enable patients to go home on the same day, therefore spending less time with staff in postoperative recovery rooms.

Starting in 1982, Medicare has covered surgical procedures provided in ASCs.\(^2,21\) There are two primary elements of total cost in a surgical procedure: the cost of physicians’ professional services and cost of the facility. Typically, providers bill for professional service separately, whereas facility costs are paid to the ASC. Currently Medicare provides separate payments for 3500 surgical procedures under the ASC payment system.\(^3\) The payment system is maintained by CMS, which adjusts fees annually to maintain budget neutrality. Through the Medicare Prescription Drug, Improvement, and Modernization Act of 2003, payments to ASCs are based on the Outpatient Prospective Payment System (OPPS) and capped at no more than 59% of what is paid to hospitals for the same service.\(^3\)

During the period of 2008 to 2011, the growth in number of ASCs plateaued. The system saved an estimated $7.5 billion for the Medicare program and its beneficiaries.\(^5\) However,
growth in procedure volume during this time was greater in ASCs than in hospitals. This suggests that physicians and patients still preferred using ASCs, possibly due to advantages in cost, convenience, comfort, or the inability for hospitals to meet demand for outpatient surgeries.

**Quality**
The majority of ambulatory surgery centers ensure safety through maintaining certification and licensing of the facilities and associated staff. There are strict regulations in place at both the state and federal level that ensure quality in ASCs, which have been detailed earlier. ASCs are held to the same high standard of care as all U.S. medical facilities. They have comparable rates of use of perioperative antibiotics, patient falls, wrong-site surgery, and use of safety checklists. Nationwide and among all procedures, from 2014 to 2016 fewer than 2% of all visits to ASCs resulted in an unplanned hospital visit within 7 days. Of these hospital visits, 1.6% were emergency department or observation stays, and 0.6% were unplanned inpatient admissions. CMS is working to develop a new quality measure that will track unplanned hospitalizations after care at an ASC. This would help create increasing data on the quality of ASCs. Due to the isolated nature of ASCs, providers are often unaware when a patient goes to the emergency department or is admitted to an unaffiliated hospital following a procedure. Thus, such a measure from the CMS will help educate providers and allow for continued quality improvement among ASCs.

Looking at the general surgical population, cholecystectomy represents a key procedure to track because it is frequently performed, requires technical skill, and can result in serious complications. Provided appropriate patient selection, outpatient cholecystectomy has been demonstrated to be safely performed in ASCs. ASCs charge significantly less for performing this procedure after controlling for the variety of indications (median of $6028 for ASCs compared to $10,876 for HOPDs).

**Ownership**
In 2017, 90% of ASCs have partial or complete physician ownership, while 25% to 30% of ASCs are at least partially owned by hospitals, as compared to HOPDs, which are by definition owned and operated exclusively by the hospital. A trend is developing for hospitals to purchase ASC ownership stakes due to their cost efficiency and as a means of diversifying revenue streams. Physicians have been a driving force in the development of ASCs through their ownership and building of new facilities. Ownership of an ASC provides many clear advantages to the physician because of the increased control and autonomy over their practice. Some advantages include ease of scheduling, shorter waiting times to get patients in for elective surgery, and the ability to hire specially trained and highly skilled staff. Physicians can also avoid the bureaucracies of a hospital, including having elective surgeries cancelled to make room for emergent surgeries and delays due to more complicated surgeries. They can also ensure that the facility has specialized equipment for their particular specialty and can design the facility to meet their specific needs. Furthermore, physicians who perform surgeries in their own ASCs receive a share of the ASC’s facility payment in addition to payment for their professional services. This could present a conflict of interest in terms of referring patients and lowering the threshold for surgery, as discussed further in the next section.

**Potential for Conflict of Interest**
Approximately 90% of ASCs nationwide have at least some physician ownership stake. Many are joint ventures between hospitals and physicians. Increasing investment in these centers may be explained in a number of ways, including an attempt by providers to assert greater control over their professional lives, such as by having greater authority in scheduling surgeries and in purchasing equipment. Alternatively, this investment trend may be explained by declining reimbursements for physician services and rising practice costs. These economic pressures have intensified providers’ interest in nontraditional revenue sources, such as ASC investment, as a means of generating income.

Ownership entitles physicians to collect a share of the facility’s profits from referrals, in addition to their professional fees. One potential conflict of interest is that physician-owners might lower their thresholds for intervention, exposing the patient and healthcare system to the harm and cost associated with unnecessary treatment. After differences between patients and healthcare markets are adjusted for, physicians with ownership in an ASC have been found to perform a higher number of procedures compared with nonowner physicians at the same facilities. It has been noted that the increase in outpatient surgery at ASCs was more than double the decline in similar procedures performed in the hospital setting.

The other potential conflict in physician ASC ownership is in regard to patient referral. There is some evidence that physicians with an ownership stake may refer well-insured patients to their own facilities while referring Medicare and Medicaid patients to hospital outpatient clinics.

Regardless of the reason for increasing investment in and utilization of ASCs, it is important to note that ASC ownership creates a potential conflict of interest for physicians. In the United States, physician financial interests are heavily regulated by the Stark Law. This can be either a financial investment, employment, or compensation agreement. To avoid conflicts of interest and potential legal violations, physicians should consult with attorneys and advisors knowledgeable in medical law and ethics prior to entering a financial relationship.

**CHALLENGES**

**Reimbursement**
Savings from use of ASCs are primarily to the patient and healthcare system. Reimbursement to ASCs for services provided is, as a direct result, lower than reimbursement paid to hospitals for the same procedures. It is up to ASCs, which operate as private corporations, to find profit and viability by keeping their costs low through efficiency and staffing. Some steps suggested to improve the financial viability of maintaining an ASC include increasing the variety of procedures offered and grouping surgeries in such a way as to maximize staffing usage.

**Patient Selection**
Safe use of ASCs is based on identification of patients who are unlikely to require admission to a hospital after their procedure. Thus far, ASCs do very well in this aspect. Unplanned admissions after ambulatory surgery occur in approximately 0.5% to 2.0% of cases. In the future, ASCs will be challenged to reduce this unplanned admission rate even further.

Patients with Medicaid insurance, lower median household income, and a greater preoperative comorbidity burden have the
highest odds of unplanned acute care use. These patients may benefit from interventions that enhance and streamline postoperative follow-up. Additionally, the potential costs associated with postoperative acute care following procedures performed in ASCs are not insignificant. Patient-specific predictors of unplanned hospital admission include age 65 years or older, anticipated operating time longer than 120 minutes, cardiac comorbidities, peripheral vascular disease, cerebrovascular disease, malignancy, positive for human immunodeficiency virus (HIV), and regional or general anesthesia use.

The strongest predictor for unplanned inpatient hospital admission was the individual patient’s own history of previous hospitalizations, particularly among older adults. African American and Hispanic individuals also have had a markedly elevated risk of inpatient hospital admission, possibly related to cultural or socioeconomic issues of access to care. These measures may provide a valuable target for quality improvement, cost improvement, and innovation.

ASCs vs. Hospital Outpatient Departments vs. Office-Based Surgical Suites
Competition is increasing among ASCs, HOPDs, and office-based surgical practices. The same improvements in anesthesia and surgical equipment that made outpatient surgery in a freestanding ASCs safe to perform have also led to a growing number of office-based surgical suites and HOPDs. Procedures such as dental, ophthalmologic, endoscopy, cosmetic surgery, and liposuction are increasingly being performed in office-based facilities.

Physicians’ offices are under lower regulatory oversight in comparison to ASCs and HOPDs. This has resulted in the phenomenon of “practice drift,” whereby physicians perform procedures outside of their typical scope of practice. Despite several high-profile adverse events in lay media, prospective studies have demonstrated office-based surgical suites to be of similar safety. A recent, large study in the area of cosmetic surgery compared hospital-based procedures with office-based surgical suites and with ASCs and found accredited office-based surgical suites to be a safe alternative to ASCs and hospitals.

Aging Population
With increased risk of complications from surgery, the older adult population present a unique challenge to ASCs. ASCs must be prepared for complications that may arise from operating on older patients, and they must do so without the same resources as a hospital. The potential benefit of outpatient surgery for elderly patients is substantial, however, as older adults often suffer from postoperative cognitive dysfunction in an inpatient setting, which may be minimized with early discharge. Additional benefits include lower cost to patients on fixed incomes and increased time and comfort at home.

CONCLUSION
ASCs represent a large benefit to society because of their potential to reduce the financial burden of the healthcare system on the economy. At the same time, ASCs also provide high-quality care which is patient and family centered and convenient for providers. ASCs reduce the length of stay and minimize surgical delays and cancellations. In the future, ASC quality will be rigorously tracked due to more regulatory oversight and data collection. The future of ambulatory surgical centers remains bright. Case numbers are increasing across many specialties, including general surgery, plastic surgery, vascular, urologic, and orthopedic practices. Surgeons should follow the maturation of ambulatory surgery closely.

REFERENCES

4. Munnich EL, Parente S. Procedures take less time at ambulatory surgery centers, keeping costs down and ability to meet demand up. Health Affairs 2014;33(5):764-769.
20. Trentman TL, Mueller JT, Gray RJ, Pockaj BA, Simula DV. Outpatient surgery performed in an ambulatory surgery center
CHAPTER 52
AMBULATORY SURGERY
CHAPTER 52
AMBULATORY SURGERY

versus a hospital: comparison of perioperative time intervals. 
21. Rechtoris M. 51 things to know about the ASC industry. Becker’s 
ASC Review. Available at: https://www.beckersasc.com/ 
asc-turnarounds-ideas-to-improve-performance/50-things-to- 
know-about-the-asc-industry-2017.html. Accessed August 27, 
2018.
22. Ambulatory Surgery Center Association. Medicare certifi- 
cation. Available at: www.asccassociation.org/federalregula-
23. Ambulatory Surgery Center Quality Collaboration. Quarterly 
port.cfm.
24. Dickson V. CMS seeks closer look at quality of care at ambu-
latory surgical centers. Modern Healthcare. 2017. Available at: 
http://www.modernhealthcare.com/article/20170711/ 
25. Paquette IM, Smink D, Finlayson SR. Outpatient cholecys-
tectomy at hospitals versus freestanding ambulatory surgical 
ownership. 2011. Available at: https://www.advancingsurgicalcare. 
com/asc/benefitsofphysicianownership. Accessed August 27, 
2018.
27. Becker’s Hospital Review. ASC ownership trends in 2017: 
more complex joint venture deals. Available at: www. 
beckershospitalreview.com/hospital-transactions-and-valuation/ 
asc-ownership-trends-in-2017-more-complex-joint-venture-
tives on physician productivity in medical groups. Health Serv 
29. Hollenbeck BK, Dunn RL, Suskind AM, Strope SA, Zhang 
Y, Hollingsworth JM. Ambulatory surgery centers and their 
intended effects on outpatient surgery. Health Serv Res. 
30. Plotzke MR, Courtmanche C. Does procedure profitability 
impact whether an outpatient surgery is performed at an ambu-
latory surgery center or a hospital? Health Econ. 20(7):817-830.
31. Mitchell JM. Effect of physician ownership of specialty hos-
pitals and ambulatory surgery centers on frequency of use of 
JW. Where do I send thee? Does physician-ownership affect 
referral patterns to ambulatory surgery centers? Health Affairs. 
Abuse Laws. Available at: oig.hhs.gov/compliance/physician-
34. American College of Surgeons. Should your health care system 
invest in an ambulatory surgery center? A decision-making 
framework [Internet]. Bulletin of the American College of 
bulletin.facs.org/2017/11/should-your-health-care-system-
invest-in-an-ambulatory-surgery-center-a-decision-making-
framework/.
35. Mezei G, Chung F. Return hospital visits and hospital readmis-
36. Twersky R, Fishman D, Homel P. What happens after discharge? 
1997;84:319-324.
37. Fox JP, Vashi AA, Ross JS, Gross CP. Hospital based, acute 
743-753.
Postoperative acute care use after freestanding ambulatory 
39. Fleisher LA, Pasternak LR, Lyles A. A novel index of elevat-
ed risk of inpatient hospital admission immediately follow-
40. Fleisher LA, Pasternak LR, Herbert R, Anderson GF. Inpa-
thospital admission and death after outpatient surgery 
in elderly patients: importance of patient and system charac-
41. Urman RD, Punwani N, Shapiro FE. Office-based surgical 
and medical procedures: educational gaps. Ochsner J 
42. Coldiron B, Shreve E, Balkrishnan R. Patient injuries from 
surgical procedures performed in medical offices: three years of 
discussion 1443.
safe? Comparing outcomes of 183,914 aesthetic surgical 
procedures across different types of accredited facilities. 
44. Aurini L, White P. Anesthesia for the elderly outpatient. 
This page intentionally left blank
BACKGROUND

Introduction
The basic American surgical training model is patterned after that established by William Steward Halsted at the Johns Hopkins Hospital in the late 19th century. By tradition, it has consisted of a regulated apprenticeship system with incremental increases in clinical responsibility for resident surgeons over a specified period of time culminating with the expectation that independent surgical practice will be possible at its conclusion. While this approach has served well throughout the 20th century, the pace of changes in healthcare delivery and society itself has driven the need for new approaches to education and training. Among the important changes that the surgical training community has had to confront are an expanding array of medical and surgical therapies, complex, new and sometimes disruptive technologies advancing at a dizzying pace, and increasing recognition of the opportunities for errors and adverse outcomes because of inadequate preparation for this new world of surgical practice.\(^1\) Appreciation of the scale and insidious consequences of medical errors began to take hold in the 1990s when the groundbreaking Institute of Medicine report “To Err is Human” presented to the public a dire picture of the frequency and implications of medical errors.\(^2\) Among the proposed new approaches to deal with this problem was the use of simulation training.

Simulation is the imitation of an actual or possible real-world condition or event. The degree to which a simulation resembles its real-world counterpart describes its fidelity. There are innumerable specific applications of simulation to train or assess human performance on anticipated real-world tasks. As such, it is a valuable tool with great potential to increase safety in high-stakes undertakings such as commercial aviation, military training, and nuclear power generation. These are a few areas where such training has been implemented successfully. With the advent of widely available digital computing, the level of technology applied to simulation and simulators has made near-complete fidelity possible. Moreover, the enormous expense associated with full motion flight simulation, for example, has been made acceptable by safety gains that benefit entire populations. In medicine, there has been gradual acceptance of the role of simulation to achieve these same goals, and this has led to an array of simulation methods and uses that specifically target surgical providers at all levels. The fundamental justification for commitment of resources to simulation is an ethical one: to reduce patient risk associated with invasive surgical procedures or management of complex clinical problems. Simulation methods provide opportunities for surgical learners to practice their skills under safe conditions in preparation for clinical experiences and to be assessed and deemed ready for those encounters.

Skills Labs and Skills Training
The history of surgical skills training outside the clinical operating room (OR) is a long one, involving practice of surgical skills and procedures using various models, including animals and cadavers. Newer training practices, including simulation use in a laboratory setting, have emerged that focus on objective assessment of skill and establishment of specific, defined levels of proficiency. These assessment-based approaches are relatively recent educational developments and are being implemented as a means to improve surgeon skill in a safe setting for both learners and patients.\(^3\) The use of inanimate benchtop models to test surgical skills required a leap into the world of validation of measurement methodologies. The most important of these pioneering efforts was the implementation of the objective
Key Points

1. Learning basic skills at the point of care imposes inefficiencies that might very well endanger support for the education mission.

2. In 2006, the Accreditation Council for Graduate Medical Education Residency Review Committee for Surgery instituted a formal requirement for simulation training in surgical residency.

3. Procedural skills training in a simulated environment has been shown to transfer to the real-life clinical setting.

4. Early studies of virtual reality training using both proficiency-based and non–proficiency-based training methods showed it to be an effective means of improving laparoscopic skill both in the lab and in the operating room compared to non–virtual reality trained controls.

5. Use of proficiency-based training in the context of a larger curriculum appears to be the best way to achieve good training results irrespective of the training platform used.

6. When assessing simulator validity, researchers have noted that the use of robotic surgery simulators does translate to the clinical environment and the learning curve for initial console training for surgeons is significantly decreased.

7. Simulation training for communication and other teamwork-pertinent nontechnical skills requires learners to be embedded in realistic scenarios pertinent to a healthcare team’s actual clinical responsibilities.

8. Simulation technology allows trainees the opportunity to execute a variety of tasks and procedures while also experiencing the cognitive demands of surgery, including error correction and surgical planning decisions.

9. The immediate future of simulation in surgery will likely see expanded use of proficiency-based training given the consistent demonstrations of effectiveness in improving surgeon skills and improved educational outcomes as measured in clinical settings.

10. Advances in wearables, motion tracking, and sensor technologies allow for a wide variety of hybrid and augmented experiences in simulation as well as extensive opportunities for the development of new performance metrics.

structured assessment of technical skills (OSATS) program by educators at the University of Toronto. Using a series of reproducible physical models of surgical tasks (excision of a skin lesion, bowel anastomosis, insertion of a T tube, and abdominal wall closure) and carefully designed rating instruments, it was possible to show validity of these measurements when compared with skill manifested during surgery in animal models. This demonstration of practical measurement of skill in the lab, where observation for skills rating purposes can be more readily achieved, was seminal in sparking both additional interest in lab-based training and in simulation as a mainstream educational method for surgeons.

The advent of laparoscopic general surgery in the late 1980s and early 1990s, a disruptive technology at that time, was a major driver for the use of surgical simulation to gain unfamiliar and nonintuitive skills needed to safely perform limited-access videoendoscopic surgery. The recognition that bile duct injury risk is increased by surgeon inexperience stimulated interest in simulation in order to separate a component of the psychomotor learning curve for laparoscopy from the clinical OR and patient by transferring it to the training lab. At the same time, interest in the science of skills acquisition and measurement made efforts to gain surgical proficiency with simulation more meaningful. The measured performance characteristics of experienced surgeons came to be appreciated as useful learning targets for less experienced surgical learners, and the concept of proficiency-based training began to be implemented successfully not only for research purposes but also for formative education.

As simulation began to be used more extensively and skills labs either evolved or merged with multidisciplinary education centers using simulation as a primary instructional method, recognition of such centers as focal points for surgical education also grew. The American College of Surgeons (ACS) Education Division recognized early on that simulation training was an important educational method that surgeons could take advantage of and began to explore ways to facilitate growth and implement training centers. In 2006, the ACS began to accredit education centers engaged in simulation training as Level 1 (comprehensive) and Level 2 (focused) Education Institutes. The consortium of 95 institutes, as of early 2017, now spans the globe with centers across the United States and Canada as well as several institutions outside of North America.

In 2007, the American College of Surgeons and the Association of Program Directors in Surgery (APDS) initiated a project to provide a standardized skills curriculum for surgical residents. These efforts produced the modular ACS-APDS skills curricula that represent the first comprehensive and widely available resource that prescribes simulation experiences as the principal means to achieve educational objectives. The resources and curricula are readily available to residency programs to address surgical resident learning needs and to facilitate simulation lab use as mandated by the Accreditation Council for Graduate Medical Education (ACGME). The three components of the curriculum are basic technical skills, procedural skills, and team skills. Although these have been acknowledged to be valuable and have been utilized to meet resident training needs, implementation has been limited according to a recent survey of residency programs. Elements of the basic skills curriculum were reported to be used by 36% of respondents while procedural and team training modules were reported to be in use by about half as many programs. Nonetheless, the article notes that simulation methods had permeated surgical education and the various stakeholder organizations had taken steps to either endorse or implement programs for simulation use. The need to explore simulation-based education, as stated by the Surgical Council on Resident Education (SCORE), suggests a pathway by which simulation methods might become fully integrated into standardized cognitive elements of surgical residency curriculum. Based on current trends, a competency framework can be envisioned that combines advanced forms of...
standards-based learning using simulation and measured performance in the OR for purposes of trainee advancement and certification.

**Scientific Underpinnings of Simulator Use**

The assumption that surgical technical skills can be effectively learned and tested outside the OR serves as the basis for simulation based assessment of skills before, during, and following training. This capability permits the explicit description of understandable surgeon performance characteristics including those that can be characterized as desirable training goals (e.g., “expert”). The effectiveness of specific models of training can also be ascertained by comparative studies examining educational endpoints, such as rate of learning, or testable performance either in simulations or in clinical settings. Simulation-based learning has been described using various pedagogical models to help educators understand and leverage the most effective strategies to achieve educational goals. The most fundamental concept used to justify time spent in skills training is the prospect of achieving progressively higher levels of skill pertinent to clinical care. For medical care providers, this progression is commonly related to the learning model described by Dreyfus and Dreyfus.\(^1\) This model depicts changes in specific mental functions (recollection, recognition, decision, awareness) associated with incremental steps in the progression from novice status to that of expert (Fig. 53-1). The relevance of Anders Ericsson’s description of deliberate practice to development of clinical skills is also widely accepted.\(^5\) This is predicated on the concept that “expert” performance is fundamentally different than normal performance and results from behaviors that “reflect a life-long period of deliberate effort to improve performance in a specific domain.”\(^6\) The drive to achieve mastery in surgery, provides an aspirational model for the surgeon’s use of simulation because the opportunity for deliberate, repetitive practice of numerous skills in the clinical setting alone is insufficient and in many cases nearly impossible. For this concept to be translated into effective educational results, simulation-based education must be of high quality, and the curricular framework created for its use must be directed toward carefully crafted, attainable, and clinically relevant educational goals.

**Simulation in Graduate Medical Education**

Surgery residency has been a particular area of focus for simulation use in assessment and training. The need for safe and rapid skills development is especially important in this group of learners whose time in training is limited and for whom every hour spent in education must be prioritized for highest impact. The added direct costs of operative care in training institutions has also provides a strong imperative to conduct training that minimizes negative implications of resident involvement in surgical cases.\(^17\) Learning basic skills at the point of care imposes inefficiencies that might very well endanger support for the education mission.

---

**Figure 53-1.** Dreyfus model describing stepwise skills development. In surgery, specific stages of expertise are achieved through cognitive learning, technical practice, and experience and are defined by specific cognitive and behavioral characteristics affecting how we perceive, process, and act in the task environment. (Reproduced with permission from University of South Australia; Teaching and Learning in Health Sciences: https://lo.unisa.edu.au/mod/book/view.php?id=611284&chapterid=104113.)
Despite early examples of excellence in simulation lab use and fairly wide availability, a survey of residency programs in 2006 suggested the quality of usage for laparoscopic skills training was probably low and that more than half did not use a defined curriculum. Also in 2006, the ACGME Residency Review Committee (RRC) for Surgery instituted a formal requirement for simulation training in surgical residency. While the core program requirement did not define any specific educational objectives, performance outcomes, or specific methods to be used, there was a mandate that simulation in surgical education be objectives-based and that it be incorporated into residency curricula within 2 years. As of 2016, the relevant core program requirement indicates that “resources must include simulation and skills laboratories” and that “these facilities must address acquisition and maintenance of skills with a competency-based method of evaluation.” Since the original notification of the need to conduct this training, mandated activities have expanded with the pairing of ACGME statements and new American Board of Surgery (ABS) requirements for certifications in Fundamentals of Laparoscopic Surgery (FLS; instituted for 2010 residency graduates) and in Fundamentals of Endoscopic Surgery (FES; instituted for 2018 residency graduates). Both of these certifications demand lab-based practice in order to successfully pass respective technical skill test components.

Both high and low technology and fidelity, off-the-shelf, simulation training solutions have become available for surgery residents’ lab-based training. Curricular content such as the ACS-APDS skills modules are readily available to facilitate a sufficient level of implementation to meet the RRC requirements, although preparation methods have not been specifically prescribed by either the ACGME or the ABS. In 2015, the ABS issued a new requirement for a comprehensive multilevel Flexible Endoscopy Curriculum (FEC) paired with FES certification. This curriculum states explicitly that experiential requirements at the lower levels can be met with simulation methods but it is left to individual programs to decide what preparatory practice in simulation might consist of for their own residents. 

Options for simulation use in surgical training are currently numerous and can be implemented with basic facilities and equipment, but the determination to use these methods successfully requires considerable effort that is greatly aided by fully motivated and engaged faculty members with protected time. There are now numerous guides and recommendations for successful surgical simulation lab start-ups.

“Bootcamps”
Preparatory training for surgery residency both before and immediately after the start of residency is now a widely used educational practice. The rationale for these programs includes early development of basic skills that would be of obvious use to the new trainee as well as determination of the status of basic skills that would serve as a needs assessment to model curricular efforts expected to be most appropriate for the individual resident. The suggestion that new interns will render safer care has been garnered from results of intensive, short-duration preparatory training that shows higher level of skills measured in simulation are feasible compared to the “control” situation of not using such training. However, none of the current reports address whether intensive preparatory training addresses the “July Effect” by improving clinical effectiveness during the earliest months of training.

Much of the training and assessment during these programs is accomplished using simulation methods. No standardized approach has yet been suggested, and at the present time the impact of specific simulation components is difficult to assess for senior students entering surgical training, or new interns. Sound recommendations on this await further study. A single meta-analysis of postgraduate “boot camp” programs both prior to or at the start of residency found that all programs utilized high and low technology simulation methods as “a key component.” Although the analysis included all medical specialties, 93% of the studies were surgical in nature, underscoring the perceived value of this preparatory training for new surgeons. The examined studies were those where pre- and posttraining effects were measured in some way. The compiled data revealed that trainees who completed the programs had uniformly strong increases in skills development, knowledge, and confidence. In 2014 the ABS, ACS, APDS, and Association for Surgical Education, citing the evidence of effectiveness of these now numerous preparatory courses, officially endorsed them as a useful method to position fourth-year students and interns for early success in residency. Whether preresidency exposure to surgical simulation can influence a medical student’s decision to pursue a surgical career remains to be determined.

### SPECIFIC SIMULATION TRAINING AND ASSESSMENT APPLICATIONS

#### Training Basic Surgical Skills
The use of the simulation lab to train open, laparoscopic, and flexible endoscopic basic skills in preparation for care and practice clinical constitutes the most accessible and widely adopted set of simulation training practices. In connection with these, the term “proficiency-based” training (sometimes used synonymously with competency-based or objectives-based) is often used, and can be defined as the use of meaningful performance standards as educational goals for training. Implicit in the term is the expectation that if proficiency standards are achieved, a learner will be better positioned to perform to a desirable standard in the clinical setting. Even these most basic skills have been shown to improve with proficiency-based training. Additional benefits of proficiency-based approaches to basic surgical skills include knowledge of a surgical learner’s specific educational needs and design of a larger skills curriculum roadmap that envisions progression to more advanced skills training.

Educators at Southern Illinois University implemented a program of intensive lab-based practice of basic skills as a preliminary requirement to participate in operative cases. Termed “Verification of Proficiency,” this program targets junior trainees for assessment of skill using OSATS-like rating instruments, but its unique feature is the definition of a minimal level of skill that would permit a resident to assume the role of operator for specific case types.

More recently, the concept of “proficiency-based progression” has been used to describe a formalized process of use of sequenced proficiency standards and a continuum of progressively more advanced and challenging simulation experiences as a potential future model of training. Such a model would span whatever period of time was necessary for each learner to progress from basic surgical skills to more advanced ones.

#### Open Surgical Skills
A variety of benchtop models for practice and assessment of basic open skills are available, the most...
Table 53-1

ACS-APDS basic skills curriculum components

- Asepsis and instrument identification
- Knot tying
- Suturing
- Skin flaps
- Skin grafts
- Urethral catheterization
- Airway management
- Chest tube insertion
- Central line insertion
- Surgical biopsy
- Laparotomy opening and closure
- Basic laparoscopy skills
- Advanced laparoscopy skills
- Hand-sewn bowel anastomosis
- Stapled bowel anastomosis
- Arterial anastomosis

well-known of which are the OSATS tasks that remain in use more than 20 years after their inception. The ACS-APDS Basic Skills Curriculum (Table 53-1) is a useful resource for this.

Generally, lower-fidelity models are more cost effective for repetitive training because durability with reuse may be better than high-fidelity physical models and replacement costs can be minimized for many tasks. Examples can include simple models for abdominal wall closure (Fig. 53-2). The use of low-fidelity models in well-formulated curricular training can be very effective in increasing resident skill, although high-quality measurement can be labor intensive. Bowel and vascular anastomosis training has been an especially important point of focus given their prominence and highly technical nature in clinical practice. In a recent analysis of resident experience with gastrointestinal anastomosis, Nemeth reported that although frequently performed (average 67 per resident), stapled anastomosis experience predominates during training (91% of laparoscopic procedures and 82% of open ones), suggesting a relatively small clinical experience with hand-sewn anastomotic methods, which remain critically relevant. Simulated bowel anastomosis models are widely available (Fig. 53-3), and there are numerous reports of successful curriculum-based development of this core technical skill generally utilizing OSATS rating instruments. Although further study is required to demonstrate a clinical effect of this training, other experience strongly suggests that use of proficiency-based anastomosis training should be the next step. In an example of how this could be used for trainee benefit, Palter conducted a randomized trial of use of a proficiency-based technical skills and cognitive curriculum for abdominal wall closure. The technical skills component utilized a low-fidelity model for an OSATS assessments of both lab and OR performance. Residents in the intervention arm performed better on both clinical abdominal wall closure and a test of procedural knowledge than controls.

**Basic Laparoscopic Manipulative Skills.** Current evidence supports the concept that basic laparoscopic skills should be developed in the training lab, avoiding safety concerns as well as the expenditure of time and effort that would result from learning in clinical settings. Specific programs for basic laparoscopic skills development coupled with skills assessment were set forth 20 years ago at a time when this need was newly articulated for surgeons at all levels. These utilized videoscopic training “boxes” for two-handed practice using laparoscopic instruments to manipulate box contents. Programs such as the Rosser drills employed various dexterity tasks emphasizing time for task completion as a measure of performance. Later, McGill investigators began to apply greater scientific rigor to the design

![Figure 53-2](image-url) Abdominal wall closure model made from commonly available materials using instructions in the ACS-APDS Basic Skills Curriculum module for this task.
and testing of basic skills curricula, with a particular focus on the ability to characterize the skills acquisition process and differences between learner groups based on careful repetitive measurement. Further study showed that performance, measured using these bench training tasks, correlated with performance measured in vivo animal models. In addition to helping learners prepare for basic tissue manipulation, training on basic tasks were found to shorten the learning curves for more complex laparoscopic tasks such as suturing.

**Fundamentals of Laparoscopic Surgery**

Fundamentals of Laparoscopic Surgery (FLS) was devised by minimally invasive surgical leaders in the Society of American Gastrointestinal and Endoscopic Surgeons as a means to assess laparoscopic surgical knowledge and skills for the purposes of certification of basic ability. This program represents the first broadly applied effort to demonstrate achievement of a specified level of basic surgical skill with such specific testing. Although most utilized by surgical residents, the resulting skills certification is applicable to a broad range of learners, including surgeons in practice as well as laparoscopic surgeons in specialties other than general surgery. FLS consists of separate tests of knowledge and technical ability. The latter component required integration of simulation tasks with a high degree of validation for both effectiveness in discerning skill and relevance to the respective clinical tasks. The McGill Inanimate System for Training and Evaluation of Laparoscopic Skills (MISTELS) tasks were selected as the technical skills assessment component. These were developed independently from the FLS program and adapted to the FLS based on the predictive value of a subset of the tasks for clinical skills manifested in the OR. These tasks have remained core features of the certifying examination and have been extensively studied both as training curriculum components and as predictors of clinical performance.

Recent work has suggested that specific proficiency-based training can increase pass rates on FLS to 100%, supporting the general suppositions about the benefits of this model of training.

The more recently available FES serves a similar purpose to FLS for flexible endoscopy. This certification adds the important feature of use of a virtual reality flexible endoscopy simulator as one of the platforms for delivery of the technical skills test.

**Bench Models for Training Specific Procedural Skills**

Procedure-specific simulation offers a unique opportunity to practice and evaluate wholesome surgical skill. This includes training and evaluating the relationship between basic and complex technical skills and surgical decision-making. While the majority of procedural simulation trainers focus on bedside procedures such as central venous catheter placement, bladder catheterization, and intubation, a number of trainers also have been designed to simulate more complex procedures such as laparoscopic ventral hernia, laparoscopic colectomy, and robotic nephrectomy. Currently, both virtual reality and physical or hybrid models are used for procedure-specific simulations. While each technology has specific benefits, there is still a critical need for fabrication and design approaches that are efficient, cost effective, and produce anatomically accurate models with realistic tissue properties. A number of groups have explored the use of three-dimensional (3D) printing with varying degrees of success. As the printing substrate materials and 3D machines continue to decrease in price and increase in ease of use, there will likely be a significant upsurge in the use of this approach to facilitate development of procedure specific simulations. In addition to full immersion virtual reality, a number of groups are still exploring the benefits of computer-based learning for training procedure-specific surgical skills.

The training and assessment benefits of procedure-specific simulation are numerous. Not only does it allow an opportunity to assess technical skill in the context of a multistep procedure, forced errors and critical decisions can be combined with this technology, allowing for a more in-depth learning experience and skills assessment. By way of example, two multistep bedside procedures were modified: bladder catheterization and central venous catheter insertion. Both simulations had embedded clinical scenarios, which if executed incorrectly in real life might produce patient injury due to incorrect technique or equipment choices. During the implementation of these modified, multistep procedures, it was possible to identify individual cognitive and technical errors that serve as opportunities for additional training. The ACS-APDS Procedural Skills Curriculum was developed to complement the Basic Skills Curriculum and offers detailed learning objectives for a variety of surgical procedures. The goal of this curriculum was to address holistic surgical skills in a context where both technical
and cognitive performance could be assessed at the same time.11

Table 53-2

<table>
<thead>
<tr>
<th>ACS-APDS procedural skills curriculum components</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Laparoscopic ventral hernia repair</td>
</tr>
<tr>
<td>• Open colon resection, lap right colon resection</td>
</tr>
<tr>
<td>• Laparoscopic sigmoid resection</td>
</tr>
<tr>
<td>• Open right colon resection</td>
</tr>
<tr>
<td>• Laparoscopic/open bile duct exploration</td>
</tr>
<tr>
<td>• Laparoscopic ventral/incisional hernia repair</td>
</tr>
<tr>
<td>(porcine model)</td>
</tr>
<tr>
<td>• Laparoscopic appendectomy</td>
</tr>
<tr>
<td>• Laparoscopic Nissen fundoplication</td>
</tr>
<tr>
<td>• Sentinel node biopsy and axillary lymph node dissection</td>
</tr>
<tr>
<td>• Open inguinal/femoral hernia repair</td>
</tr>
<tr>
<td>• Laparoscopic inguinal hernia repair</td>
</tr>
<tr>
<td>• Laparoscopic/open splenectomy</td>
</tr>
<tr>
<td>• Laparoscopic/open cholecystectomy</td>
</tr>
<tr>
<td>• Gastric resection and peptic ulcer disease</td>
</tr>
<tr>
<td>• Parathyroidectomy/thyroidectomy</td>
</tr>
</tbody>
</table>

Transfer of Training

The determination that simulation-based education is effective requires assessment of training effect in either the clinical setting or one with a demonstrated relationship to a clinical setting. The latter might result from comparison to a validated “gold standard” training method. Establishing a relationship between lab-based training and performance in the OR requires validated measures of operative skills such as the widely used Global Assessment of Operative Laparoscopic Skills (GOALS) method described by Vassiliou.65 The majority of studies of laparoscopic simulation training that examine transfer of skills to the clinical OR make comparisons to control groups without highly structured and non-simulation based training. The results of these studies should eliminate any doubts that surgical learners, especially students and residents, can achieve training benefits from both proficiency-based and time- or repetition- or session number-based simulation practice. When the results of proficiency-based training are dissected out from the other models of training, consistently higher levels of clinical or animal OR performance is observed with such training compared to without.64,67 In one such review of skills transfer studies conducted between 2007 and 2013, Dawe reported results for general surgery and gynecology procedures using different virtual reality and non-VR laparoscopic simulation platforms (Table 53-3).48,67,77 For the 12 randomized controlled trials with surgical residents as study subjects, all but one showed significantly better clinical performance for intervention groups compared to non–simulation-trained controls.

The comprehensive reviews of skills transfer data underscore the wide disparities in study design characteristics, metrics, simulator types, and the difficulties in comparing effectiveness of different simulation interventions. The “transfer effectiveness ratio” (TER) has been forwarded as a means of expressing the relative magnitude of the training effect and may provide a basis for comparison of cost or time efficiency of different training methods.78 To determine TER for a simulation training effort, one would calculate the difference in clinical effort (time or some other measure such as number of cases) between simulator-trained and alternative-trained groups to achieve a desired level of clinical performance, divided by training time received by the simulator-trained group.

Korndorffer raised concerns with studies of transfer of training when, irrespective of the improved results in lab performance with proficiency-based practice, residents studied by their group did not readily meet performance standards for more advanced skills set by expert surgeons in practice.79 Looking for opportunities to maximize the effectiveness of simulation training, Stefanidis proposed training to “automaticity” by adding a secondary visuospatial task to practice with laparoscopic suturing, but only after proficiency levels were achieved with more standard laparoscopic suturing practice.80 It was suggested that the added attentional challenge likely prompted the achievement of automaticity, the end result of which was much greater performance than was seen for proficiency-based practice alone. It is not clear, however, that simply adding to task difficulty improves training results if learner capabilities are not taken into account. In a separate study, Stefanidis also reported that increased task difficulty during proficiency-based training causes measurable increases in cognitive workload and that when confronted with these challenges, novice learners did not perform better than those in proficiency-based training at a lower level of difficulty, suggesting a possible mismatch between task challenges and capacity to learn.81 More work is required to characterize the proper balance of training difficulty and the capacity for learning in specific learners or learner groups.

Virtual Reality

The use of virtual reality (VR) simulation as a way to deliver training experiences in surgery was proposed by Satava in 1993.82 Within a few years, practical applications of this technology led to the first commercially available laparoscopic simulators and studies to determine their value. The earliest and most functional VR platform was MIST-VR which permitted manipulation of abstract virtual objects using a realistic physical interface that transduced instrument motion into actions that could be observed in the virtual environment (Fig. 53-4).83 Even without a force feedback apparatus and haptic, or “sense of touch,” cues expected with instrument–instrument or instrument–object interactions, learners could experience the psychomotor challenges of videolaparoscopy and iteratively train until
### Table 53-3
Randomized trials studying the effects of virtual reality training on surgical and OB-GYN resident operative performance vs. control trainees without virtual reality training

<table>
<thead>
<tr>
<th>AUTHOR (YEAR)</th>
<th>PARTICIPANTS (N) AND SIMULATOR</th>
<th>PROCEDURE ASSESSED</th>
<th>CONCLUSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahlberg et al (2007)</td>
<td>PGY 1–2 surgery residents IG (7); Lap Sim VR simulator CG (6) no simulation</td>
<td>Laparoscopic cholecystectomy</td>
<td>IG made fewer errors ($P = 0.004$), exposure ($P = 0.040$), clipping and tissue division ($P &lt; 0.008$), and dissection ($P &lt; 0.031$) compared with CG</td>
</tr>
<tr>
<td>Banks et al (2007)</td>
<td>PGY 1 GYN residents IG (10); Limbs &amp; Things box trainer CG (10) no simulation</td>
<td>Bilateral tubal ligation</td>
<td>IG scored higher than CG with all three evaluation tools: task-specific checklist ($P = 0.002$), OSATS ($P = 0.003$), pass-fail grade ($P = 0.003$)</td>
</tr>
<tr>
<td>Cosman et al (2007)</td>
<td>Junior surgical trainees IG (5); Lap Sim VR simulator CG (5) no simulation</td>
<td>Laparoscopic cholecystectomy (clip application and cystic artery division)</td>
<td>IG had fewer errors ($P = 0.05$), better bimanual coordination ($P = 0.05$), higher global score ($P = 0.04$) than CG</td>
</tr>
<tr>
<td>Gala et al (2013)</td>
<td>GYN residents IG (48); FLS box trainer CG (54) no simulation</td>
<td>Pomeroy bilateral tubal ligation</td>
<td>IG had higher OSATS progression score than CG ($P = 0.03$)</td>
</tr>
<tr>
<td>Hogle et al (2009)</td>
<td>PGY 1 surgery residents IG (6); Lap Sim VR simulator CG (6) no simulation</td>
<td>Laparoscopic cholecystectomy</td>
<td>No significant difference between IG and CG in GOALS Domain areas of depth perception, bimanual dexterity, efficiency, tissue handling, autonomy</td>
</tr>
<tr>
<td>Larsen et al (2009)</td>
<td>First- and second-year OB-GYN registrars IG (13); Lap Sim VR simulator CG (11) no simulation</td>
<td>Salpingectomy</td>
<td>IG had higher score than CG in OSA-LS scale ($P &lt; 0.001$). IG completed procedure faster than CG ($P &lt; 0.001$)</td>
</tr>
<tr>
<td>Sroka et al (2010)</td>
<td>PGY 1–3 surgery residents IG (8); FLS box trainer CG (8) no simulation</td>
<td>Laparoscopic cholecystectomy (excision from liver bed)</td>
<td>IG had higher total GOALS score than CG ($P &lt; 0.001$) and better GOALS domain scores for bimanual dexterity ($P = 0.04$), tissue handling ($P = 0.04$)</td>
</tr>
<tr>
<td>Van Sickle et al (2008)</td>
<td>PGY 3, 5–6 surgery residents IG (11); MIST-VR simulator CG no simulation</td>
<td>Nissen fundoplication (placement of intracorporeal sutures)</td>
<td>IG completed task in less time ($P &lt; 0.003$), committed fewer errors ($P &lt; 0.01$), and had fewer needle manipulations ($P &lt; 0.05$) than CG</td>
</tr>
<tr>
<td>Zendejas et al (2011)</td>
<td>Surgery residents IG (26); Guildford MATTU TEP hernia task trainer CG (24) no simulation</td>
<td>TEP inguinal hernia repair</td>
<td>IG faster on first procedure ($P &lt; 0.001$) and had higher participation rates ($P &lt; 0.001$). For subsequent repairs IG remained faster than CG. GOALS score higher for IG ($P = 0.001$). Complications and overnight stay less likely for first TEP procedure in IG ($P &lt; 0.05$).</td>
</tr>
<tr>
<td>Palter et al (2012)</td>
<td>PGY 2–4 surgery residents IG (9); curriculum including simulation training on LapSim VR simulator CG (9) no simulation</td>
<td>Right hemicolecctomy</td>
<td>IG had higher OSATS score ($P = 0.030$) than CG. IG able to perform more operative steps than CG ($P = 0.021$)</td>
</tr>
<tr>
<td>Palter et al (2013)</td>
<td>PGY 1–2 surgery residents IG (9); curriculum including simulation training on LapSim VR simulator and on FLS box trainer CG (9) no simulation</td>
<td>Cholecystectomy</td>
<td>IG had higher OSATS scores for first four procedures ($P = 0.004$, $P = 0.036$, $P = 0.021$, $P = 0.023$)</td>
</tr>
</tbody>
</table>

PGY = postgraduate year; IG = intervention group; CG = control group; VR = virtual reality; OSATS = Objective Structured Assessment of Technical Skills; GOALS = Global Operative Assessment of Laparoscopic Skills; OSA-LS = Objective Structured Assessment of Laparoscopic Salpingectomy; TEP = totally extraperitoneal.

Simulators: LapSim VR simulator (Surgical Science, Gothenburg, Sweden); laparoscopic simulator and Minimal Access Therapy Unit (MATTU) (Limbs and Things, Bristol, UK); Fundamentals of Laparoscopic Surgery (FLS) Training Box simulator (SAGES, Los Angeles, California, USA); Minimally Invasive Surgical Trainer—Virtual Reality (MIST-VR; Mentice, Gothenburg, Sweden).

performance goals for precision, efficiency, and error avoidance were achieved. Performance measurement was automated and included time, instrument motion, and electrosurgery use metrics, as well as a tally of the occurrence of predefined errors. All metrics were free of human observer bias. Early studies of VR training using both proficiency-based and non-proficiency-based training methods showed it to be an effective means of improving laparoscopic skill both in the lab and in the operating room compared to non-VR trained controls. Since the first studies of this type were performed almost 20 years ago, several comprehensive reviews of the growing body of literature on VR have continued to support the conclusion that skills acquired in VR transfer to the clinical setting (Table 53-3), not only for laparoscopy, but also for flexible endoscopy, sinusoscopic surgery, and endovascular interventions. Largely due to small study sizes and some design limitations, the quality of evidence of these studies is consistently described as below level I.

Based on available evidence, expanded use of VR for skills training could be justified, but few comparisons of training effectiveness have been made between physical laparoscopic video trainer (“box” trainer) and laparoscopic VR simulator-based training. Crossover studies designed to determine if training in one environment improves performance in the other have not been especially helpful in defining the value of either. Only a few studies have compared the effects of the two training methods on OR performance (Table 53-4). Although some advantage has been suggested, the prevailing view is that both can be used for highly effective laparoscopic practice. Until better comparisons are made, use of proficiency-based training in the context of a larger curriculum appears to be the best way to achieve good training results irrespective of the training platform used.

Table 53-4

<table>
<thead>
<tr>
<th>AUTHOR (YEAR)</th>
<th>STUDY ARMS (N)</th>
<th>SUBJECTS</th>
<th>PROFICIENCY-BASED TRAINING?</th>
<th>CONCLUSIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diesen et al (2011)</td>
<td>VR (10) Box trainer (8)</td>
<td>Interns Medical students</td>
<td>No</td>
<td>No difference in interval blinded video assessments of animal OR task performance</td>
</tr>
</tbody>
</table>

VR = virtual reality; OR = operating room.
VR simulator systems are significant capital investments for simulation centers that have competing procurement priorities. In a prospective randomized trial, Orzech analyzed the cost impact of each type of training across Canadian residency programs and found that the transfer of training effect was greater for VR as compared to the box trainer group. However, box trainer use was found to be more cost effective except in larger residencies. Although the reasons for the latter finding require further analysis, the high acquisition costs of VR systems can be offset by an economy of scale benefit not realized with physical models where staffing for purposes of task setup and assessment and consumable items related to many tasks represent additive costs. Flexible endoscopy VR simulator device sharing was shown to work effectively for a skills acquisition program across a network of Texas institutions where procurement of numerous such simulators might not be practical.

The role of haptics has been debated for laparoscopic VR simulators since the inclusion of sophisticated force feedback hardware adds substantially to the cost of surgical VR as well as to system computing demands. In simulated endoscopic and laparoscopic procedures, effective haptic cues are important contributors to the fidelity of the experience. For basic skills acquisition these features have not been shown to offer significant advantages over nonhaptic VR systems, although for more advanced skills haptic cues may permit greater precision of instrument use. However, surgical VR has advanced to the point where inventories of procedure types offered on specific commercially available systems are quite extensive and these will inevitably increase and encompass entire procedures. The realism of the user experience, including the haptic experience, may prove essential to effective learning of advanced surgical skills in a new generation of VR devices. The current generation of laparoscopic simulators are highly capable devices with a high degree of graphical realism, full haptic features, and numerous tasks available for training basic and procedural skills for general surgical, urologic, and gynecologic procedures (Fig. 53-5).

Despite encouraging progress made in VR simulation, two separate publications in 2007 and 2015 identified ongoing computational challenges in development of very high-fidelity simulations for surgery. These include the unique VR problems of modeling human tissues and the added demands of rendering the appropriate deformations when tissues are manipulated. One of the practical examples of this is that of guidewire behavior during VR as compared to real-life endovascular procedures, where variable elasticity characteristics of blood vessel walls were observed to result in real tissue or wire deformations that the VR software could not depict accurately. Also identified were disparities between levels of resourcing for surgical VR versus mainstream computer gaming, raising important questions on how the full power of this technology can be realized to more fully simulate complex operations with a high degree of fidelity and realism.

**Virtual Reality for Flexible Endoscopy**

The use of virtual reality simulation to train for flexible endoscopy procedures is now well established and supported by consistent evidence of both skills transfer. Sedlack compared colonoscopy skills between small groups of inexperienced VR-trained and nontrained gastroenterology fellows and found that VR training resulted in farther progression into the colon, better inspection skills, and a higher percentage of completed studies (those that reached the cecum). This performance advantage extended out to 30 posttraining procedures. In the years since that report, four randomized controlled trials of VR training with blinding to training status during posttraining clinical colonoscopy have been conducted. Despite methodological issues with each of these trials (no proficiency-based training, unsupervised VR practice on one, vaguely defined training characteristics of control groups), all but one showed training benefits for

---

**Figure 53-5.** Three different current generation laparoscopic virtual reality simulators. **A.** LapVR VR Simulator, CAE Healthcare, Sarasota, FL. **B.** Simbionix LAP Mentor VR simulator, Littleton, CO. **C.** LapSim VR simulator, Surgical Science AB, Gothenburg, Sweden.
previously inexperienced VR-trained residents versus controls, especially in the earlier posttraining clinical cases based on either subjective determination of competency or measurement of procedure length. The study that did not show improved clinical performance compared to controls was a noninferiority comparison to control subjects with undefined patient-based training where the two groups performed comparably well. Randomized trials of VR training for upper endoscopy have shown significant performance advantages in both procedure length and subjective assessment of competency compared to both untrained and patient-based training control groups.

Based on the highly standardized test environment that can be experienced in VR and concurrent validity to clinical endoscopy, the SAGES FES program utilizes VR flexible endoscopy simulation for certification of skills. Systematic review of published literature on VR flexible endoscopy skills cite the need for better quality evidence in support of best training practices and improved patient outcomes. Further study is required, especially in the area of proficiency-based VR flexible endoscopy training, which has been generally underutilized.

Virtual Reality for Endovascular Interventions

Virtual reality simulation has been shown to be an effective alternative to training with animal or cadaver models for catheter-based vascular interventions. Current simulations allow development of basic guidewire and catheter handling skills, as well as practice in use of fluoroscopy, angioplasty, and stenting techniques. Procedural training for coronary, carotid, renal, neuro, peripheral vascular, and other interventions are feasible with measurement of learner performance in numerous areas such as procedure and fluoroscopy time, contrast use, and intervention effectiveness. Simulators display fluoroscopy images and allow for the selection and insertion of virtual catheter, balloon, and stent types, which are pertinent to the procedure being performed (Fig. 53-6). In small randomized trials, VR training has been shown to increase residents’ peripheral angioplasty skills with transfer of lab-acquired skills to the clinical OR. After having previously shown that VR simulator assessment can effectively discriminate the level of clinical experience with carotid artery stenting among experienced interventionists, Van Herzeel demonstrated that experienced interventionists could also significantly increase carotid artery stenting skills following a 2-day intensive course of didactic and VR training for this procedure. Following training, decreased procedure and fluoroscopy time and decreased time for placement and retrieval of the embolic protection device were observed. Although differing clinical outcomes were not seen in this small study, it was noted that internal carotid artery spasm frequency decreased after training.

The incorporation of actual patient vascular anatomic information taken from computerized tomography data into a vascular interventional simulation as an aid to procedural planning or technical procedure performance has been referred to as “mission rehearsal” or more commonly now, procedure specific rehearsal or procedure specific simulation. It has been described most extensively for carotid artery stenting procedures and highlights how VR simulation can be directed toward the immediate problems of clinical practice. Cates’ brief report of a single procedure was followed by small studies further demonstrating the feasibility and general impressions of the value of this method. In a randomized comparison of trainees who performed either part-task rehearsal or rehearsal of the entire procedure, Willaert reported that a similar performance benefit could be achieved, suggesting a potentially more time-efficient way to train, although the embolic protection device was in place slightly longer in the part-task-trained group.

Figure 53-6. Representation of patient-specific aortic vascular anatomy during simulation of endovascular aortic replacement (EVAR) for abdominal aortic aneurysm, with distal graft limbs in the ballerina (A, crossed) and standard (B, uncrossed) configurations. This example of patient-specific rehearsal requires preparation the virtual aorta from a DICOM (Digital Imaging and Communications in Medicine) format file of patient computed tomography (CT) or CT angiogram imaging data. Image is rendered on ANGIO Mentor VR simulator (3D Systems, Littleton, CO). (Reproduced with permission from Pakelian D, Van Herzeel I, Lachat ML, et al: EVAR 2020: Training Future Aortic Specialists. Emerging needs and the role of simulation. Endovascular Today 2017 March;16(3):95-100.)
In a recent multinational European study, Desender and colleagues randomized a series of 100 patients scheduled to undergo elective endovascular aneurysm repair (EVAR) for infrarenal aortic aneurysm to either have their procedure rehearsed with VR simulation preoperatively or to have the procedure performed without rehearsal. There were 26% fewer minor errors, 76% fewer major errors, and a 27% fewer errors causing procedural delay in the VR rehearsal group. In addition, this group had significantly fewer angiograms performed to visualize proximal and distal graft landing zones. In a follow-up of this study, Desender reported that patient-specific rehearsal before EVAR resulted in alteration of the operative plan for proximal landing zone (54%), distal landing zone (76%), stent graft main body size (16%), contralateral limb size (34%) or orientation (16%), and iliac extension size (28%). Ninety-two percent of these changes were implemented during the actual EVAR case.

**Patient-Specific VR Surgery Simulation**

In addition to patient-specific rehearsal for endovascular interventions, VR simulations for OR surgical procedures have begun to use patient imaging data to rehearse procedures preoperatively. While many surgeons consciously and subconsciously mentally rehearse procedures before entering the operating room, this process does not allow for fully explicit information sharing between team members. Moreover, even when the surgeon verbalizes a plan for other members of the OR team after the mental rehearsal, it is not uncommon to unintentionally exclude important details that team members may value. The use of anatomically accurate VR simulations, based on patient-specific anatomy, may allow for team-based rehearsals and reduce the risk of human error. In addition, VR-based rehearsals may also facilitate doctor-patient communication.

Patient-specific VR simulations have recently emerged for a variety of complex operations including pancreatectomies, hepatectomies, renal surgery, and hand surgery. For one of the renal surgery simulations, patient-specific computed tomography (CT) data was captured and used to create 3D imaging for incorporation into the simulation. The anatomical accuracy of various structures such as arteries, veins, ureters, and even tumors was reported to be high. Another group compared the appearance of individual vascular structures while performing several patient-specific virtual hepatectomies simultaneously with real-life hepatectomies and also noted a high degree of accuracy. In addition, similar to the way CT angiograms are used in the OR for surgical planning, this group was able to increase and decrease the transparency level of the patient-specific VR simulation and use it both as a real-time operative guide (minimal transparency mode) as well as an operative planning guide (high transparency–vessel only view). This enabled one system to be used seamlessly throughout the actual operation.

According to reports, the time needed to create patient-specific VR simulations is relatively short. On average, it took approximately 2.5 hours each for both the hepatectomy and pancreatectomy simulations. In addition, compared to the use of 3D printed simulations, patient-specific VR simulations are readily reusable and do not consume as many resources. While these recent advances are quite promising, patient-specific VR simulators are a new technology, and thus additional studies are required to more fully understand the pros and cons of introducing this technology into the patient care arena. As there continues to be major improvements in patient-specific 3D rendering, including organ and tissue deformation in reaction to surgical manipulation, this increases the possibility of even more sophisticated and accurate VR simulations that can be used for preoperative planning and rehearsing for complicated procedures.

**Robotic Surgery Simulators**

After the da Vinci surgical system was first introduced in the United States in 1999, a number of simulation systems for teaching robotic surgery emerged. While there have been several versions of the da Vinci system deployed worldwide, the basic system components usually include dual hand controls, foot pedals, and a controllable 3D camera. Consistency in these system components allows for similar consistency in simulation design and delivery. Currently, there are four different simulators geared towards imparting some level of competency in using the da Vinci System: the SEP-Robot (SurgicalSim Educational Platform Robot; SimSurgery, Oslo, Norway); RoSS (Robotic Surgery Simulator-Simulated Surgical Systems, San Jose, CA); dV-Trainer (Mimic Simulation, Seattle, WA), and the da Vinci Skills Simulator (Intuitive Surgical, Santa Clara, CA). The SEP-Robot is a desktop-like system for training robotic skills in a VR graphical interface. The da Vinci Skills Simulator, also called the “backpack,” is a hardware system that loads VR simulations into the actual da Vinci console. The RoSS and dV-Trainer systems are stand-alone devices with surgical controls resembling those of the da Vinci system. These simulators largely focus on hand-eye coordination, tissue manipulation, suturing, and knot tying. The major benefit of VR simulators for training da Vinci robotic skills is that they produce performance metrics including time, error measures, and motion analysis. These simulators are increasingly being used for training novice surgeons in robotic skills for a variety of surgical specialties.

When assessing simulator validity, researchers have noted that the use of robotic surgery simulators does translate to the clinical environment and the learning curve for initial console training for surgeons is significantly decreased. Unfortunately, the available robotic surgery simulators still come with a high sticker price and varying agreement on the level of fidelity that is currently present in these technologies. These deficiencies are likely due to the early stage of the robotic surgery approach, and it is likely that cheaper and more sophisticated systems will be available in the near future.

**Fundamentals of Robotic Surgery**

The Fundamentals of Robotic Surgery (FRS) is a robotic surgical skills training and assessment program designed to provide a proficiency-based curriculum of basic technical skills to prepare surgeons for performing robotic surgery procedures across a wide range of specialties. The FRS program was developed over a 2-year period by subject matter experts from multiple surgical societies, surgical educational societies, surgical boards, and other governing organizations through a series of four consensus conferences, which included over 80 international robotic surgery experts, behavioral psychologists, medical educators, statisticians, and psychometricians. The multidisciplinary team of experts agreed upon the critical skills and tasks to be included in a comprehensive basic curriculum, and a task deconstruction was performed to identify the tasks, subtasks, and errors that needed to be measured. A modified Delphi methodology was then used to create a matrix of specific robotic surgery tasks, common errors, desired outcomes, and quantitative metrics to
Table 53-5

The four online modules for the fundamentals of robotic surgery curriculum

| Module 1 | Introduction to Surgical Robotic Systems, includes an overview of minimally invasive surgery, advantages of robotic assisted surgery, components of robotic systems, and system functionality |
| Module 2 | Didactic Instructions for Robotic Surgery Systems, provides an overview of robotic surgery systems, as well as detailed information regarding the pre-, intra-, and postoperative phases |
| Module 3 | Psychomotor Skills Curriculum, consists of background and general principles of the psychomotor tasks, an introduction to the physical model on which the tasks are performed, and general scoring guidelines for all the tasks, followed by detailed descriptions of each task, including the targeted primary and secondary skills and metrics |
| Module 4 | Team Training and Communication Skills, includes background on the degradation of situation awareness and the TeamSTEPPS process followed by detailed content covering communication, situational awareness, mutual support, leadership, the preoperative phase, robotic docking, intraoperative phase, postoperative phase, and a review of five scenarios |

support those outcomes. Finally, a second round classic Delphi anonymous rating was used to ensure concurrence, prioritize the task rankings, and eliminate low-scoring tasks.

All trainees must first complete an online curriculum consisting of four modules. Each of the four online modules is followed by a short quiz, requiring a minimum of 70% correct to proceed in the training (Table 53-5). The curriculum also includes a cumulative, cognitive test following completion of all modules.

Following online course completion, trainees must complete seven psychomotor exercises using a surgical robot, a simulated abdomen, and an 18-cm, removable dome model (Fig. 53-7).

The psychomotor exercises consist of the following tasks: docking/instrument insertion, ring tower transfer, knot tying, railroad track, 3rd arm cutting, puzzle piece dissection, and vessel energy dissection (Table 53-6).

Nontechnical Skills

For surgeons, the term “nontechnical skills” refers to the cognitive knowledge and teamwork-related abilities that must be integrated with psychomotor skills and abilities. There are no sharp demarcations between these areas of skill, but different simulation methods are suitable for training each, and all should be addressed.

High-Fidelity Patient Simulation and Team Skills

Recognition of the role of human factors in the occurrence of preventable errors has spawned various efforts to train behaviors conducive to high-performing teams. Crew resource management (CRM) training utilizing simulation has been credited with increased safety in aviation. Lessons learned from CRM have been adapted to medical training with simulators focusing on medical team performance in complex clinical situations. This development came about in the 1990s driven primarily by anesthesiologists responsible for establishing the first high-fidelity simulation environments. These were developed for simulation of crisis-level events where management could be practiced under realistic but safe conditions. Such training could be scaled to involve a single learner to focus on clinical management up to an entire care team able to practice team processes. Development of a program for such training requires an understanding of the principles underlying team effectiveness and the specific characteristics of an expert team. In a 2012 discussion paper on team-based health care emerging from the Best Practices Innovation Collaborative of the Institute of Medicine (IOM) Roundtable on Value & Science-Driven Health Care, such principles were clearly laid out following a careful analysis of effective medical teams across the country (Table 53-7). This document provides an excellent review of characteristics and values that surgical teams can aspire to. The necessity to train these skills has been widely accepted as...
a means of increasing the safety of healthcare. There are methods to accomplish medical team training that do not involve simulation, but high-fidelity patient simulation has proven to be highly effective in increasing health care team competency, and systematic reviews have given evidence based endorsement of this approach.\textsuperscript{138}

Simulation training for communication and other teamwork-pertinent nontechnical skills requires learners to be embedded in realistic scenarios pertinent to a healthcare team’s actual clinical responsibilities where activities and interactions prompted by the simulated clinical circumstances can be practiced and observed. The computer-driven high-fidelity manikin simulator serving as the “patient” at the center of these activities can be monitored and controlled to demonstrate realistic physiology consistent with the clinical condition needed for the scenario.\textsuperscript{139} Software-driven, physiologic changes from the baseline state can occur in response to either manual commands or programmed adjustments to accurately depict, for example, new, ongoing and unexpected clinical developments with blood loss, sepsis, or myocardial ischemia. Basic interventions such as airway management, drug administration and wound care for moulaged body parts can also be performed. Hybrid simulations using both manikin and open abdominal or laparoscopic surgical simulators have also been used to extend scenarios to an operating room setting, with all members of the surgical team engaged in their role-specific tasks.\textsuperscript{136,140,141} These events can be conducted in a dedicated simulation suite or in an actual clinical area where it would be termed in situ simulation.

The postsimulation debriefing is an essential component of simulation-based team training. This is where learning points are reinforced and progress towards desired knowledge, attitudes, and behavior can be developed.\textsuperscript{142,143} Participants are prompted to reflect on the events of the simulation and to openly discuss positive and negative aspects of the experience. The debriefing environment and discussion ought to be open, nonjudgmental, and directed at improvement in individual and overall team performance. A facilitator with strong content knowledge should ensure that the discussion includes identification of gaps between the observed and desired performance. An
effective facilitator is cognizant of the need to keep the debriefing learner-centric and to keep discussion focused on opportunities for improvement. Although debriefing is uniformly viewed as essential to health care team simulation effectiveness and structured debriefing models are frequently cited as highly effective, there is no consensus on which specific methods, including video review debriefings, represent best practices. Quality of implementation is consistently cited as the most important contributing factor to effectiveness of debriefing.

The complexity of team-managed clinical events makes measurement of team performance challenging, but several assessment tools have been developed and used successfully in simulation settings. NOTECHS (Non-Technical Skills) and the NOTSS (Non-Technical Skills for Surgeons) instruments have been used to study nontechnical abilities of individuals in surgical teams. The principal focus of these rating scales is on the quality and effectiveness of situational awareness and communication. Instruments such as the Mayo High Performance Teamwork scale or the surgery-specific OTAS (Observational Team Assessment Scale) place focus on the team dynamics that extend beyond the single team member. These have been used to detect changes in team performance with training. Rosen suggested a framework for a best practices approach to team performance assessment in simulation that details specific applications of measurement techniques in the simulation training environment.

Simulation training directed at nontechnical skills has been shown to improve clinical performance and increase knowledge and attitudes about team functioning as measured in simulated surgical settings such as the trauma bay and the OR for interdisciplinary surgical teams and for surgical trainees. Two systematic reviews of simulation training to increase team skills in the operating room consistently bear out this result, but both also cited the small number of studies where simulation training effects were investigated in the clinical OR setting and lack of evidence of improved clinical outcomes with such training.

### ERROR PREVENTION

Error avoidance and prevention are the overarching goals of surgical care and the time-honored focus of surgical training. Faculty who are responsible for training the next generation of highly qualified surgeons must facilitate the delicate balance between resident autonomy and patient safety. From a training perspective, gradual increases in patient responsibility, autonomous decision-making, and operative action provide a critical opportunity for independent hands-on performance, critical thinking, and action-based skill assessment. However, resident autonomy must be balanced with the goal of delivering high-quality, error-free patient care. A key component to achieving residency training goals includes exposure to techniques and strategies for avoiding errors. The relationship between errors and patient safety is well established in the literature and is also the cornerstone of the case review process for surgical morbidity and mortality conferences. Simulation technology allows trainees the opportunity to execute a variety of tasks and procedures while also experiencing the cognitive demands of surgery, including error correction and surgical planning decisions. Fig. 53-8 presents a framework for categorizing surgical errors that may take place on the cognitive-motor continuum.

Simulation-based curricular approaches to exposing and training surgical errors include (a) the use of error-enabled

<table>
<thead>
<tr>
<th>Cognitive</th>
<th>Cognitive-motor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error prevention</td>
<td>Error recognition</td>
</tr>
<tr>
<td>What should you be concerned about?</td>
<td>Which of these may result in an error?</td>
</tr>
<tr>
<td>What will prevent an error?</td>
<td>Estimate error risk for: technical approach and patient selection</td>
</tr>
<tr>
<td>Knows - the anatomy; right technique; natural hx of disease; how to avoid an error</td>
<td>Identify - proper technique; potential error: error risk</td>
</tr>
<tr>
<td>Select - the right operation; the right stitch; the right instrument</td>
<td>Knows - surgical options</td>
</tr>
<tr>
<td>Example: Plans ahead of time to adjust tension during knot tying based on tissue type</td>
<td>Example: Recognizes that the first knot was a little tight then makes an adjustment to get the second knot just right</td>
</tr>
<tr>
<td>Example: First knot pulled through</td>
<td>Rescue: Freshens tissue edge, places a new stitch</td>
</tr>
</tbody>
</table>

Figure 53-8. A framework for understanding error prevention, error recognition, and error rescue/correction during both the cognitive phase as well as the cognitive-motor phase where there is fluid integration and updating of motor and cognitive decisions. This fluid integration allows for an error loop, as each decision is based on the results of each motor action in a dynamic fashion where an error can take place with each decision or action. (Reproduced with permission from Pugh CM, Santacaterina S, DaRosa DA, et al: Intra-operative decision making: more than meets the eye, J Biomed Inform. 2011 Jan;44(3):486-496.)
simulations, where the learner can make any of a variety of errors during the course of the task; and (b) the use of forced-error simulations, where the learner experiences a specific, usually unexpected error and demonstrates error management as well as options for correction. Error management is a human factors process that includes making, detecting, and correcting errors.\textsuperscript{160,161} Several studies show that there is wide variation in the timing and type of operative decisions that residents make when interacting with an error-enabled simulation.\textsuperscript{112,162-164} Specific error training simulations have been developed using a cognitive task analysis approach to guide the design of error scenarios. Cognitive task analysis involves the use of structured interviews to explore how experts approach the error management process, including intraoperative decision-making and technical approaches.\textsuperscript{165,166} The error-enabled approach allows for a broad assessment of learning needs. The forced-error approach allows instructors to engage in specific error management skills and metrics. Error-enabled and forced-error surgical simulation trainers have been used in a variety of research and training modules to allow observation and assessment of surgical residents as they independently perform operative procedures and practice the error management process.\textsuperscript{112,166,167}

Another approach to error training and assessment is the use of assessments that focus on errors. Current assessment tools for surgical skills include task-specific and global rating scales, final product analysis, and documentation of critical failures.\textsuperscript{160,168} These existing tools for assessing surgical residents largely focus on manual techniques and procedure time and do not capture the cause of resident performance failures.\textsuperscript{169} Use of error-related theories in human factors allows for the development of error-centric skills assessments.\textsuperscript{161,170,171} One study used previously developed human error classifications, including omission versus commission and cognitive versus technical as a means of assessing surgical residents.\textsuperscript{162} Omission errors were defined as failure to perform a step entirely. Commission errors represented failure to perform a step correctly. For example, failure to measure the hernia defect was categorized as an omission error, whereas measuring the hernia defect with an inaccurate method was categorized as a commission error. Errors in information, diagnosis, and strategy were categorized as cognitive, and errors in action, procedure, or mechanics were classified as technical. Use of assessment surveys with this type of differentiation allows for development of error metrics in surgery and focused error training and feedback.\textsuperscript{172}

### SIMULATION AND PATIENT OUTCOMES

Surgical simulation training is intended to make patient care safe and free of avoidable errors and to maximize opportunities for good clinical outcomes. The Kirkpatrick four-level scale (Table 53-8) characterizes educational intervention effects, including those that might improve surgeon performance in clinical settings (Kirkpatrick level 3) or those that might actually improve patient outcomes (Kirkpatrick level 4).\textsuperscript{173} Despite the expanding use of simulation in proficiency-based practice models over at least 15 years, there are surprisingly few studies of simulation-based surgical training that can be described as Kirkpatrick level 4. On the other hand, there are numerous studies showing that clinician performance during the course of clinical care is improved after simulation training (Kirkpatrick level 3). As detection of changes in clinical outcomes can be quite challenging, it may be difficult to isolate simulation training effects from numerous other factors that can also affect patient outcomes. When looking specifically at a low-frequency complication event such as bile duct injury with laparoscopic cholecystectomy, the detection of a small, positive training effect is statistically improbable. However, the use of available Kirkpatrick level 3 data to support assumptions about training benefits is fully supportable given the preponderance of literature showing such benefit.

The best current evidence for improved patient outcomes with simulation is that of technical and cognitive training for central venous catheter (CVC) insertion. In 2009, Barsuk et al monitored catheter-related bloodstream infection incidence in an ICU setting over a 32-month period before and after institution of proficiency-based simulation training. After simulation-trained medical residents began performing CVC insertion, an 85% reduction in these infections was observed (3.2 per 1000 catheter-days reduced to 0.50 infections per 1000 catheter-days).\textsuperscript{174} Subsequently, the same group reported that the financial savings realized with these improved outcomes amounted to a 7:1 return of the investment for the training.\textsuperscript{175} In a similar single unit observational study, Burden et al reported a reduction in catheter-related bloodstream infection incidence from 6.47 per 1000 catheter days to 2.44 per 1000 catheter days after training intervention and comparable financial savings with shorter ICU and hospital stays.\textsuperscript{176} Single cohort studies have inherent weaknesses, but a somewhat smaller randomized trial of simulation training versus traditional apprenticeship model-trained controls also showed this reduced infection incidence (1.0 vs. 3.4 infections per 1000 catheter-days, respectively).\textsuperscript{177}

Riley and colleagues conducted a study in three small community hospitals, administering TeamSTEPPS (Team Strategies and Tools to Enhance Performance and Patient Safety) didactic team training to perinatal care teams at one hospital, TeamSTEPPS with an accompanying program of in situ simulations at a second, and no intervention at the third hospital, which

<p>| Table 53-8 |</p>
<table>
<thead>
<tr>
<th>Kirkpatrick level scale of educational outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>KIRKPATRICK LEVEL</strong></td>
</tr>
<tr>
<td>Level 1</td>
</tr>
<tr>
<td>Level 2</td>
</tr>
<tr>
<td>Level 3</td>
</tr>
<tr>
<td>Level 4</td>
</tr>
</tbody>
</table>

served as the study control. The simulation interventions were designed to involve triage, labor and delivery, and OR components of care. Perinatal outcomes were assessed using the Weighted Adverse Outcomes Score (WAOS). Simulation training resulted in improved clinical results with a 37.4% decrease in WAOS observed in the simulation intervention group, while similar benefit was not observed in either the didactic-trained or control groups.

TeamSTEPPS with supplemental simulation training was also studied by Capella and colleagues for team performance and patient outcomes in trauma care. Trauma resuscitations for successive 2-month periods separated by didactic and scenario-based simulation training to subjectively-defined proficiency levels (33 before training, 40 after training) were assessed in multiple teamwork domains including overall performance using the TPOT (Trauma Team Performance Observation Tool). Significant improvements were observed post training in leadership, situational monitoring, mutual support, communication, and overall scoring. In addition, time to OR was also observed to decrease. Steinemann and colleagues reported on a larger number of trauma resuscitations, 141 before and 103 after in situ trauma bay patient simulation training with accompanying didactic instruction. Although significant improvements in patient outcomes were not observed in this study, resuscitation time did decrease following training.

In a comprehensive review of literature examining simulation impact on patient outcomes, Zendejas identified 50 studies purporting to show patient outcomes. For the majority of these, the quality of clinical outcomes evidence was considered low, and appropriate validity data were reported for results in only a small number of published reports. In a subsequent review of 1328 articles resulting from an exhaustive literature search, Cox identified 12 individual articles that reported sufficient patient outcomes data to be considered Kirkpatrick level 4. Concerns regarding the relatively few studies of this nature are primarily methodological and relate to the preponderance of use of the observational pre- and post-training study model with its inherent bias risk, rather than randomized controlled trials. These studies have also tended to be quite limited in size. In surgery, the general areas of surgeon performance that have lent themselves to study of educational outcomes in the clinical setting after simulation training include technical abilities and observable behaviors during team-based activities. The use of decreased operative time or technical errors as clinical outcomes after simulation training might suggest, but do not clearly establish, a patient safety or other efficacy benefit. However, these metrics have frequently been used to make the case for the potential for such benefits in skills transfer studies for procedures such as laparoscopic cholecystectomy and laparoscopic hernia repair. There is a clear need to obtain high quality evidence of how proficiency-based simulation training impacts surgical patient outcomes.

In a recent study examining intern response to pediatric codes, simulation training was shown to significantly reduce the time to request help, to initiate bag-mask ventilation, and to initiate chest compressions. Citing the rarity of pediatric codes, investigators measured this effect in in mock code situations using patient simulators. This use of a high-fidelity simulation environment as a surrogate for clinical events, that might otherwise be difficult to observe, may be the only practical model available for the study of provider performance outcomes with educational interventions.

**Simulation Training for the Practicing Surgeon and Maintenance of Skill**

Given the current requirement for FLS and FES certification in residency, some authors have posed the important question of whether such certification should be applied more broadly to surgeons in practice. In a 2012 press release, both the ACS and SAGES made the recommendation that all surgeons performing laparoscopy obtain FLS certification. The potential to lower malpractice litigation risk under a self-insurance model was used as justification to certify 37 surgeons in FLS in the Harvard system in 2009. In a recent report, surgical oncologists new to minimally invasive inguinal lymph node dissection (MILND) were FLS tested prior to performance of their first clinical procedures, which were video-assessed using the GOALS scoring method. The FLS score was shown to correlate with both GOALS results and operative time but not lymph node yield. However, evidence that a lab-based technical and cognitive skills test predicts observed operative technical skill in practicing surgeons is promising and warrants investigative follow-up.

At the present time, the American Board of Surgery’s requirements for Maintenance of Certification (MOC) do not specifically include any certification of technical skills that might use surgical simulation. Nonetheless, simulation training and testing can be made available to surgeons in practice to provide an avenue for specific training. Although most investigations of VR use for laparoscopy have examined basic skills acquisition in the lab setting far in advance of any measured impact in the clinical OR, a recent innovative study of the use of a VR simulator for “warm-up” practice immediately before a procedure showed that this improved OR performance. As new surgical procedures and technologies are introduced to clinical practice, simulation training solutions could serve just as important a purpose to prepare for these as simulation training methods currently serve for laparoscopic surgery. In a recent publication relating to practicing surgeons, Sullivan et al. provide a framework for development of simulation-based certification models for both trainees and faculty surgeons.

**Future Considerations**

Simulation-based training and assessment is firmly established in surgical education, especially in graduate medical education where learners have been more extensively studied than any other simulation user group. The immediate future of simulation in surgery will likely see expanded use of proficiency-based training given the consistent demonstrations of effectiveness in improving surgeon skills and improved educational outcomes as measured in clinical settings.

The question of what types of simulation-based assessments and training activities might be possible raises questions of where technological advances might open new opportunities. The advancement of virtual reality is inevitable, and much richer virtual experiences entering the mainstream in surgical education seems likely in the near future. Three-dimensional printing technology has been growing in its use and applications and has greatly facilitated the development of anatomically accurate bench top simulations for complex surgical procedures. Some of the materials allow for elegant instrument-based dissections, including the use of electrosurgery.
Advances in wearables, motion tracking, and sensor technologies allow for a wide variety of hybrid and augmented experiences in simulation as well as extensive opportunities for the development of new performance metrics. One study using sensor technology for evaluating clinical breast examination skills noted key performance differences in experienced physicians. Both the sensor data and sensor-guided video analysis allowed for skill quantifications that were previously unknown but critical to performance excellence.\(^{103-106}\)

Opportunities for remote collaboration are now greatly improved with higher internet speeds, improvements in augmented reality technology, and ever-increasing camera resolution. VIPAR (Virtual Interactive Presence and Augmented Reality) allows for the visual field of a surgeon to be converted to a simulation and projected in a remote location.\(^ {107-108}\) As such, the system allows for intraoperative collaboration and telementoring.\(^ {109}\)

Irrespective of what new simulation and engineering technologies emerge, these technologies are here not just to stay but to grow as assessment and educational tools. This presents abundant opportunities for simulation leaders in surgery to improve the delivery of care by defining best practice in simulation applications and keeping step with current and future changes in surgical practice.

REFERENCES

Entries highlighted in bright blue are key references.


### Web-Based Education and Implications of Social Media

**Lillian S. Kao and Michael E. Zenilman**

#### INTRODUCTION

Surgical education has changed significantly over the past two decades. Disruptive forces such as work hour restrictions and the advent of laparoscopy have forced educators to rethink how and where to teach residents. Technologies, including the internet and web-based applications, have further enabled educators to redesign surgical education (Fig. 54-1). The internet has become an integral tool not just in surgical education but also in Americans’ lives by changing the way that people communicate with each other, access information, and conduct their daily lives. Today, almost 9 in 10 American adults use the internet. Furthermore, the internet has revolutionized education by allowing for expanded reach, asynchronous learning whereby students and instructors do not have to be on the same time schedule, and multimedia materials.

Like internet usage, social media has seen a rise in adoption over the past decade. Social media is a term that encompasses multiple computer-mediated platforms that are used for creating and sharing information, ideas, and other content. Social media facilitates communication and interactions across virtual networks. Commonly used platforms include Facebook, Twitter, Snapchat, and Instagram. Social media can be used for multiple purposes including social and professional networking; however, this chapter will focus on its uses in surgical education.

#### WEB-BASED EDUCATION

Web-based educational resources include lectures and webinars, simulators, assessment tools, and interactive mentoring and coaching. Furthermore, entire web-based curricula have been developed that can link to online resources such as journal articles, interactive anatomy modules, and videos of operations. There are multiple advantages to web-based education. For surgical trainees, web-based educational materials allow access regardless of time of day or night, provide interactive tools for learning (i.e., anatomy), and videos for viewing operations and procedures. Furthermore, for practicing surgeons, web-based educational resources include forums for sharing challenging cases and procuring advice, activities for obtaining continuing education, and rapid access to information about new technologies and research. However, there may also be disadvantages in terms of costs and technical problems.

Multiple studies have evaluated learning outcomes after implementation of web-based educational interventions. In surgical education, these interventions may be used to teach patient care and decision-making via online case studies, convey knowledge using online didactic materials, or introduce surgical skills. However, studies evaluating these interventions tend to be nonrandomized, small, and single center. A 2008 systematic review and meta-analysis by Cook et al evaluated the effect of internet-based learning across healthcare in general. The review suggested that internet-based learning is better than no intervention but has similar effectiveness as traditional educational methods. A more recent 2015 systematic review by Jayakumar et al focused on web-based education in surgery. They reported a positive effect, but the majority of studies included in the review lacked a control. Based on the current literature, the internet should be considered one tool among many that can facilitate learning. However, further studies are necessary to identify the key elements that improve effectiveness. Web-based educational materials should be developed keeping adult learning theories and principles in mind.

Web-based surgical curricula have been developed both at an institutional and at a national level. A widely-used curriculum is the Surgical Council on Resident Education (SCORE) curriculum, which is available via an online portal. Developed in 2006, SCORE is based on the six core competencies required of a graduating resident: patient care, medical knowledge, professionalism, communication, practice-based learning, and systems-based practice. SCORE is the result of an ongoing collaborative effort of the American Board of Surgery, American College of Surgeons, American Surgical Association, Association of Program Directors in Surgery, Association for Surgical Education, Residency Review Committee for Surgery of the Accreditation Council of Graduate Medical Education, and Society of American Gastrointestinal and Endoscopic Surgeons. The SCORE curriculum provides content for topics to be covered during a 5-year general surgery residency and is adding fellowship-level content as well. The SCORE Portal modules for each topic include learning objectives, discussion
Questions, text resources and videos, and self-assessment quizzes. Although improvements in quality examination performance among residencies that subscribe to SCORE are promising, no studies have definitely demonstrated that SCORE use improves resident knowledge, skills, or clinical performance.

Web-based education can also be used for assessing and teaching surgical skills. In 2013, Birkmeyer et al performed a study that correlated surgical skills in bariatric surgery, based on blinded reviews of videotaped operations, to clinical outcomes. The ability to discriminate surgeons with good and poor technical skills using video-based assessments has significant implications for training surgeons and for evaluating their performance. With regards to training, multiple web-based and virtual reality simulators have been developed that allow residents to practice tasks and skills repetitively at their own pace and on their own time. These simulators can quantify efficiency of motion and time to complete a task as well as provide real-time feedback. These metrics have been demonstrated to have construct validity (in that they measure what they are supposed to be measuring) and criterion validity (in that they correlate with operative performance). For example, randomized trials have demonstrated that surgical simulation training correlates with decreased operative time and improves subjectively rated performance on technical skills in the live setting. Web-based assessments have also been used in combination with physical simulators to provide similar metrics.

With regards to surgeons in practice, there has been increased enthusiasm for the use of video-based coaching to complement intraoperative teaching. In particular, postoperative review of videotaped procedures allows surgeons to receive individualized feedback about opportunities for improvement without the time constraints or pressures of the operating room. Randomized trials of surgical coaching in simulated settings suggests benefits over traditional simulator training, and larger trials in a live setting are ongoing. Although most coaching occurs face-to-face, there are opportunities to use web-based coaching. As an example, telementoring has been used to mentor surgeons in the operating room, even across the globe.

Barriers to web-based education include the up-front costs for development and the need for technical expertise. As already noted, web-based education for teaching knowledge may not be more effective than traditional methods. On the other hand, simulation and video-based coaching hold significant promise in improving training and assessment of surgical skills. However, widespread implementation of video-based coaching will require a culture shift for surgeons to accept assistance and resources such as time, availability of coaches, and finances. Ongoing studies will provide data regarding the effectiveness of these educational strategies.

**Social Media—Based Education**

**What is Social Media?**

Social media is a term describing websites and web-based applications that enable users to share ideas, information, and
content through virtual networking. Although social media is often used to interact with friends and family, social media can also be used for educational and professional purposes. Examples include Twitter-based journal clubs, Facebook-based discussion forums, and professional networking sites such as LinkedIn or ResearchGate. Social media platforms can serve different purposes including social networking, microblogging, blogging, photo sharing, video sharing, and crowdsourcing.

Commonly used social media platforms in surgery include Facebook, Twitter, and YouTube. Facebook is the most popular social networking site; it can be accessed via desktops, laptops, and mobile phones. It allows users to exchange information, photos, and videos with specified contacts or “friends” with whom there is a two-way relationship. Twitter is another popular social media platform. It is a microblogging site that, like Facebook, allows exchange of messages and photos but limits messages or tweets to 140 characters or less. Twitter users may have one or two-way relationships with other users. Followers of a user receive all of that person’s tweets in their Twitter feed. YouTube is a social media platform that allows users to share videos.

Social media usage in surgery may be unidirectional or bidirectional (Fig. 54-2). For example, journals such as the New England Journal of Medicine may have a large number of followers but may be following very few users back. Given that the number of journal articles published daily has risen exponentially, particularly with the advent of open access journals, keeping up with the surgical literature can be overwhelming. Following journals on social media is one strategy for staying updated. Although conventional media outlets such as newspapers and news channels may draw attention to practice-changing studies, social media is another platform by which such information can be promoted and disseminated by journals. Many surgical journals have an online and social media presence, and many have social media editors who curate the posted materials. Popular social media platforms for journals include Facebook and Twitter. Both platforms allow journals to post text, figures, and links to abstracts or journal articles. Both platforms allow others to share information or comment on articles. However, Twitter restricts text to 140 characters. A recent innovation that may counteract the limited number of allowed characters is the visual abstract, which is a concise pictorial representation of an article’s key points (Fig. 54-3). Recently, a prospective, case-control crossover study was performed whereby tweets about articles from Annals of Surgery were either accompanied by a visual abstract or tweeted with text alone. Accompaniment of a tweet with a visual abstract resulted in a threefold increase in article visits. Thus, the majority of journal followers may merely receive the information about new publications (unidirectional flow of information). However, users may also choose to respond to posts with comments (bidirectional flow of information).

**Figure 54-2.** Social media usage can be (a) unidirectional or (b) bidirectional. User names on Twitter are denoted by “@”. A. @Twitteruser2 is following @Twitteruser1. She is receiving all of his messages in her Twitter feed. However, @Twitteruser1 does not follow her back and therefore does not receive her messages in return. B. @Twitteruser1 and @Twitteruser2 follow each other. Therefore, they each receive each other’s messages in their Twitter feeds.

**Figure 54-3.** A visual abstract is a graphical summary of the main results of a journal article.
Journal Clubs

Journals may promote bidirectional flow of information by hosting social media–based journal clubs. From an educational standpoint, journal clubs have traditionally served not only as an adjunct to lectures but also as a forum to teach about critical appraisal of the literature. Furthermore, when facilitated by faculty with clinical expertise on the subject being discussed, surgical trainees can better evaluate how to incorporate the evidence into practice. However, barriers to traditional journal clubs may include poor participation, lack of a convenient time, or absence of local expertise in either the clinical topic or research methodology. Social media–based journal clubs can help to overcome these barriers by allowing for asynchronous discussion and expert moderators. Online journal clubs can be carried out in real time, but they also allow respondents to comment hours or even days later to a conversation. Multiple specialties, including surgery, have developed social media–based journal clubs. Tips for successfully launching an online journal club can be garnered from the expanding experience with them (Table 54-1).

Social media–based journal clubs in surgery have been conducted via Facebook, Twitter, or a combination of the two platforms. They have also taken the form of a blog. Although commonly associated with personal journals or diaries, blogs can also be found on professional websites that are updated frequently by a person or group (i.e., by a journal or surgical society). Conversations from journal club discussions can also be compiled and summarized into a transcript either manually or using web-based applications such as Storify. These transcripts can be posted on the journal website or shared. Furthermore, summaries from either in-person or online journal clubs can be added to the PubMed citation via PubMed Commons.

Preliminary data suggests that online journal clubs increase discussion about articles, views of the abstract, and downloads (Fig. 54-4). For example, the International General Surgery Journal Club held four moderated discussions of journal articles on Twitter between March and June 2014. The reviewed articles covered topics relating to bariatric surgery (March), venous thromboembolism in trauma (April), diverticulitis (May), and contralateral prophylactic mastectomy for breast cancer (June). Although the authors and invited experts only moderated discussions for 3 days, Twitter activity increased in the days preceding and following these discussions. Furthermore, daily views of the article and downloads increased correspondingly. Thus, online journal clubs are a potential strategy for increasing surgeon education about seminal articles.

While social media–based journal clubs hold much appeal, there is a paucity of data regarding their effectiveness in teaching participants about critical appraisal skills. A multicenter randomized trial compared journal clubs moderated by a faculty member to online discussions. Surgical residents in both arms utilized modules developed by the Evidence Based Reviews in Surgery Steering Committee; these modules include the relevant guide to critical appraisal and a methodological and clinical review. Residents randomized to the moderated group scored higher on a validated test evaluating critical appraisal skills. Further study is required to assess the effectiveness of social media–based journal clubs in disseminating new knowledge as well as in teaching critical appraisal.

### Table 54-1

<table>
<thead>
<tr>
<th>Tip</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Create an online home page that serves as a launching pad for your journal club discussions.</td>
<td></td>
</tr>
<tr>
<td>2. Develop and register a hashtag on Twitter.</td>
<td></td>
</tr>
<tr>
<td>3. Incorporate not only Twitter but other collaborative platforms as part of your online journal club.</td>
<td></td>
</tr>
<tr>
<td>4. Ensure that the time of the journal club is convenient for your target audience.</td>
<td></td>
</tr>
<tr>
<td>5. Help prepare participants by aggregating other online resources relevant to the article.</td>
<td></td>
</tr>
<tr>
<td>6. Consider inviting the authors of the featured article or other experts in the field.</td>
<td></td>
</tr>
<tr>
<td>7. Suggest journal club participants consider using specific Twitter management applications during the Tweet chat.</td>
<td></td>
</tr>
<tr>
<td>8. Engage the participants to cultivate and incentivize more discussion.</td>
<td></td>
</tr>
<tr>
<td>9. Connect to the online community by following and engaging with other relevant social media accounts.</td>
<td></td>
</tr>
<tr>
<td>10. Link back to the original paper by inserting a comment on PubMed Commons.</td>
<td></td>
</tr>
</tbody>
</table>


Live-Tweeting Conferences

Traditionally, surgeons have attended regional and national conferences to network, learn new information relevant to their practice, and exchange ideas. However, with the advent of social media, surgeons no longer have to physically attend a conference to perform all those activities. Live-tweeting is a term used to describe the posting of comments on Twitter about an event while it is ongoing. Multiple surgical and nonsurgical societies have adopted Twitter to expand the reach of their conferences. By denoting tweets as emanating from a specific conference with a unique hashtag, the reach and number of impressions can actually be measured (Fig. 54-5). The reach refers to the number of unique recipients of messages from a specific group of Twitterers (or people posting on Twitter). Impressions refers to each time a message was delivered to a recipient; a recipient may receive the same message more than once. Neither reach nor impressions measure whether the recipient read the tweet. As an example of how reach and impressions can be used to provide metrics for social media, the Healthcare Hashtag Project allows registered conference hashtags to track the latest tweets, the most prolific Twitterers of conference-related tweets, the most commonly mentioned Twitterers, and the number of impressions (Fig. 54-6) (https://www.symplur.com/healthcare-hashtags/). As noted in the figure, social media rapidly and exponentially increases the spread of information. Transcripts of conference-related tweets can also be assembled to allow a conversation thread to be organized into a cohesive discussion.

Interactive Forums and Communities

Interactive forums and communities are another method by which both the internet and social media can be used for
Figure 54-4. Impact of a social media–based journal club on Twitter activity, hypertext markup language (HTML) views, and portable document format (PDF) downloads. A. Activity during a Twitter-based journal club such as the International General Surgery Journal Club (IGSJC) can be tracked by denoting the journal club related tweets with a hashtag (#IGSJC). Twitter activity increased during each of four journal clubs. B. Daily HTML views and PDF downloads of featured articles also increased around the time period of the four Twitter journal clubs. (Unpublished data from Sarah Bryczowski and Michael E. Zenilman.)

Figure 54-5. Difference between reach and impressions in Twitter. If there are two Twitter users and each has three unique followers and one shared follower, then there are seven unique recipients of their combined tweets. Their total reach is seven unique users. If both users tweet the same message, then one user will have received the message twice. However, each time the message was delivered counts as an impression; thus, the followers will have a total of eight impressions.
The #Surgery Conference Influencers

<table>
<thead>
<tr>
<th>Top 10 by Mentions</th>
<th>Top 10 by Tweets</th>
<th>Top 10 by Impressions</th>
</tr>
</thead>
<tbody>
<tr>
<td>@User1 123</td>
<td>@User3 27</td>
<td>@User13 103,937</td>
</tr>
<tr>
<td>@User2 65</td>
<td>@User11 25</td>
<td>@User24 80,956</td>
</tr>
<tr>
<td>@User3 62</td>
<td>@User7 23</td>
<td>@User42 42,976</td>
</tr>
<tr>
<td>@User4 46</td>
<td>@User4 23</td>
<td>@User3 40,911</td>
</tr>
<tr>
<td>@User5 42</td>
<td>@User12 19</td>
<td>@User6 28,966</td>
</tr>
<tr>
<td>@User6 32</td>
<td>@User13 13</td>
<td>@User14 22,534</td>
</tr>
<tr>
<td>@User7 21</td>
<td>@User14 12</td>
<td>@User52 14,566</td>
</tr>
<tr>
<td>@User8 19</td>
<td>@User15 11</td>
<td>@User72 13,514</td>
</tr>
<tr>
<td>@User9 17</td>
<td>@User2 10</td>
<td>@User35 11,965</td>
</tr>
<tr>
<td>@User10 16</td>
<td>@User1 9</td>
<td>@User68 9,862</td>
</tr>
</tbody>
</table>

The Numbers
- 443,601 Impressions
- 298 Tweets
- 72 Participants
- 2 Avg Tweets/Hour
- 4 Avg Tweets/Participant

Figure 54-6. Example of conference analytics from Healthcare Hashtags (https://www.symplur.com/healthcare-hashtags/). Surgery conferences can tag tweets by using a prespecified hashtag (i.e., #SurgeryConference) to denote conference-related messages. These can then be tracked. A mention occurs when a user includes another user’s name in the tweet. Note in the example that even if there are only a few users tweeting about a conference, the number of impressions can be large if several of those users have a large number of followers. For example, @User13 only tweeted 13 times but had 103,937 impressions.

PITFALLS IN WEB AND SOCIAL MEDIA–BASED EDUCATION

Despite the many advantages of web and social media–based education, significant potential pitfalls still remain. First, widespread adoption and utilization is a challenge. Although internet and mobile phone usage is prevalent, effectiveness of web-based educational materials is dependent upon trainees’ and surgeons’ uptake of the technology and available resources. Despite increasing availability of the internet and mobile technology, not all surgeons use it for educational purposes. For example, a systematic review and meta-analysis by Guraya et al found that three out of four medical students use social networking sites, but only one out of five uses them for educational purposes. Furthermore, lack of utilization can impact educational effectiveness. In the previously mentioned multicenter trial comparing a moderated journal club to an online version, low participation in the Internet journal club was postulated to be a significant factor in the poorer performance on a critical appraisal test. More attention to instructional design may improve uptake and effectiveness. A systematic review and meta-analysis by Cook et al determined that features such as interactivity, practice exercises, repetition, and feedback are associated with improved learning outcomes. The authors also noted that the evidence base upon which to design internet-based learning programs is limited by poor study methodology, failure to use conceptual frameworks, and lack of adherence to reporting standards. Thus, while the aforementioned features should be considered in designing future web and social media–based educational programs, further evidence-based guidance is needed.

Second, the quality of information available online and via social media may not be accurate or reliable. Multiple reports have been published regarding the inaccuracies of web-based educational materials. Surgeons and surgical trainees should carefully evaluate the source of educational material, search for conflicts of interest that may result in biased information, assess how recently the information was updated, and cross-check references. The lay public may have more difficulty in identifying trustworthy surgical educational materials on the internet. Healthcare providers should guide patients to reputable websites and to encourage discussion regarding the accuracy of the content.

Third, useful dialog and advice about difficult cases must be balanced with ethical considerations surrounding patient confidentiality and privacy. Appropriate safeguards must be taken...
to ensure that patients cannot be identified based on provided information, that patients have consented to have their information posted anonymously, and that all case-related comments are appropriate and professional. While common sense should be utilized in posting about patient cases, only a few organizations have published guidelines for how to safeguard against potential pitfalls. In 2013, the American College of Physicians Ethics, Professionalism, and Human Rights Committee; the American College of Physicians Council of Associates; and the Federation of State Medical Boards Special Committee on Ethics and Professionalism published a position statement about online medical education (Table 54-2). The paper stated that “Maintaining trust in the profession and in patient–physician relationships requires that physicians consistently apply ethical principles for preserving the relationship, confidentiality, privacy, and respect for persons to online settings and communications.”

Fourth, conflicts of interest must be clearly stated. Journals require authors to declare relevant conflicts of interest, but multiple studies suggest that these often go unreported. Similarly, conflicts of interest should be disclosed on social media. However, such disclosures may be more difficult on social media due to the limited number of allowable characters (i.e., 140 characters for Twitter) or to the way information is propagated. For example, a surgeon may disclose an industry relationship on an original tweet, but the disclosure may not appear in subsequent comments of a discussion thread. Surgeons posting on social media must make it their ethical and professional obligation to disclose their conflicts of interest. Furthermore, users of social media content must be aware of the potential for bias introduced by undisclosed conflicts of interest and perform due diligence in assessing the reliability of the source. Lastly, regulatory bodies and professional organizations should publish standardized guidelines for disclosing on social media or develop mechanisms by which disclosure can be publicly accessed (such as the Open Payments database).

Fifth, professionalism must always be maintained. Surgeons posting content on the web or on social media must be aware that information will be widely disseminated and available for posterity; messages posted on social media cannot be fully retracted. As already mentioned, surgeons must strive to maintain patient privacy, ensure accuracy of information, and disclose conflicts of interest. Furthermore, surgeons must be aware of unintentional interpretations of messages (i.e., as discriminatory or unprofessional). Multiple studies of healthcare providers’ social media sites have identified potentially and clearly unprofessional content; these studies have included medical students, residents, and practicing surgeons. Despite the prevalence of unprofessional content, few surgical residency programs have formal institutional social media policies. Furthermore, the American College of Physicians and Federation of State Medical Boards position statement only addresses a few of the issues surrounding web-based activities including for patient and physician education (Table 54-3).

### Table 54-2
The American College of Physicians Ethics, Professionalism and Human Rights Committee; the American College of Physicians Council of Associates; and the Federation of State Medical Boards Special Committee on Ethics and Professionalism published a position paper on online medical professionalism

<table>
<thead>
<tr>
<th>Position</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position 1</td>
<td>Use of online media can bring significant educational benefits to patients and physicians, but it also pose ethical challenges. Maintaining trust in the profession and in patient–physician relationships requires that physicians consistently apply ethical principles for preserving the relationship, confidentiality, privacy, and respect for persons to online settings and communications.</td>
</tr>
<tr>
<td>Position 2</td>
<td>The boundaries between professional and social spheres can blur online. Physicians should keep the two spheres separate and comport themselves professionally in both.</td>
</tr>
<tr>
<td>Position 3</td>
<td>Email or other electronic communications should only be used by physicians in an established patient–physician relationship and with patient consent. Documentation about patient care communications should be included in the patient’s medical record.</td>
</tr>
<tr>
<td>Position 4</td>
<td>Physicians should consider periodically “self-auditing” to assess the accuracy of information available about them on physician-ranking websites and other sources online.</td>
</tr>
<tr>
<td>Position 5</td>
<td>The reach of the internet and online communications is far and often permanent. Physicians, trainees, and medical students should be aware that online postings may have future implications for their professional lives.</td>
</tr>
</tbody>
</table>


### Table 54-3
Online medical professionalism and education

**POSITIONS ON PROFESSIONALISM IN ONLINE PATIENT AND PHYSICIAN EDUCATION**

- The Internet can be a powerful tool for education.
- Physicians should guide patients to high quality online resources that are accurate and objective. These sites should have peer-reviewed content or have verifiable mechanisms for quality control of information.
- Online resources for learning can be used by patients and physicians.
- The internet and social networking can be used to improve public health. Physicians engaged in online communities should ensure the security of the networks and restriction of participation to verified users. Clinical scenarios should not contain any personal identifying information, and patient consent should be obtained before sharing the vignette.
- Discussion of frustrations online undermines trust and professionalism and should be avoided.

Lastly, studies evaluating the effectiveness of web-based education have had varied results. A recent systematic review by Taveira-Gomes et al assessed 251 articles using computer-based learning methodologies in medical education; the number of articles on this topic has increased over time. The most commonly used metrics for evaluating the effectiveness of these methodologies were assessments of knowledge, attitudes, and skills. The majority of studies reported positive effects on these outcomes, although the more rigorous studies (i.e., randomized trials) were less likely to find a positive effect. Online activity (i.e., number of posts or views) was tracked in a few studies, but results were conflicting regarding whether increased engagement correlated with improved performance. This review suggests that high-quality studies are needed of web-based educational interventions and that these studies need to include measures of clinical performance and outcomes.

**IMPLICATIONS AND FUTURE DIRECTIONS**

The Society of University Surgeons’ Social and Legislative Committee issued a position statement entitled: “Social media is a necessary component of surgery practice.” Given the rapid pace with which technology is advancing and the familiarity of the current generation (Generation Z) with the internet, surgeons have no choice but to harness the power of the internet and social media or risk being left behind. Many journals are phasing out print versions, and several journals are already online only. Furthermore, textbooks may also become a relic from the past as publishers move towards developing digital versions that include interactive graphics, audio, and video.

As surgical education continues to evolve, future directions may include broader indications for utilization of web- and social media-based resources. For example, video-based coaching with face-to-face discussions have been used postoperatively to supplement intraoperative teaching. However, preoperative crowdsourcing in planning a challenging case or intraoperative video telementoring are other applications of video-based coaching.

Future research efforts should focus on identifying the most effective formats and components of web and social media–based educational interventions, using rigorous methods to compare educational methods, and measuring clinical outcomes. Moreover, standardized guidelines should be instituted in order to safeguard against ethical and professional misconduct.

**BIBLIOGRAPHY**


Bresnahan ER, Huynh DTK, Jacob B. Social media and education in hernia repair. In: Hope WW, Cobb WS, Adrales GL, eds. *Textbook of Hernia.* Cham: Springer International Publishing; 2017:373-379. This book chapter describes the International Hernia Collaboration, which is a Facebook-based community of practicing surgeons, trainees, and industry representatives. The goals of the collaboration include exchange of ideas and advice regarding patient care, discussion and debate of controversial topics in hernia management, and dissemination of information to a global audience.

Chan TM, Thoma B, Radecki R, et al. Ten steps for setting up an online journal club. *J Contin Educ Health Prof.* 2015;35(2):148-154. Written by a multidisciplinary group of authors, this article provides 10 helpful hints for setting up an online journal club, using examples from existing journal clubs.

Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Instructional design variations in internet-based learning for health professions education: a systematic review and meta-analysis. *Acad Med.* 2010;85(5):909-922. This systematic review and meta-analysis of 51 studies, including 30 randomized trials, identified several key features of internet-based learning interventions that were associated with improved learning outcomes: interactivity, practice exercises, repetition, and feedback. However, the evidence base upon which to design internet-based learning programs is limited by poor study designs, failure to use conceptual frameworks, and lack of adherence to reporting standards.

Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Internet-based learning in the health professions: a meta-analysis. *JAMA.* 2008;300(10):1181-1196. This meta-analysis of 201 studies reported that internet-based instruction had a positive effect on educational outcomes as compared to no intervention, but that there was similar effectiveness between internet-based instruction and traditional educational methods.

Farnan JM, Synder Sulmasy L, Worster BK, Chaudhry HJ, Rhyne JA, Arora VM; American College of Physicians Ethics, Professionalism and Human Rights Committee; American College of Physicians Council of Associates; Federation of State Medical Boards Special Committee on Ethics and Professionalism. Online medical professionalism: patient and public relationships: policy statement from the American College of Physicians and the Federation of State Medical Boards. *Ann Intern Med.* 2013;158(8):620-627. This position statement provides recommendations on issues relating to professional use of the internet and social media including but not limited to: use of social media for nonclinical purposes, patient confidentiality in online discussions, and use of web-based patient education resources.

Guraya SY. The usage of social networking sites by medical students for educational purposes: a meta-analysis and systematic review. *N Am J Med Sci.* 2016;8(7):268-278. This review of 10 articles published between 2004 and 2014 suggested that although 75% of medical students use social networking sites, only 20% use them for academic or educational purposes. Furthermore, none of these studies evaluated whether use of social media improved academic performance.

Ibrahim AM, Lillemoe KD, Klingensmith ME, Dimick JE. Visual abstracts to disseminate research on social media: a prospective, case-control crossover study. *Ann Surg.* 2017;266(6):e46-e48. This prospective, case-control crossover study compared the impact of tweets alone versus tweets accompanied by a visual abstract on article visits; there was a threefold increase when visual abstracts accompanied the tweets.


McLeod RS, MacRae HM, McKenzie ME, Victor JC, Brasel KJ; Evidence Based Reviews in Surgery Steering Committee. A moderated journal club is more effective than an Internet journal club in teaching critical appraisal skills: results of a multicenter randomized controlled trial. *J Am Coll Surg.*
This multicenter randomized trial reported that surgical residents who participated in a journal club moderated by a faculty member scored higher on a validated test of critical appraisal than those who participated in an online journal club.


Taveira-Gomes T, Ferreira P, Taveira-Gomes I, Severo M, Ferreira MA. What are we looking for in computer-based learning interventions in medical education? A systematic review. *J Med Internet Res.* 2016;18(8):e204. This systematic review assessed recent studies on computer-based learning (CBL) for types of software platforms and interventions and adherence to current recommendations for CBL research.
This page intentionally left blank
# Index

Note: Page numbers followed by t indicate tables; those followed by f indicate figures.

## A
A-A index in vascular trauma, 205
AAASC. See American Association of Ambulatory Surgery Centers
Abatacept, 362
Abbe flap in lip reconstruction, 634
ABC technique, 10
Abdominal aortic aneurysms, 919–928
Abdominal compartment syndromes, 238–240, 239f, 241f. See also Compartment syndromes
Abdominal pain
in appendicitis, 1332
in colorectal disorders, 1267
differential diagnosis in, 1335
in mesenteric ischemia, 929–930, 1250
in peptic ulcer disease, 1123
Abdominal reconstruction, 2011–2012
Abdominal surgery, enhanced recovery
Abdominal trauma, 200–204, 225–234
blunt, 202
algorithm on evaluation of, 203f
emergency exploration in, 211
in children, 244, 1752–1753
colon, 230–232
compartment syndrome in, 238–240, 239f, 241f. See also Compartment syndromes, abdominal
damage control surgery in, 217f
diagnostic peritoneal lavage and aspiration in, 202, 202f, 202t
in pregnancy, 240
emergency exploration in, 210–213, 211f–213f
of gallbladder, 226
genitourinary tract, 233–234
grading of injuries in, 204, 204t, 1753, 1753r
hemorrhage in, 203f, 204, 211–212
of liver, 225–227, 1752, 1753r
of pancreas, 229–230
penetrating, 200–204
algorithm on evaluation of, 202f
emergency exploration in, 211
in pregnancy, 240
of rectum, 230–232
in seat belt injuries, 244, 1753, 1753f
of small intestine, 228–229, 1752–1753, 1753f
of spleen, 227–228, 1753
of stomach, 228–229
vascular injuries in, 232–233, 235–237
Abdominal wall, 1549–1562, 1740–1744
acquired disorders of, 1553–1557
anatomy of, 1549–1550, 1550f–1552f
blood supply of, 1549–1550, 1552f
congenital disorders of, 1553, 1740–1744
tumors of, 2020
desmoid tumors, 1556
embryonic development of, 1553, 1740
gastrostomies of, 1741–1743
erinomas of, 1553–1556. See also Hernia, abdominal
incisions in
description of, 1550–1553, 1552f
erina of, 1554–1555
innervation of, 1550, 1552f
lipomas of, 1556–1557, 1557f
muscular layers of, 1549, 1551f, 1600, 1603f
neurofibromas of, 1556–1557
omphalocoele of, 1740–1741
pannus of, 2020
in patent urachus, 1740
physiology of, 1550
in prune-belly syndrome, 1743
tumors of, 1556–1557
vascular injuries in laparoscopy and, 1821, 1821f
Abdominoperineal resection, 1274
Abductor muscles of hand, 1928–1929
Ablative carbon dioxide laser, 260
ABO blood groups, 115
Abortion, induced, 1804–1805
Abscess
abdominal, 172, 426
anorectal
in Crohn's disease, 1285
description of, 1313–1314
Bartholin's glands, 1791
brain, 1868
breast, 550
subareolar, 554–555, 555f
treatment in, 554–555
cryptoglandular, 1313–1314
epidural, 1868–1869
hand, 1948
liver, 172, 1369–1370, 1370f
anebic, 1369–1370
pyogenic, 422
lung, 706r, 706–708, 707f, 708t
pancreas, 172, 1464f
pelvic, in diverticulitis, 1287f
psosas, in Crohn's disease, 1284
retroperitoneal, 1561
spleenic, 1523t, 1530
Absorption
in gallbladder, 1397
in large intestine, 1263
in small intestine, 1222–1225
Abuse, of children, 244, 1745
Accelerated partial breast irradiation, 583
Accommodation response
of rectum, in defecation, 1264
of stomach, 1110
Acidemia, metabolic, 1707
Acid-base balance, 91–93
Acid substances
Acetylcholine, 35, 52f
Acetylcholinesterase inhibitors, 2032
N-Acetylcysteine in acetaminophen overdose, 1361
Achalasia, 1056–1057, 1057f
Achalasia, open, 1065
Acid secretion, gastric, 1105–1108
Acid-base balance, 91–93
Acid substances
Acetylcholine, 35, 52f
Acetylcholinesterase inhibitors, 2032
N-Acetylcysteine in acetaminophen overdose, 1361
Achalasia, 1056–1057, 1057f
bird’s beak appearance in, 1056, 1057f
botulinum toxin injection in, 1063–1064, 1068
esophageal myotomy in, 1057f
Heller technique, 1063–1065
laparoscopic, 1065
peroral, 472f
outcome assessment in, 1065–1068
pneumatic dilation in, 1066, 1068
esophageal perforation in, 1084, 1084f–1085f
vigorous, 1056, 1060, 1065
Acid secretion, gastric, 1105–1108
analysis of output in, 1116
basal output in, 1108, 1117
food ingestion and, 1107–1108
gastrin as stimulant of, 1106–1107, 1108f, 1109f in gastrinoma, 1137
in Helicobacter pylori infections, 1122
maximal output in, 1108, 1117
parietal cells in, 1106–1107, 1107f
in peptic ulcer disease, 1122, 1122f
physiology of, 1105–1108, 1107f–1108f
somatostatin in regulation of, 1107, 1109f in Zollinger-Ellison syndrome, 1137
Acid substances
ingestion of, 1085
by children, 1721
skin injuries caused by burns, 1960
description of, 520–521
Acid-base balance, 91–93
buffering systems in, 91
in children, 1707
hemorrhagic, 142, 143f
homeostasis, 91
in shock, 152–153
Acidemia, metabolic, 1707
Acidosis
metabolic, 91f, 91–93, 93f
in children, 1707
etiologies of, 92, 93f
lactic acid, 92
in renal failure, 99
in shock, 137
in trauma, 237
bloody vicious cycle and, 207, 215, 216f
respiratory, 93
acute, 91f
in children, 1707
chronic, 91f
etiologies of, 93, 93f
Acinic cells of pancreas, 1435–1436, 1436f, 1442, 1442f
Acrochordons, 528
Acrodermatitis enteropathica, 279
Acromegaly, 1289
Acromegaly, 1289
Acromegal, 1289
Acromegal, 1289
Acrocyanosis, 2075
Acromioclavicular joint injuries, 1883
Actinomycosis, 526, 708
Actinomyces, 708
Adenocarcinoma, 216
Adenocarcinoma in situ, 670–671
Adenocarcinoma in situ, 670–671
Adenoid hypertrophy, 622–623
Adenoidectomy, 622–623
Adenoiditis, 622
Adenolipoma, 553
Adenoma
adrenal gland. See Adrenal gland, adenoma
of breast, 602
Adenoid cystic carcinoma, 650
Adenoid hypertrophy, 622–623
Adenoidec...
Adopitive transfer in cancer
immunotherapy, 344
Ado-trastuzumab emtansine, 343f
Adrenal artery, 1682, 1683f
Adrenal gland, 1681–1701
adenoma of
adrenalectomy in, 1691
differentiated from carcinoma, 1690
hyperaldosteronism in, 1685, 1687
nonfunctioning, 1695, 1695–1696
sex steroid excess in, 1692–1693
adrenalectomy of, 1698–1701. See also
Adrenalectomy
anatomy of, 1682, 1683f
postoperative, 427
Cushing’s syndrome in, 1688–1691
ectopic tissue, 1681
embryonic development of, 1681–1682
historical studies of, 1681
hormones released by, 1682–1685
aldosterone, 38, 1682
catecholamines, 38, 1684–1685
glucocorticoids, 1683–1684, 1685f
mineralocorticoids, 1682
sex steroids, 1684, 1692–1693
in shock response, 135
hyperaldosteronism of, 1685–1688
hyperplasia of
adrenalectomy in, 1691
congenital, 1681, 1693
ambiguous genitalia in, 1746–1747, 1747f
Cushing’s syndrome in, 1689
hyperaldosteronism in, 1685, 1687
incidentalomas of, 1695–1697
diagnosis of, 1695–1696
differential diagnosis in, 1695f, 1696–1697
treatment of, 1696–1697, 1697f
pheochromocytoma of, 1693–1695
physiology of, 1682–1685
Adrenal vein, 1682, 1683f
Adrenalectomy, 1698–1701
in adrenocortical carcinoma, 1698
in breast cancer, 598
complications of, 1701
in Cushing’s syndrome
description of, 1691, 1698
historical aspects of, 1681
in hyperaldosteronism, 1688
in incidentaloma, 1696–1697
laparoscopic, 1698–1700, 1699f
lateral transabdominal approach, 1699–1700
posterior retroperitoneal approach, 1700f
open, 1700f–1701f
anterolateral approach, 1700–1701
lateral approach, 1701
posterior approach, 1701
in pheochromocytoma, 1695, 1698
β-Adrenergic antagonists, 881–882
Adrenergic receptors, 1685, 1687f
Adrenocorticotropin hormone in adrenocortical insufficiency, 1697
in Cushing’s syndrome, 1688f,
1688–1689, 1691
deficiency of, 1697
description of, 35–36
in glucocorticoid production, 36f,
1683–1684
in lung cancer, 682
pituatory tumor secreting, 1688
in shock, 135
stimulation test, in adrenocortical insufficiency, 1698
in stress response and injury, 35–36
Adrenogenital syndrome, 1747f
Advance directives, 2036, 2064
Advanced Trauma Life Support, 183
Advancement flaps
description of, 1978, 1979f
endorectal
in fistula in ano, 1317
in rectovaginal fistula, 1318, 1319f
Adverse events, 400f
risk management in, 413
Surgical Care Improvement Project on, 406–407, 407f
Aerobic glycolysis, 38
Aerodigestive tract, upper
Alginic acid, 1037
Alfentanil, 2031
Aldosterone
-Adrenergic antagonists, 881–882
Albuterol, 95f
Alkalosis
Alkaloids, plant, 340, 341f
Alkaline phosphatase, 1353
Alkaline agents
ingestion of, 1085
in children, 1721
skin injuries from, 520–521, 521f
Alkaline phosphatase, 1353
to antibiotics, 1357
growth, 1459–1460, 1460f
Alkali, 303
Alginic acid, 1037
Allografts, 1758
Algorithmic integrative monitoring, 436
ALK gene, 321t
Alkaline phosphatase, 1353
Alkaloids, plant, 340, 341f
Alkalosis
metabolic
in children, 1707
description of, 92–93, 93f
etiology of, 92–93, 93f
respiratory
in children, 1707
description of, 93, 93f
Alveolar hypoventilation, 55
Alveolar macrophages, 1453
Anaplastic carcinoma of thyroid, 1657, 1657f

Anastomosis
  in aortic coarctation repair, 762–763
  biliary enteric, 1414, 1416f
  in colorectal surgery, 1274, 1275f–1276f
  end-to-end, 1274, 1275f–1276f
  end-to-side, 1274, 1275f
  ileal pouch anal anastomosis in, 1272–1273, 1273f, 1278
  in familial adenomatous polyposis, 1292
  in ulcerative colitis, 1283
  side-to-end, 1274
  side-to-side, 1274, 1275f
  stapled technique, 1274, 1275f–1276f
  in esophageus caustic injury, 1088f
  gastrointestinal, 2168f
  in liver transplantation, 382–383
  in lung transplantation, 391, 391f
  in pancreaticoduodenectomy, 1493, 1495–1497

Roux-en-Y. See Roux-en-Y anastomosis

Anastrozole, 596–597
Anesthesia
  administration of, 2028
  American Society of Anesthesiologists
    physical status classification, 2034, 2034r
  in anorectal procedures, 1278–1279
  cardiovascular complications of, 2040
  in colon surgery, 1278–1279
  complications of, 2039–2040
  depth of, 2033
  epidural, 2039
  fasting before, 2036, 2036r
  in hand and wrist injuries, 1933–1934, 1937f
  history of, 2027–2028
  in inguinal hernia repair, 1620
  intraoperative management of, 2036–2039
  local, 2031
  malignant hyperthermia caused by, 428, 2040
  in minimally invasive surgery, 457
  modern era of, 2027–2028
  monitored anesthesia care, 2039
  monitoring during, 2032r–2033r, 2032–2033
  neurologic complications of, 2040
  perioperative evaluation and preparation for
    airway, 2034
    ASA physical status classification, 2034, 2034r
    cardiovascular disease, 2034–2035, 2035f
    endocrine disease, 2036
    fasting, 2036
    hepatic disease, 2036
    laboratory studies, 2033
    medical history-taking, 2033
    overview of, 2033–2034
    pulmonary disease, 2035
    renal disease, 2036
    risk estimation, 2036
    peripheral nerve blocks, 2039
    pharmacodynamics of, 2028–2029
    pharmacokinetics of, 2028–2029
    pharmacology of, 2028–2029
    postoperative nausea and vomiting, 2030, 2039
    postoperative pain managed with, 2039–2040
    psychiatric complications of, 2040
    recovery from, 2039–2040
    regional, 2039
    respiratory failure after, 2040
    spinal, 2039
    truncal blocks, 2039
  Anesthesiologist, in enhanced recovery multidisciplinary team, 2114

Anesthesiology, in bariatric surgery, 1179
Angiography, 900–903
Angiogenesis
  in tumors, 316
  in breast cancer, 579
  in wound healing, 273, 274r

Angioplasty
  of aorta. See Aorta, aneurysms of
  in Behçet’s disease, 971
  of carotid artery, 917, 918f
  in fibromuscular dysplasia, 917
  cerebral, 1852–1853
  locations of, 1852r
  subarachnoid hemorrhage in, 1852f, 1852r, 1852–1853, 1853f
  of coronary arteries in Kawasaki syndrome, 971
  of iliac artery, with abdominal aortic aneurysm, 924
  of left ventricle, postinfarction, 833–834, 834f
  of splenic artery, 1531
  splenectomy in, 1523r, 1531
  Angina pectoris, 801–802
  in aortic stenosis, 824–825
  in coronary artery disease, 809
  indications for bypass surgery in, 809
  disability assessment in, 803, 803r
  Angioembolization, 224
  Angioplasty
    of splenic artery, 1531
    splenectomy in, 1523r, 1531
    in coronary artery disease, 809
    indications for bypass surgery in, 809
    disability assessment in, 803, 803r
  Angioembolization, 224
  Angiogenesis
    in tumors, 316
    in breast cancer, 579
    in wound healing, 273, 274r
  Angiography, 900–903
  bronchial, in hemoptysis, 719
  in carotid artery occlusive disease, 909–910, 910f
  in colorectal disorders, 1265
  complications of, 417
  computed tomography. See Computed tomography angiography
coronary preoperative, 903
digital subtraction technique, 902–903, 903f
  in carotid artery occlusive disease, 910
  in renal artery stenosis, 937
  in gastrinoma, 1137
  in lower extremity arterial occlusive disease, 951
  magnetic resonance technique, 951.
  See also Magnetic resonance angiography
  in mesenteric ischemia, 931
  nephropathy from contrast agents in, 417
  in neurologic examination, 1831
  in portal hypertension, 1366
  in renal artery fibromuscular dysplasia, 935, 936f
  of spleen, 1540–1541
  in subarachnoid hemorrhage, 1852, 1853f
  in Takayasu’s arteritis, 969f, 970
  in thoracic aortic aneurysm, 859
  of upper extremity, 1933
  Angiopathy, amyloid, 1851
  Angioplasty
    in aortofemoral occlusive disease, 948–949
    balloons used in, 905, 905f
    in carotid artery occlusive disease, 911–912, 914–917
    compared to endarterectomy, 911–912
    embolic protection devices in, 915f, 915–916
    in fibromuscular dysplasia, 973
    technique in, 914–917, 915f
coronary
    in cardiogenic shock, 150
    compared to bypass surgery, 810r, 810–811
    prior to other surgeries, 903
    in lower extremity arterial occlusive disease, 961f, 961–962
    complications of, 965–966
cryooplasty technique, 965
    subintimal technique, 962–963, 965
    in mesenteric artery occlusive disease, 864, 933
    percutaneous transluminal, 948, 961–962
    in renal artery occlusive disease, 939–940
  Angiopoietins, 316
  Angiosarcoma
    of breast, 602
    description of, 535–536, 843, 1570, 1579, 1584, 1913
  Angiosomes, 1980
  Angiotensin
    description of, 1682
    renin and. See Renin-angiotensin system
  Angle classification, 2002, 2002f
  Angle of His, 1099, 1187
Aspergillus infections
description of, 711–714, 713f
invasive, 714

Aspiration
foreign body, in children, 1716–1717
gastric contents, 1036–1037
Aspiration procedure, peritoneal, 202, 202f
Aspirin
in carotid artery occlusive disease, 910, 916
peptic ulcer disease caused by, 1122–1123, 1124t
perforation in, 1134
risk factors for, 1123, 1124t
in persistent pain, 2070t
platelet dysfunction caused by, 110
Association of Program Directors in Surgery, 2164

Asthma
in gastroesophageal reflux, 1035, 1037
in obesity, 1200

Astrocytoma

Asthma
in gastroesophageal reflux, 1035, 1037
in obesity, 1200

Aspiration
infections
Aspergillus
intestinal, 1724–1725, 1725f
esophagus, 1717–1721.
See also biliary, 1737–1739.
Esophagus, atresia of
intestinal, 1724–1725, 1725f
of duodenum, 1720, 1724–1725, 1725f
of jejunum, 1725f
mitral valve, 773
tricuspid valve, 770–773
Atrial fibrillation
description of, 838–839
in mitral insufficiency, 819
in mitral stenosis, 819
peripheral artery embolism in, 952
signs and symptoms in, 802
Atrial natriuretic peptides, 60
Atrial septal defect, 751–755
anatomy of, 752, 753f
in cor triatriatum, 768–769
diagnosis of, 753–754
in Ebstein’s anomaly, 777
embryology of, 751–752
pathophysiology in, 752–753
percutaneous closure in, 755
treatment of, 754
in tricuspid atresia, 771
Atrioventricular canal defects
classification of, 789f, 789–790
complete, 787, 789, 789f
operative repair of, 789
partial, 787
Atrium
arrhythmias of
fibrillation. See Atrial fibrillation
lesion sets, 838–839
postoperative, 421
in cor triatriatum, 768f, 768–769
myxomas of, 842–843, 843f
septal defect of. See Atrial septal defect
Atrophic gastritis, 1138–1139
Atrophy
cystitis caused by, 1761
in central venous catheterization, 416
perforation in, 1134
risk factors for, 1123, 1124t
in persistent pain, 2070t
platelet dysfunction caused by, 110
Association of Program Directors in Surgery, 2164

Asthma
in gastroesophageal reflux, 1035, 1037
in obesity, 1200

Astrocytoma

Asthma
in gastroesophageal reflux, 1035, 1037
in obesity, 1200

Aspiration
infections
Aspergillus
intestinal, 1724–1725, 1725f
esophagus, 1717–1721.
See also biliary, 1737–1739.
Esophagus, atresia of
intestinal, 1724–1725, 1725f
of duodenum, 1720, 1724–1725, 1725f
of jejunum, 1725f
mitral valve, 773
tricuspid valve, 770–773
Atrial fibrillation
description of, 838–839
in mitral insufficiency, 819
in mitral stenosis, 819
peripheral artery embolism in, 952
signs and symptoms in, 802
Atrial natriuretic peptides, 60
Atrial septal defect, 751–755
anatomy of, 752, 753f
in cor triatriatum, 768–769
diagnosis of, 753–754
in Ebstein’s anomaly, 777
embryology of, 751–752
pathophysiology in, 752–753
percutaneous closure in, 755
treatment of, 754
in tricuspid atresia, 771
Atrioventricular canal defects
classification of, 789f, 789–790
complete, 787, 789, 789f
operative repair of, 789
partial, 787
Atrium
arrhythmias of
fibrillation. See Atrial fibrillation
lesion sets, 838–839
postoperative, 421
in cor triatriatum, 768f, 768–769
myxomas of, 842–843, 843f
septal defect of. See Atrial septal defect
Atrophic gastritis, 1138–1139
Atrophy
gastritis in, 1142, 1142f
intestinal, in parenteral nutrition, 74
Attrition bias, 2144
Atypical ductal hyperplasia (ADH), 553
Auditory canal tumors, 645–646
Auerbach’s myenteric plexus
description of, 1102
in Hirschprung’s disease, 1734
Augmentation mammoplasty, 2018, 2021f
Austin Flint murmur, 827
Authoritative leadership, 13
Authorship, 2073t, 2073–2074
Autografts
in aortic valve replacement, 817, 828
in spinal arthrodesis, 1866
Autoimmune disorders
Graves’ disease in, 1634
hemolytic anemia in, 1523t,
1524–1525
pancreatitis in, 1454
of peripheral nerves, 1867
Autologous fat grafting, 2024
Autonomic nervous system
nervous system
in shock response, 134
sympathetic. See Sympathetic nervous system
Autonomy principle, 2061
in informed consent, 2064
in withdrawing or withholding of
life-sustaining therapies, 2065
Autophagy, 40
Autophagy in cancer cells, 315
Autotransplantation
description of, 356
of islet cells in chronic pancreatitis,
1480–1481
AVDO2. See Arteriovenous oxygen content difference
AVMs. See Arteriovenous malformations
Axillary artery
in axillofemoral bypass, 946
palpation of, 899
Axillary lymph nodes
anatomy of, 546, 546f
in breast cancer, 591–593, 600
dissection after chemotherapy,
595–596
early invasive, 583–584
historical aspects of surgery in, 542
natural history in, 563, 563f
sentinel node dissection and, 590–591
staging of, 575, 576f–577f
in unknown primary cancer, 600
Axillary vein
anatomy of, 982
thrombosis of, 995–996
Axillofemoral bypass in aortoiliac
occlusive disease, 946
Axis fractures, 1843
Axitinib, 343t
Axon injury
description of, 281, 1848
diffuse, 1837
Axonotmesis, 281, 1848
Azathioprine
in Crohn’s disease, 1238, 1281
immunosuppressive uses of, 357, 359f,
360, 361t
side effects of, 360, 360t
in ulcerative colitis, 1281
Azithromycin, 167t
Aztreonam, 167t

B
B cells
in chyle, 742r
in small intestine immune function,
1226
in transplant rejection, 358

Bacille Calmette-Guérin in bladder cancer, 1771
Bacillus anthracis, 158
as biologic warfare agent, 178
Back pain, 681
Bacteria
biofilm effects on, 299
colonization of, 289, 424
in cutaneous microflora, 159, 163
in intestinal microflora, 159, 1263
in necrotizing enterocolitis, 1728
translocation of, 172
transcription in, 483
Bacterial contamination, 289
in surgery, 164, 169t, 286
in transfusions, 121, 122r
Bacterial infections, 161–162, 163t
anorectal, 1320
aortic aneurysm caused by, 856
appendicitis caused by
description of, 1331–1338
organisms associated with,
1331–1332
breast, 550
in central venous catheterization, 416
cystitis caused by, 1761
Bacterial infections (Cont.):
empyema caused by, 740–741
epididymo-orchitis caused by, 1761–1762
liver abscess caused by, 172, 1369
lung
abscess of, 706–708
description of, 706–717
peritonitis caused by, 171
pharyngitis caused by, 622
pyelonephritis caused by, 1761
sinusitis caused by, 617
skin and subcutaneous tissues, 517–518
infections of, 524–526
surgical site, 554, 286, 287t–288t
in transplant recipients, 364
vulvovaginitis caused by, 1789t, 1793, 1794f
Balanitis, 1762
Balanoposthitis, 1762
Balloon dilatation, in aortic coarctation, 762
Balloon dilatation, 468–469
Balloon expulsion test, 1266
Balloon pump, intra-aortic in cardiogenic shock, 149
description of, 835
Balloon valvotomy, 757
Balloon valvuloplasty, 819
Balloon-occluded retrograde transvenous obliteration, 1367
Banding of stomach. See Gastric banding
Banting, Frederick, 1437
Barbed sutures, 1260
Barbiturates, 2030
Bardach two-flap palatoplasty, 1986
Bariatric surgery
in adolescents, 1207
anemia risk in, 1159
biliopancreatic diversion, 1174f,
1189f–1190f, 1189–1190, 1207
candidates for, 1173–1174
complications of nonsurgical, 1204–1205
postgastric bypass hypoglycemia, 1204
surgical, 1203–1204, 1206–1207
treatment of, 1206–1207
contraindications for, 1173
cost-effectiveness of, 1207
deep venous thrombosis in, 1178
duodenal switch, 1189f–1190f;
1189–1190
enhanced recovery after surgery in, 2123
follow-up for, 1191–1192
future questions about, 1209
gastric banding in. See Gastric banding
ghrelin levels after, 1110, 1111f
history of, 1167–1169, 1168f, 1174f
indications for, 1169, 1173
insufficient weight loss after, 1206
intra gastric balloons for, 1191
investigational practices in, 1190–1191
laparoscopic adjustable gastric banding. See Adjustable gastric banding,
laparoscopic
laparoscopic Roux-en-Y gastric bypass. See Roux-en-Y gastric bypass,
laparoscopic
longer-term studies of, 1193–1198,
1194f–1196r
mechanisms of, 1169, 1175–1177
National Institutes of Health criteria, 1173
nonsurgical treatment versus, 1193
obstructive sleep apnea in, 1177–1178
plastic surgery after, 1208f–1209f;
1208–1209
pneumoperitoneum in, 1180
postoperative care for, 1191–1192
pregnancy after, 1179
preoperative issues for
esophageal surgery, 1205–1206
present-day status of, 1169, 1173
quality assurance of, 1207–1208
quality of life effects of, 1200
operative, 1205–1207
results of, 1192–1203
Roux-en-Y gastric bypass.
See Roux-en-Y gastric bypass
selection criteria for, 1175f
short-term outcomes of, 1192–1193
sleeve gastrectomy, laparoscopic. See
Sleeve gastrectomy, laparoscopic
special equipment and infrastructure for, 1180
utilization patterns for, 1169f
venous thromboembolism risks, 1179
vitamin deficiencies after, 1191–1192
weight regain after, 1206
weight loss after insufficient, 1206
plastic surgery after, 1208–1209
weight regain after, 1206
Barium studies
in colorectal cancer screening, 1294,
1294r
in colorectal injury, 1031–1322
of eosinophilic esophagitis, 1051
of esophagus, 1018–1019, 1021f
in diverticula, 1059f
in leiomyoma, 1082, 1082f
in sarcoma, 1080, 1081f
in gastric disorders, 1115, 1116f
in hiatal hernia, 1018–1019, 1021f
1047 in Hirschsprung’s disease, 1734
in large intestine disorders, 1265
in ulcerative colitis, 1282
Baroreceptors, 86
in shock response, 134
Barotrauma, 446
Barrett, Norman, 1035
Barrett’s esophagus, 1034–1035,
1044–1045
adenocarcinoma in description of, 1035, 1036f, 1038, 1068
risk for, 1068
surgical treatment of, 1074
antireflux surgery in indications for, 1038
outcome of, 1044–1045, 1045r
reoperation of, 1045
surveillance for, 1035
columnar-lined esophagus in, 1018,
1035, 1068
drug therapy in, 1037
in duodenal reflux, 1033, 1035
endoscopy in, 1018, 1019f
intestinal metaplasia in, 1018, 1035
structurally defective sphincter in, 1033r
surveillance for, 1045
Bartholin’s glands
cyst and abscess of, 1791
description of, 1785
Basal cell carcinoma, 528–529
of anus, 1305
of hand, 1955, 1956f
Mohs microsurgery in, 529
types of, 528–529
Basal cell nevus syndrome, 528
Base deficit
in shock, 152–153
hemorrhagic, 142, 143f
in trauma, 237
Basilar artery injuries, 1841
Basilar skull fractures, 1839
Basilic vein, 982
Basiliximab, 359, 359f
Bassini repair in inguinal hernia, 1599,
1610, 1611f
Battle’s sign, 196, 1835
Baux score, 254
Bax death signal protein, 579
Bayesian statistics, 2148
Bcl-2 proteins, 314
in breast cancer, 579
Bcr-Abl, 491, 491f
Beck’s triad, 150, 189
Beckwith-Wiedemann syndrome, 1747
Beclin gene, 315
Beer’s law, 444–445
Beger procedure in chronic pancreatitis,
1475, 1476f, 1479–1480
Behçet’s disease, 971
Belatacept, 359–360t, 362
Bell stages in necrotizing enterocolitis, 1729
Bell’s palsy, 615
Belsey fundoplication, lower esophageal sphincter myotomy with, 1064
Beneficence principle
description of, 2061
in withdrawing or withholding of
life-sustaining therapies, 2065
Bennett fracture, 1934
Bentall procedure, 863, 865f
Bentorimide test, in chronic pancreatitis,
1463
Benzodiazipines
anesthetic uses of, 2030
ketamine and, 2031
Bereavement, 2072
Bernard-Burrow technique, 634, 636f
Bernard-Soulier syndrome, 107
Berne procedure, 1478, 1479f
Best, Charles, 1438
Betel nut/quid, head and neck tumors
associated with, 630
Bevacizumab, 343
Bevezy, 1154, 1154f
Bias
attrition, 2144
in biomedical literature, 2137–2138
Bicarbonate
in gastrointestinal secretions, 87f
in parenteral nutrition, 94r
therapy with
in hyperkalemia, 95, 95f
in metabolic acidosis, 92
Bicaval anastomosis, 389, 390f
Bifunctional RNAi, 505
Biglycan, 30–31
Bile
composition of, 87t, 1352, 1396–1397, 1403f
formation of, 1352, 1396–1397
functions of, 1352
gallbladder storage and regulation of, 1397
leak in cholecystectomy complications, 1420, 1420f
reflux of, gastritis in, 1158
Bile acids, 1352, 1396
in fat metabolism, 1225
in gastroesophageal reflux, 1034, 1034f
Bile ducts
anatomy of, 1350, 1394–1396, 1396f
anomalies in, 1396
carcinoma of, 1377–1378, 1423–1425, 1424f. See also Cholangiocarcinoma
cholangiography of. See Cholangiography
common
anatomy of, 1394–1396
choledocholithiasis in, 1406–1407, 1413–1414. See also Cholelithiasis
cysts of, 1417, 1418f, 1739–1740
idopathic dilatation of, 1739
laparoscopic exploration of, 1413–1414, 1415f
outlet stenosis of, 1415
hamartoma of, 1376
radionuclide scans of, 1399
strictures of, 1418–1419, 1419f
trace of, 1419–1421
in cholecystectomy, 1415, 1419f, 1419–1421, 1420f
ultrasonography of, 1398
Bile salts, 1352, 1396
Biliary disease, traditional Chinese
Gallbladder
gallbladder in.
See also
cysts of, 1417, 1418f, 1739
classification of, 1417, 1418f, 1739
clinical presentation in, 1739
diagnosis of, 1739
etiology of, 1739
treatment of, 1739–1740
diagnostic studies of, 1398–1401
gallbladder in. See Gallbladder
gallstones in, 1401–1414. See also Gallstones
in breast cancer, 574–575
in chest wall tumors, 722
in soft tissue sarcomas, 1571–1572
in thyroid nodule, solitary, 1643–1644
in hand tumors, 1953
incisional, in soft tissue sarcomas, 1572
liver
in cirrhosis, 1364–1365
in incidental mass, 1373
lung
description of, 685
in lymph node assessment, 686
mediastinal tumors, 729–730
melanoma, 530–531, 532f–533f
in neck mass, 646
open, 332–333
rectal, in Hirschsprung’s disease, 1734
sentinel lymph nodes
in breast cancer, 338–339, 590–591
in melanoma, 531, 532f–533f
in vulvar cancer, 1810
soft tissue sarcomas, 1571–1572
stomach cancer, 1145
thyroid, in solitary nodule, 1643–1644
vagina, 1790
vulva, 1790
Bird’s beak appearance in achalasia, 1056, 1057f
Birth injuries, 1916–1917
maternal, perineal lacerations in, 1805
neonatal, brachial plexus palsy in, 1915
Bishop-Koop anastomosis, 1727, 1728f
Bismuth-Corlette classification of bile duct carcinoma, 1423, 1423f, 1424
Bispectral index, 448, 2033
Bisphosphonates
in bone pain, 2071f
in hypercalcemia, 1680f
in hyperparathyroidism, 1670
Bisulfite sequencing, 498–499
Bite wounds, 519–520
of hand, 1950
in human bites, 1950
risk for infections in, 519–520, 520f
Bivalirudin, 989
Black pigment gallstones, 1403
Bladder
anatomy of, 1760
cancer of, 1770–1771
cystectomy in, 1771
epidemiology of, 307t, 308f–309f, 310f
urinary diversion procedures in, 1771
fetal surgery for anomalies of, 1754
fistula of, with vagina, 1791
infections of
Candida albicans, 425
description of, 1761
pressure measurement of
in abdominal compartment syndrome, 238, 239f, 447
description of, 447
trauma of
extraperitoneal, 234, 1766, 1767f
in inguinal hernia repair, 162f
intraperitoneal, 234, 1766, 1767f
in laparoscopy, 1621, 1821
Blalock-Taussig shunts, 771
Blastocysts, 492
Blastomyces dermatitidis, 364
Blastomycosis, 717
Blatchford score in peptic ulcer disease, 1125, 1126

Blood
banked whole, 115
fresh whole, 115, 207
oxygen-carrying capacity of, 117
transfusion of. See Transfusions
volume of, in children, 1707–1708

Blood gas analysis, 445
Blood groups, ABO, 115

Blood pressure, 434–435
anesthetic monitoring of, 2032
ankle-brachial index of, 899, 900f
in aortic dissection, 881–882
preoperative use of, 2052

diastolic, 434
in hemorrhagic shock, 141, 142f, 144
in hypertension. See Hypertension
in hypotension. See Hypotension

Bone(s)
growth of, 1915
growth plate injuries of, 1916

Bone morphogenic proteins, 281

Bone marrow
micrometastasis in, 337
neutrophil mobilization from, 54
spleenectomy in disorders of, 1528–1529

Bone mineral density, 1667, 1670
Bone morphogenetic proteins, 281
Borst’s elephant trunk technique, 865, 1079

Bowel preparation, preoperative, 1279

Bowel, transit time of, 292

Bowel obstruction, 1158

Bother, 1484

Botulism, 1063–1064

Botulinum toxin, 1063–1064, 1068

Bowel preparation, preoperative, 1279

Bowel, transit time of, 292

Bowel obstruction, 1158

Bowel, transit time of, 292

Bone(s)

Bone morphogenic proteins, 281

Bone marrow
micrometastasis in, 337
neutrophil mobilization from, 54
spleenectomy in disorders of, 1528–1529

Bone mineral density, 1667, 1670
Bone morphogenetic proteins, 281
Borst’s elephant trunk technique, 865, 1079

Bowel preparation, preoperative, 1279

Bowel, transit time of, 292

Bowel obstruction, 1158

Bowel, transit time of, 292

Bone(s)

Bone morphogenic proteins, 281

Bone marrow
micrometastasis in, 337
neutrophil mobilization from, 54
spleenectomy in disorders of, 1528–1529

Bone mineral density, 1667, 1670
Bone morphogenetic proteins, 281
Borst’s elephant trunk technique, 865, 1079

Bowel preparation, preoperative, 1279

Bowel, transit time of, 292

Bowel obstruction, 1158

Brain
death
diagnosis of, 364, 1842
organ donation process in, 364–365

Brain stem
anatomy of, 1828
compression of, 1833, 1833f

Branchial cleft
cyst of, 649, 649f
fistula of, 1711

Braun enterointerostomy in bile reflux
gastritis, 1158

BRCA1/BRCA2, 8, 558–561
in breast cancer, 558–561, 559
in breast-ovarian syndrome, 320, 321t,
323–324
carriers of, 559
mutations in, testing for, 559–560
in ovarian cancer
description of, 558–559
risk management in, 560–561

BRCPRA model on breast cancer risk, 556

Breast
541–602
aberrations of normal development
and involution (ANDI), 551t,
551–552
abscess of, 550
subareolar, 554–555, 555f
treatment in, 554–555
accessory, 543
aesthetic surgery of, 2018
age-related changes in
benign breast disorders and, 551, 551t
description of, 544, 547–548, 548f
anatomy of, 543–547, 544f
benign disorders and diseases of,
551–555
classification of, 551, 552r
nonproliferative, 552–553
pathology in, 552–553
proliferative, 553–554
blood supply of, 545–547, 546f
calcium deposits in, 553
cancer of. See Breast cancer
cysts of, 552, 554, 572f
development of
aberrations of normal development
and involution, 551t, 551–552
description of, 543–544, 547–548
disorders in, 551, 551t
embryonic, 543–544
hormones affecting, 544, 547f, 547–548
duct ectasia of, 553
ductal ectasia of, 552
fibroadenomas of
description of, 551, 551f–552f, 553
treatment of, 554
fibrocystic disease of, 553
fungal infections of, 550
in gynecostasia, 549. See also
Gynecostasia
hyperplasia of, 552, 552f, 552t
cancer risk in, 552, 552t
pathology in, 553
inactive and active, 545
infectious and inflammatory disorders
of, 550
inversion of, 545–547
lobular involution of, 552
lymphatics of, 545–547, 546f
mammary milk line and, 543, 544f
mastitis of. See Mastitis
myotic infections of, 550
nipple discharge from, 599–600
nodularity of, 552
Paget’s disease of, 550, 565
phyllodes tumors of, 600–601
phylloides and involution, 551
physiology of, 545–547
plastic and reconstructive surgery of.
See Breast plastic and
reconstructive surgery
in pregnancy and lactation
description of, 544–545, 545f, 548f,
548–549
epithelial hyperplasia of, 552
radial scars of, 553
sarcoma of, 1584
sclerosing adenos of, 552–554
senescence-related changes in, 548–549
Breast cancer, 555–602
adenoid cystic, 602
advanced local-regional, 585–587
angiogenesis in, 579
anti-HER-2/Neu antibody therapy in,
579, 598–599
apocrine, 602
axillary lymph nodes in. See Axillary
lymph nodes, in breast cancer
biomarkers in, 336, 575, 578–580, 579t
biopsy in, 574–575
excisional, with needle localization,
588, 588f–589f
incisions in, 588, 588f
mammography in, 588, 589f
of sentinel lymph nodes, 590–591
breast conservation surgery in, 591.
See also Breast conservation
surgery
chemoprevention of, 348, 557–558
chemotherapy in, 340, 594–597.
See also Chemotherapy, breast
cancer treated with
circulating tumor cells in, 336
conserving surgery, 543
diagnosis of, 567–575
ductography in, 571, 571f
magnetic resonance imaging in,
572–574
mammography in, 567–571,
568f–570f
physical examination in, 567, 567f
studies used in, 581t
ultrasonography in, 571–572,
572f–574f
ductal carcinoma. See Ductal carcinoma
of breast
early invasive, 582–585
dermatofibrosarcoma in, 596–597
ablative, 598
neoadojuvant, 596–597
epidemiology of, 306–309, 310t,
561–563
treatment of, 311f, 561
international variations in, 561
mortality rates in, 307f, 308f–309f,
311f, 561
survival rates in, 561–562, 562f
ER-positive, locally advanced, 587
estrogen and progesterone receptor
status, 578
fatality rates, 2083
invasive factors in. See Genetic factors,
in breast cancer
growth signs of, 542
hereditary influences on, 562
histopathology of, 563–567
historical development of therapy in,
541–543
immunotherapy in, 491
in situ (stage 0)
histopathology in, 563–565
treatment of, 580–582
inflammatory
description of, 601f, 601–602
noninflammatory cancer versus,
601–602, 602r
invasive
description of, 565–567
risk for, in ductal carcinoma in situ,
564–565, 565f
lobular carcinoma. See Lobular
breast carcinoma of breast
locally advanced, 586f
local-regional recurrence of, 587
lymphatic mapping in, 338f, 339
in male, 600
mammary lymph nodes in, internal, 587
mastectomy for, 542–543. See also
Mastectomy
medullary carcinoma, 565–566
metastasis of, 316–317, 587
ablative endocrine therapy in, 598
to axillary lymph nodes. See Axillary
lymph nodes, in breast cancer
bone marrow micrometastasis in, 337
distant (stage IV), 563, 587
to mammary lymph nodes, 587
natural history in, 562–563, 563f
staging of, 575, 576f–577f
trastuzumab for, 598–599
mucinous carcinoma, 565–566
multicentricity in, 564, 564f
multifocality in, 564
natural history of, 562–563
in axillary lymph node metastasis,
563, 563f
in distant metastases, 563
in primary cancer, 562
nipple discharge in, 599–600
oncoplastic surgery in, 591
platin dye assay in, 334
papillary carcinoma, 566
in pregnancy, 600
prognosis in, 587–588
radiotherapy in, 594. See also
Briefings, preoperative, 402–403, 403r
BRIP1, 321t
Broad ligament of uterus, 1786, 1787
Bromochal arteries
description of, 717
hemoptysis in bleeding from, 719
Bronchial gland tumors, 672
Bronchiectasis, 708–710, 709f
in children, 1716
diagnosis of, 709
etiologies of, 708–709
hemoptyis in, 709–710
management of, 709–710
types of, 709
Bronchioles, 668
Bronchiolitis obliterans, 720, 720f
Bronchogenic cysts, 735, 1083, 1716
Bronchopleural fistula
- drainage in, 737f
- after thoracic surgery, 705
Bronchoscopic complications of, 416–417
description of, 707
in hemoptyis, 718–719
in lung cancer, 683, 685
in tracheal neoplasms, 666
Brooke ileostomy
description of, 1272, 1276, 1277f
functional results in, 1278
Brown pigment gallstones, 1403–1404
Brown-Séquard syndrome, 198, 1845, 1846
Brown pigment gallstones, 1403–1404
Burkitt’s lymphoma, 602
Burch procedure, 1808
Burkitt’s lymphoma, 602
Burn(s), 251–262, 1958–1960
- acute management in, 1958–1959
- burn center referral for, 251
- caustic substances as cause of, 520–521, 1086, 1086r
- chemical, 252, 520–521, 1960
- classification of, 252–253
- compartment syndromes in, 258–259
- complications in, 258–259
- Curling ulcer in, 1123
death caused by, 2084
depth of, 253–254
- disasters, 261–262
electrical, 252
esophageal, 1086, 1086r
- excision and grafting in, 259, 1958–1959
- fluid therapy in, 252, 254–255
- future research on, 262
- hand
description of, 1958–1960
- heterotopic ossification secondary to, 261
- hypermetabolic response in, 257–258
- hyperthermic, 521
- inhalation injuries in, 255–257
- initial evaluation in, 251–252
- nutrition in, 257–258
- prevention of, 261
- prognosis in, 254
- radiation, 262
- referral to specialized burn center in, 251, 252r
- rehabilitation in, 260
- rule of nines on, 252, 253f
- size estimations for, 262
- skin substitutes in, 259–260
- sulfur mustard as cause of, 262
- surgical management in, 1959
- topical agents for wound care in, 257
- total body surface area involved in caloric requirements affected by, 258
- description of, 251, 253f, 254
- transfusions in, 255
- ventilatory support in, 257
- zone of hyperemia in, 253
- Burned out pancreatitis, 1460
- Burst fractures, 1843
- of cervical spine, 1901
- of thoracic and lumbar spine, 1902
- Buschke-Lowenstein tumor, 1305
- Buttock lifts, 2021
- Bypass
- aortobifemoral, 944–945, 945f, 949–950
- aortoiliac, 949–950
- axillofemoral, 946
- femorofemoral, in aortoiliac occlusive disease, 960f
- iliofemoral, 946, 946
- iliocaval, 946
- in coronary artery disease, 947
- in congenital diaphragmatic hernia, 1707
- in cervical spine repair, 1901
- in gastric surgery affecting, 1159
- in familial hypocalciuric hypercalcemia, 1668
- in parathyroid hormone deficiency, 1669–1670
- Calcium channel blockers, 882
- Calcium chloride
- in hyperkalemia, 95
- in hypercalcemia, 95
- Calcium gluconate in hyperkalemia, 95
- Calcification
- of breast, 553
- of pulmonary nodule, solitary, 677–678
- Calcitonin
- description of, 1664–1665
- secretion of, 1629
- serum levels of
description of, 1634
- in medullary thyroid cancer, 1656
- therapeutic uses of
- in bone pain, 2071t
- in hypercalcemia, 1680r
- Calcitonin gene-related peptide, 1435
- Calcium absorption
- description of, 1225, 1665
- gastric surgery affecting, 1159
- extracellular, 1958
- in intracellular signaling pathways, 141
- in parenteral nutrition, 94r
- serum levels of
- in cancer, 99
- of description of, 90, 95
- in hypercalcemia, 89, 90
- See also Hypercalcemia
- in hyperparathyroidism, 1668–1669, 1669r
- hypocalcemia, 72r, 90
- See also Hypocalcemia
- parathyroid gland affecting, 1663–1664, 1665f–1666f
- supplemental, in children and infants, 1707
- urine levels of
- in familial hypocalciuric hypercalcemia, 1668
- in hyperparathyroidism, 1669r
- Calcium channel blockers, 882
- Calcium chloride
- in hyperkalemia, 95
- in hypercalcemia, 95
- Calcium gluconate in hyperkalemia, 95
- Calcoli
gallstones. See Gallstones
urinary in hyperparathyroidism, 1667
- Callus formation, in fracture healing, 281
- Caloric requirements, 61f
- age-related differences in, 1708r
- in burns, 258
- in enteral nutrition
- for children, 1708r
- description of, 69–70
- estimation of
- in burns, 258
- description of, 66–68, 68r
- overfeeding in overestimation of, 68
- in fasting, 62
- hypocaloric feeding, 68
- in injury, 62, 62f
- in parenteral nutrition, 72
- sources of, 61r
- Calorimetry, indirect
- in robotic surgery, 467
- Camera systems
- in endoscopy, 463–465
- in robotic surgery, 467
- Canadian Cardiovascular Society anginal classification, 803, 803t
- Cancer, 305–350. See also Adenocarcinoma; Carcinoma; Tumor(s)
- Cancer antigens
- See also CAHPS.
- CAGE questionnaire, 2050
- CAHPS. See Consumer Assessment of Healthcare Providers and Systems
- Cakal interstitial cells, 1110, 1585
- Calcaneal fractures, 1893–1894
- Calcification
- of breast, 553
- of pulmonary nodule, solitary, 677–678
- Calcineurin inhibitors, 362r
- Calcitonin
- description of, 1664–1665
- secretion of, 1629
- serum levels of
description of, 1634
- in medullary thyroid cancer, 1656
- therapeutic uses of
- in bone pain, 2071t
- in hypercalcemia, 1680r
- Calcitonin gene-related peptide, 1435
- Calcium absorption
- description of, 1225, 1665
- gastric surgery affecting, 1159
- extracellular, 1958
- in intracellular signaling pathways, 141
- in parenteral nutrition, 94r
- serum levels of
- in cancer, 99
- of description of, 90, 95
- in hypercalcemia, 89, 90
- See also Hypercalcemia
- in hyperparathyroidism, 1668–1669, 1669r
- hypocalcemia, 72r, 90
- See also Hypocalcemia
- parathyroid gland affecting, 1663–1664, 1665f–1666f
- supplemental, in children and infants, 1707
- urine levels of
- in familial hypocalciuric hypercalcemia, 1668
- in hyperparathyroidism, 1669r
- Calcium channel blockers, 882
- Calcium chloride
- in hyperkalemia, 95
- in hypercalcemia, 95
- Calcium gluconate in hyperkalemia, 95
- Calcoli
gallstones. See Gallstones
urinary in hyperparathyroidism, 1667
- Callus formation, in fracture healing, 281
- Caloric requirements, 61f
- age-related differences in, 1708r
- in burns, 258
- in enteral nutrition
- for children, 1708r
- description of, 69–70
- estimation of
- in burns, 258
- description of, 66–68, 68r
- overfeeding in overestimation of, 68
- in fasting, 62
- hypocaloric feeding, 68
- in injury, 62, 62f
- in parenteral nutrition, 72
- sources of, 61r
- Calorimetry, indirect
- in robotic surgery, 467
- Camera systems
- in endoscopy, 463–465
- in robotic surgery, 467
- Canadian Cardiovascular Society anginal classification, 803, 803t
- Cancer, 305–350. See also Adenocarcinoma; Carcinoma; Tumor(s)
Cardiac tamponade
diagnosis of, 150–151, 189, 189f
obstructive shock in, 150
treatment of, 150–151, 189–190
emergency department thoracotomy in, 189, 189–190, 190f–191f
pericardiocentesis in, 151, 189, 189f
Cardiogenic shock
description of, 131, 148f, 148–150
differential diagnosis of, 193
hemodynamic response to, 135f, 148
in trauma, 149, 193
Cardiomyopathy
ischemic coronary bypass grafting in, 831–832
in mitral regurgitation, 831–832
peripheral artery embolism in, 954
restrictive, 841
Cardiomyotomy, laparoscopic, 1065, 1066f–1067f
Cardioplegia, 808
Cardiopulmonary bypass, 807–808
in aortic valvotomy for congenital stenosis, 754
in atrial septal defect, 754
in atrioventricular canal defect repair, 754
in atrioventricular canal defect repair, 754–755
in esophageal cancer, 1072–1073
in atrioventricular canal defect repair, 754
in aortic aneurysm, 865
in blunt chest trauma, 1867
in left ventricular dysfunction, 148
in tricuspid valve disease, 148
in ventricular septal defect repair, 788
Cardiopulmonary reserve, 1072–1073
Cardiopulmonary surgery, in elderly, 2055
Cardiovascular disease
in anesthesia, 2034–2035, 2035f
in obesity, 1199
Cardiovascular disorders
acquired, 801–844
arterial, 897–975
congenital, 751–794
in kidney transplantation, 369
physical examination in, 803
in postoperative complications, 406–407, 421
preoperative assessment of
in esophageal cancer, 1072–1073
in noncardiac surgery, 804
in thoracic aortic aneurysm repair, 861
venous, 981–1001
Cardiovascular surgery, in elderly, 2055
Cardiovascular system
anesthetic monitoring of, 2032–2033
preoperative assessment of, in elderly, 2050–2051
Cardioverter-defibrillator, implantable, 837
Carditis, 1033
Career development, 20–21
Carney complex, 1343, 1689
Carnitine, 64, 65f
Caroli’s disease, 1374, 1739
Carotid artery, 907–919
aneurysms of
description of, 917, 918f
in fibromuscular dysplasia, 917
arteritis of, radiation-induced, 972
dissection of
description of, 917, 917f
traumatic, 919, 1841–1842
eelongation and kinking of, 916, 916f
eductectomy of, 911–914. See also
Endarterectomy, carotid
fibromuscular dysplasia of, 916–917, 917f, 973
injuries of
carotid-cavernous fistula, 1841
description of, 220, 919, 1841–1842
operative approach for exposure in, 209f, 210
transposition procedure for repair in, 214, 215f
metastasis involving, 648
occlusive disease of, 907–916
angioplasty and stenting in, 911–912
arrest, 912
complications of, 916
contraindications to, 914, 915f
embolic protection devices in, 915f, 915–916
stents approved for, 915, 916f
technique in, 914–917, 915f
asymptomatic, 911
athroembolism in, 907, 908f
diagnostic evaluation in, 909–910
endarterectomy in, 911–914
epidemiology and etiology of, 907–916
grading of stenosis in, 909, 909f
physical examination in, 899
signs symptoms in, 898, 908
stroke risk in, 911
transient ischemic attacks in, 907–908, 910
injury caused by, 909
palpation of, 899
thrombosis of, 1849
Carotid body tumors, 918f, 918–919
Carotid endarterectomy, 418
Carotid-cavernous fistula, 1841
Carotid-cavernous fistula, 1841–1842
Carpal bones
anatomy of, 1929
with angioplasty balloon, 905, 905f
treatment in, 910
Carpal tunnel
anatomy of, 1928, 1929
median nerve compression in, 1867, 1944
Carpenter classification of mitral regurgitation, 832
Carpometacarpal joint
anatomy of, 1926, 1928f
in degenerative wrist changes, 1945, 1946f
Carpal ligament, transverse, 1928f, 1929
Carpal tunnel
anatomy of, 1928f, 1929
median nerve compression in, 1867, 1944
Catecholamines
cardiogenic shock treated with, 149
description of, 38, 66, 1684–1685, 1686f, 1687f
in fight or flight response, 131
in neuroblastoma, 1749
in shock response, 135
urine levels of
in neuroblastoma, 1749
in pheochromocytoma, 1694
Catepsin B, 1442
Catheterization
arterial, 416
cardiac, 806–807, 807f
in aortic aneurysm, 859–860
in aortic insufficiency, 827
in aortic stenosis, 825
in atrial septal defect, 753–754
in chronic constrictive pericarditis, 841
in total anomalous pulmonary venous connection, 766
central venous complications of, 415–416
oxygen saturation measurements in, 440
pulmonary artery catheterization versus, 440–442, 441f
directed thrombolysis, 990–991
in endovascular therapy access for, 904, 904f
with angioplasty balloon, 905, 905f
catheters in, 905
guidewires for, 904
 introducer sheaths for, 905, 905f
fluid therapy in trauma using, 187, 188f
infections associated with
in burns, 259
description of, 416
in parenteral nutrition, 74
postoperative, 174–175
urinary, 425
pulmonary artery
in burns, 254
description of, 437–438
rupture of pulmonary artery in, 416
in trauma, 237
urinary
in acute urine retention, 1768, 1768f
complications of, 425
for renal monitoring, 447
Cauda equina syndrome, 1865
Caudate lobe of liver, 1347
Caveolin-1, 568
Cavopulmonary anastomosis in tricuspid atresia, 772–773
CCDs. See Charge-coupled devices
CD14, 1355
CD18, 57f
CD20, 362
CD25, 359
CD29, 57
CD31, 57
CD52, 35
CD117, 158
CDH1 in hereditary diffuse gastric cancer, 321, 325
CDK (cyclin-dependent kinases), 486
CDK4, 321, 325
inhibitors of, 321, 486
CDK4
description of, 321, 325
inhibitors of, 486
CDKN2A gene, 321
description of, 321, 325
CDKN2a(p14) gene, 321
transfection of, 499
in injury response of, 50–52
in inflammatory response in, 311–312
in cell cycle of, 486
apoptosis of, 486–487
See also Apoptosis
CD29, 57
cancer
apoptosis alterations in, 314–315
autophagy in, 315
cell cycle of
in chemotherapy, 340
circulating, 336–337
dysregulation in, 311–312
metastasis of, 490, 490f
proliferation and transformation of, 490f
in radiation therapy, 347
culture of, 499f, 499–500
cycle of, 486f, 486–487
in cancer, 311–312
and chemotherapy, 340
in radiation therapy, 347
phases in, 311–312, 313f, 486, 486f
hypoperfusion in shock, 137
injury response of, 50–52
inflammatory response in, 56–58
signaling pathways in, 50–52
signaling pathways of, 50–52, 140–141.
See also Signal transduction pathways
stress response by. See Cellular stress response
transfection of, 499f, 499–500
Cell-free mitochondrial DNA, 30
CellSearch, 336
Cellular injury, 28–34
Cellular necrosis, 40
Cellular stress response
apoptosis, 40
autophagy, 40
necroptosis, 40–41
pyroptosis, 41–42
reactive oxygen species in, 39
unfolded protein response, 39–40
Cellular xenotransplants, 392
Cellulitis
description of, 173, 524–525
of hand, 1947–1948
in lymphedema, 1002–1003
Center for Evidence Based Medicine, 2140, 2141f
Centers for Medicare & Medicaid Services
Surgical Care Improvement Project, 17f
Central cord syndrome, 197, 1845, 1846f
Central nervous system, 1829
brain in. See Brain
genital and developmental anomalies of, 1872–1874
infections of, 1867–1868
in inflammatory response, 34–39, 36f
spinal cord in. See Spinal cord
tumors of, 1854–1861
Central venous catheterization
complications of, 415–416
oxygen saturation measurements in, 440
pulmonary artery catheterization versus, 440–442, 441f
Central venous pressure, for anesthetic monitoring, 2033f
Cephalexin, 165t
Cephalic vein, 982
c-erbB-2, 313
Cervicectomy, 1811–1813
cancer of, 1811–1813
biopsy of, 1790
epidemiology of, 307f, 310f
pelvic exenteration in, 1813, 1814f
radical hysterectomy in, 1799–1800f, 1812–1813f
screening for, 331f, 1789–1790
staging of, 1811f
trachelectomy in, 1813
cervical technique, 1805
condylomata of, 1792
polyphs of, 1794
cesarean delivery
incisions in, 1806, 1806f
indications for, 1806
Cetuximab, 343f
c-fos, 313
CFTR in chronic pancreatitis, 1451, 1454
CGCAHPS. See Clinical and Group Consumer Assessment of Healthcare Providers and Systems
Chagas’ disease, 1308
Chamberlain procedure, 730
Chance fractures, 1843
Chancroid, 1320, 1791, 1792t
Charcot’s triad, 1408
Charge-coupled devices, 454
Charitable surgical delivery platforms, 2087–2088, 2088f
Cheek reconstruction, deltopectoral flap for, 1981f
Chemicals
burns from, 252, 520–521, 1960
carcinogenic
description of, 326, 327t
soft tissue sarcoma associated with, 1568
Chemodectomas, 734, 919
Chemomoelization, 1381
Chemokines, 30, 57
Chemoprevention of cancer
of breast, tamoxifen in, 557–558, 596
description of, 348
Chemoreceptors in shock response, 134
Chemotherapy, 340–342, 491
adjuvant, 340, 492
adrenocortical carcinoma treated with,
1692
agents used in, 340–342, 341t
administration of, 341–342
combination of, 341, 492
resistance to, 345–346, 346t
toxicity of, 341
anal and perianal tumors treated with,
1305
bladder cancer treated with, 1771
breast cancer treated with, 594–597
adjuvant, 594t, 594–597
advanced local-regional, 585–587
distant metastasis, 587
neoadjuvant, 595–597
in pregnancy, 600
tratuzumab, 598–599
cholangiocarcinoma treated with, 1377,
1424
cholecystectomy performed with, 1298,
1301
colorectal cancer treated with, 1069, 1073,
1077–1079, 1079t
salvage esophagectomy after, 1076–1077
everalization injuries in, 521
fallopian tube cancer treated with, 1817
laryngeal cancer treated with, 641
liposarcoma treated with, 724t
liver tumors treated with, 1381
in children, 1751
hepatic artery pump chemoperfusion
of, 1381
lung cancer treated with
early-stage, 697
locoregional advanced, 700t
preoperative, 700t
survival rates in, 700t
malignant fibrous histiosarcoma treated
with, 724t
melanoma treated with, 533
mesothelioma treated with, 744
nasopharyngeal cancer treated with, 645
neoadjuvant, 595–597
pancreatic cancer treated with, 1490–
1491, 1498–1499
paranasal sinus tumors treated with, 644
postoperative, 340
preoperative, 340
primitive neuroectodermal tumors
treated with, 724t
principles of, 340
radiation therapy and, 347
response to, 340
rhabdomyosarcoma treated with
of chest wall, 723, 724t, 725
in children, 1589
sarcoma treated with
bone, 723, 724t
chest wall, 723, 724t, 725
in children, 1589
Ewing’s, 723, 724t
extremity, 1578–1579
isolated regional perfusion
technique in, 1576–1578
with radiation therapy, 1577
metastatic, 1581–1582
response to, 723, 724t
stomach cancer treated with, 1147–1148
thyromma treated with, 731
thymic cancer treated with, 1652
anaplastic, 1657
in lymphoma, 1657
vulvar cancer treated with, 1810
Wilms’ tumor treated with, 1748
wound healing in, 283–284
Chest pain
in cardiac disorders, 801, 809, 1031
differential diagnosis in, 1031–1032
noncardiac causes of
description of, 801–802, 1057
in esophageal motility disorders,
1056–1057
in gastroesophageal reflux disease,
1031–1032
Chest trauma, 199–200, 222–225
blunt
description of, 199, 224
heart injury in, 149, 193, 223
cardiac tamponade in, 189
in children, 244, 1752
in elderly, 243
flail chest in, 186, 187f, 224
heart injuries in
blunt, 149, 193, 223
penetrating, 223, 223f
hemothorax in, 188, 188f, 199, 199f, 222
operative treatment in
approaches and exposure in, 210, 211f
indications for, 222, 222–225
penetrating
description of, 199–200
heart injuries in, 223, 223f
pneumothorax in, 186, 187f, 199, 222
respiratory disorders in, 224
vascular injuries in, 224
of aorta, 199, 200f, 200t, 201f, 222, 223f
hematoma in, 199, 200f
of innominate artery, 222, 222f
Chest tubes
complications of, 417
in empyema, 741
in lung abscess, 708
in pleural effusions, 737f
in pneumothorax, 186, 186f, 705
Chest wall, 720–726
anatomy of, 2012
reconstructive surgery of
flaps in, 726, 726f
prosthetic materials in, 726, 726f
in tumors, 726, 726f
trauma of, 224
tumors of, 720–726
benign, 722
biopsy in, 721–722
in bone, 722–723
clinical approach to, 720–721, 721f
laboratory tests in, 721
malignant, 721, 723–726
radiography in, 721
reconstructive surgery in, 726, 726f
in soft tissue, 723–726, 724t
Chiari malformations, 1873–1874, 1874f
Chief cells
of parathyroid gland, 1664, 1664f
pepsinogen secretion of, 1105, 1108
of stomach, 1105, 1105f–1106f, 1106r
Children and infants, 1705–1754. See also
Neonates; Premature infants
abdominal wall deformities in,
1740–1744
abused, 244, 1745
acid-base balance in, 1707
ambiguous genitalia in, 1746–1747
anorectal malformations in, 1735–1737,
1736f
aortic coarctation in, 761–763
appendicitis in, 1337–1338, 1731–1733
clinical presentation of, 1731–1732
differential diagnosis in, 1337, 1733
imaging studies in, 1732, 1732f
perforation and rupture in, 1337,
1732f, 1732–1733
astrocytoma in, 1861
blood volume in, 1707–1708
bone growth in, 1915
brachial plexus palsies in, 1915
branchial cleft anomalies in, 1711
bronchiectasis in, 1716
bronchogenic cysts in, 1716
burns in, 252
fluid therapy in, 254
skin grafts in, 259
caustic and corrosive substances
ingested by, 1085, 1721
cerebral palsy in, 1915
cervical lymphadenopathy in, 1710
choloideal cysts in, 1739–1740
chondroma in, 722
congenital diaphragmatic hernia in,
1712f, 1712–1714
congenital pulmonary airway
malformation in, 1715, 1715f
craniopharyngioma in, 1858–1859
cryptorchidism in, 1777
Cushing’s syndrome in, 1916
emphysema in, congenital lobar, 1714,
1714f
eosinophilic granulomas in, 722
ependymoma in, 1856
esophageal atresia and
tracheoesophageal fistula in,
1717–1721
esophageal injury from corrosive
substance ingestion in, 1721
fluid and electrolyte balance in, 1707
foreign bodies in, 1716–1717
fractures in
description of, 1916–1917
growth plate injuries in, 1916
ganglioneuroblastoma in, 734
gastroesophageal reflux in
description of, 1721–1722
eosophageal atresia and
tracheoesophageal fistula, 1720
gastrochisis in, 1741–1743
general principles of surgery in, 1707–1710
heart disorders in, congenital, 751–794
hemangiomias in, 626–627
Hirschsprung’s disease in, 1734–1735, 1735f
historical aspects of surgery in, 1705–1706
hydronephrosis in, 1777
hypertrophic pyloric stenosis in, 1722–1723, 1723f
hypospadias in, 1776–1777
hypothyroidism in, 1638
informed consent for, 2063–2064
inguinal hernia in, 1604, 1606f, 1743–1744
clinical manifestations of, 1743f, 1743–1744
embryology of, 1743
incarcerated, 1744
laparoscopic repair of, 1744
intestines in
atresia of, 1724–1725
duplications of, 1733
malrotation of, 1725–1726
obstruction of, 1723–1724
transplant of, 386
intravenous catheter placement in, 2037
intussusception in, 1253, 1731, 1731f
jaundice in, 1737–1740
laparoscopy in, 472
Legg-Calvé-Perthes disease in, 1918
liver transplantation in, 472
in biliary atresia, 1738
in, 1419–1421
retained or recurrent stones in, 1408, 1409f
complications of, 422
bile duct trauma in, 1415, 1419f, 1419–1421, 1420f
bile leak in, 1420, 1420f
in gallstones, 1410–1411. See also Gallstones, cholecystectomy in
laparoscopic, 1412–1413
in acute cholecystitis, 1406
bile duct strictures after, 1418–1419, 1419f
bile duct trauma in, 1415, 1419f, 1419–1421
in chronic cholecystitis, 1405
in elderly, 474
in gallstones, 1405–1406, 1411, 1412f, 1412–1413
in pregnancy, 472–473, 1405
technique in, 1412f, 1412–1413
in natural orifice transluminal endoscopic surgery, 471f
in single-incision laparoscopic surgery, 471
in Trends in Mongolia, 2100–2101, 2101f
cholecystitis
acute, 1405–1406
acalculous, 1405, 1415, 1417
clinical manifestations in, 1405–1406
diagnosis of, 1406, 1406f
in gallstones, 1405–1406
pathogenesis in, 1405
radionuclide scans in, 1399, 1406
management of, 1405
I-21
Index
preoperative, 123–124
thromboelastography in, 124, 124f
Coagulation devices, bipolar, 465, 465f
Coarctation of aorta, 761–764
Cobalamin malabsorption, 1225
Cobblestone appearance
in Crohn’s disease, 1280
in gastroesophageal reflux disease, 1019f
Cocarcinogens, 326
Coccidioides immitis, 1225
Coccidioidomycosis, pulmonary, 715, 717
Coccyygeus muscle, 1784
Cockett perforator veins, 291, 981
Codon, 484
Coercive leadership, 12–13
Cognitive assessment, preoperative, in elderly, 2049–2050, 2050r
Cognitive function
in dying patients, 2069
postoperative changes in, 418, 418r
Cognitive impairment, postoperative
delirium associated with, 2049, 2050r
Cognitive task analysis, 2178
Cohort studies, 306
Cold exposure, frostbite injury in, 523
Cold preservation of organs for transplantation, 365
Colectomy, 1270–1272
anastomosis in
description of, 1274, 1275f–1276f
ileorectal, 1283, 1292
anterior, 1273, 1278
bowel preparation for, 1279
in colorectal cancer
description of, 1270, 1272f, 1296
laparoscopic, 1302
in emergency, 1270
in familial adenomatous polyposis, 1292
functional results in, 1278
in Hartmann’s procedure, 1273–1274
left, 1271
extended, 1271
minimally invasive technique, 1270
in colorectal cancer, 1302–1303
mucus fistula in, 1273–1274, 1277
positioning of patient in, 1279
right, 1270
segmental, in Crohn’s disease, 1285
sigmoid
description of, 1271–1272
in diverticulitis, 1287
subtotal, 1272
total, 1272
transverse, 1270–1271
types of, 1270–1272, 1272f
in ulcerative colitis, 1283
Colic, biliary, 1401, 1404
Colic arteries, 1260
Colipase, 1436
Colitis
zystica profunda, 1306
diversion, 1263
fulminant, 1280, 1282–1283, 1285
indeterminate, 1279, 1286
infected, 1309
ischemic
in aortic aneurysm repair, 922
diagnosis of, 1309
signs and symptoms in, 1309
treatment of, 1309
psuedomembranous, 1309
ulcerative
diagnosis of, 1282
differentiated from Crohn’s disease, 1280
elective surgery in, 1283
emergency surgery in, 1283
epidemiology of, 1279
etiology of, 1279–1280
indications for surgery in, 1282
nonoperative management of,
1281–1282
pathology and differential diagnosis in, 1280
Collagen
description of, 516
in Ehlers-Danlos syndrome, 855, 970
synthesis of
description of, 274–275, 275f
in gastrointestinal tract, 279, 280f, 280r
types of
description of, 274, 516
in inguinal hernia, 1605
in wound healing
in cartilage, 281
corticosteroid therapy affecting,
283–284
description of, 273f–274f, 274–275
in gastrointestinal tract, 279, 280f
in hypertrophic scars and keloids, 292–293
nutrition affecting, 284, 285f
in tendons, 281
Collagen disorders, inguinal hernia in,
1605
Collagenase, 279, 280r
Collagenolysis
description of, 275
in gastrointestinal tract, 279, 280f
Collar button abscess, 1948–1949, 1949f
Collateral circulation
in aortoiliac occlusive disease, 941, 942f
in mesenteric artery occlusion, 929, 929f
Collateral ligaments of knee injuries, 1898
Collett and Edwards classification, 764, 764f
Collins, Jim, 4, 4f
Collis gastroplasty, 1042, 1043f, 1049
in hiatal hernia, 1049
in scleroderma of esophagus, 1050
Colloidal carcinoma, 565–566
Colloid solutions
in burns, 255
description of, 94
Colography, computed tomographic, 1294
Colon, 1259–1324
anatomy of, 1260–1263
landmarks in, 1260
lymphatic, 1261
nerve supply in, 1261
vascular, 1260–1261, 1261f
anesthesia in surgery of, 1278–1279
cancer of. See Colorectal cancer
carcinoid tumors of, 1303
clinical evaluation of disorders, 1264–1270
colitis of. See Colitis
composition of secretions, 87r
diverticular disease of, 1286–1288
embryonic development of, 1259–1260
endoscopy of, 1264–1265. See also
Colonoscopy
as esophageal substitute, 1089
in composite with stomach and jejunum, 1090
in vagal esophagectomy, 1090–1091
fistula with pancreas in chronic pancreatitis, 1467–1468
gastrointestinal stromal tumors of, 1303–1304
inflammatory bowel disease of, 1279–1286
leiomyoma and leiomyosarcoma of, 1303
lipoma of, 1303
lymphoma of, 1303
microflora in, 1263
motility of, 1263–1264
multidisciplinary approach to disorders of, 1279
preoperative preparations for surgery of, 1279
pseudo-obstruction of, 1308–1309
resection of, 1270–1274. See also
Colostomy
sarcoma of, 1584
trauma of, 230–232, 231f, 1321–1322
blunt, 1321
colostomy in, 230–232, 1321r
damage control surgery in, 216
iatrogenic, 1321–1323
penetrating, 1321
suture repair and anastomosis in, 230
volvulus of, 1306–1307
sigmoid, 1306–1307, 1308f
transverse, 1307–1308
Colon surgery, traditional Chinese medicine in, 2127
Colonization, bacterial, 289, 424
Colonography, computed tomographic
in colorectal cancer, 1294
description of, 1265
Colonoscopy, 1264
perforation in, 1322
in screening for colorectal cancer, 1294, 1294r
virtual, in colorectal cancer screening, 1294, 1294r
Colony-forming units, 159
Colony-stimulating factors, granulocyte-
colonies of, 44, 47, 277f
Colorectal cancer, 1288–1303
carcinogenesis model for, 312f
clinical presentation in, 1295–1296
diagnostic algorithm on, 1300f
differential diagnosis in, 1335
drug resistance in, 1246
epidemiology of, 306, 309,
1288–1289
compared to other cancers, 308f–309f, 310r
familial, 1291–1293, 1295r
fetal occult blood testing in, 1266–1267, 1293–1294
field defect in, 1296
follow-up and surveillance in, 1302
in heterozygosity loss, 1290
in replication errors, 1290
Colorectal cancer (Cont.):
hereditary nonpolyposis, 322t, 324–325, 1267, 1292–1293
Amsterdam criteria on, 324, 325t
screening guidelines in, 1295r
in inflammatory bowel disease
risk for, 1282, 1285, 1289
screening for, 1295r
locally advanced, 1301
metastasis of, 1295–1296, 1299, 1301–1302
to liver, 1295, 1299, 1378–1379
hepatectomy in, 1386
to lymph nodes, 1296, 1298, 1301
staging of, 1296, 1297r
natural history and routes of spread in, 1295
pathogenesis of, 1289–1290
in polyps, 1290–1291
preoperative evaluation in, 1296
recurrent, 1298, 1302
resection procedures in
in colon carcinoma, 1296, 1298
description of, 1270, 1271f–1272f, 1296, 1298–1299
minimally invasive, 1302–1303
in rectal carcinoma, 1299–1300
risk factors for, 1288–1289
screening for
advantages and disadvantages in, 1294r
description of, 330, 331r, 1293–1295
guidelines on, 1294–1295, 1295r
second primary tumors in, 1296, 1302
staging of
description of, 1295–1296, 1297t
treatment options based on, 1298–1299
survival rates in, 1296, 1297r
total mesorectal excision in, 338, 1300
treatment of, 1296
in colonic carcinoma, 1298–1299
in rectal carcinoma, 1299–1302
in recurrence, 1302
in tumor markers in, 334–335, 1267
malignant neoplasms in, 311
ultrasonography in, endorectal, 1296, 1298f
Colorectal surgery, enhanced recovery
after surgery in, 2122–2123
Colostomy, 1277–1279, 1278f
in colorectal trauma, 230–232, 232f, 1321r
complications of, 1278
diversion colitis in, 1263
dissection of, 1277, 1278f
Hartmann’s pouch in, 1273–1274, 1277 loop, 1277
mucus fistula created in, 1273–1274, 1277
Colostrum, 545
Colpocelesis, 1808
Colporrhaphy, for pelvic organ prolapse
in, 1807–1808
Colposcopy, 1790
Coma
hepatic, 1361
myxedema, in hypothyroidism, 1639
Commissurotomy, mitral
description of, 819
operative technique in, 821–822
Communication
as core competency, 4r
in culture of safety, 401–402
implementation of, 403–404
risk management in, 413–415
in handoffs of patient care
description of, 401–402
sign outs in, 403
on medical errors, 413–415, 2074
in operating room, 401–404
safety checklist on, 402–403, 404f
in surgical counts, 410–411
in wrong-site surgery, 411
in palliative care, 2067–2068, 2068r
in postoperative debriefings, 403
in preoperative briefings, 402–403, 403r, 405f
in teleconsultation and telenursing, 475
Compartment syndromes
abdominal, 238–240, 239f, 241f
bladder pressure in, 238, 239r, 447
in burns, 258–259
closure technique in, 240, 241f–242f
complications in, 240, 242f
decompression in, 239
grades of, 239r
multiple organ impact of, 239, 240f
postoperative, 424
in burns, 258–259
description of, 2005
in fractures, 1883
of lower extremity
fasciotomy in, 237, 238
in ischemia treatment complications, 956–958, 957f, 957r
in trauma, 237
postoperative, 423–424
thoracic, in burns, 259
of upper extremity, 1939, 1943
Complement system, 48–49
in microbial invasion response, 160
in shock response, 140
Complex regional pain syndrome in upper extremity trauma, 1943
Compliance with therapy, in compression therapy for venous insufficiency, 998
Complications
postoperative, 415–428. See also Postoperative period
in robotic surgery, 415
in minimally invasive surgery, 455
in neurologic examination, 1829
in pancreatic cancer, 1487–1489, 1488f
in pancreatitis acute, 1443f, 1443–1444, 1446, 1447f
in chronic, 1458–1459, 1459f
of parathyroid gland, in preoperative localization studies, 1671, 1672r
in pheochromocytoma, 1694, 1694f
in sinusitis, 617f, 618–619
in small intestine obstruction, 1229f, 1229–1231, 1231f
in soft tissue sarcomas in initial assessment, 1570f
in retroperitoneal, 1582, 1582f
in surveillance for recurrence, 1580–1581
in spine trauma, 1845
of spleen, 1521f, 1540
of stomach, 1115–1116, 1116f
in cancer, 1145
description of, 291, 998–1000, 999f
ulcerations and, 291, 998–1000
Computed tomography
in abdominal trauma, 201, 204, 204f
in adrenal cancer, 1691, 1691f
in aortic aneurysm
abdominal, 902f, 921, 921f
in endoleaks, 927f, 927–928
in postoperative surveillance, 925
thoracic, 859
in aortic dissection, 881f, 884
in aortoiliac occlusive disease, 943f
in appendicitis, 1332–1333
in biliary disorders, 1398–1399, 1399f
in carcinoid artery occlusive disease, 910, 910f
in cerebrovascular disorders, 1833, 1834f
in cervical spine trauma, 197
in chest wall tumors, 721, 723f
in children, 1732, 1732f
in esophageal cancer, 1069–1070, 1073
in gallbladder cancer, 1422, 1422f
in goiter, 1642f
of hand and wrist
description of, 1932
of tumors, 1953
in head trauma, 196, 197f, 1836
in children, 1752
in skull fractures, 1835, 1835f
in inguinal hernia, 1607–1608, 1608f
in large intestine disorders, 1265
in mesenteric ischemia, 931, 931f
in mesothelioma, 1265
in myxedema, in hypothyroidism, 1639
in neck trauma, 1902
in neurologic examination, 1829
in minimally invasive surgery, 455
in mesenteric ischemia, 931, 931f
in pyloric stenosis, 1334
in pyogenic abscess, 1369, 1370f
in lower extremity arterial occlusive disease, 951, 951f
of lungs, 678f
in cancer assessment, 683, 685, 687, 689
screening for, 675
in solitary nodule, 677–678
in mediastinal tumors, 727–728
in mesenteric ischemia, 931, 931f
in minimally invasive surgery, 455
in neurologic examination, 1829
in pancreatic cancer, 1487–1489, 1488f
in pancreatitis acute, 1443f, 1443–1444, 1446, 1447f
in chronic, 1458–1459, 1459f
of parathyroid gland, in preoperative localization studies, 1671, 1672r
in pheochromocytoma, 1694, 1694f
in sinusitis, 617f, 618–619
in small intestine obstruction, 1229f, 1229–1231, 1231f
in soft tissue sarcomas in initial assessment, 1570f
in retroperitoneal, 1582, 1582f
in surveillance for recurrence, 1580–1581
in spine trauma, 1845
of spleen, 1521f, 1540
of stomach, 1115–1116, 1116f
in cancer, 1145

of thyroid, 1634
virtual colonoscopy in, 1265, 1294, 1294
Computed tomography angiography, 900–902, 902
in abdominal aortic aneurysm, 902f, 921f
in aortoiliac occlusive disease, 943f
in carotid artery occlusive disease, 910, 910f
in lower extremity arterial occlusive disease, 951, 951f
in mesenteric ischemia, 931f
Computer-enhanced surgery, 453, 467
Concussion in head trauma, 1836–1837
Conditional knockout, 503
Condyloma acuminata
Conditional knockout
Conn’s syndrome, 1685
Congenital melanocytic nevi, 2000
Congenital disorders
abdominal wall, 1553, 1740–1744
adenomatoid malformation in, cystic, 1715, 1715f
adrenal hyperplasia in, 1681, 1693, 1746, 1747f
ambiguous genitalia in, 1746–1747, 1747f
aortic stenosis in, 755–759
in hypoplastic left-heart syndrome, 773
branchial cleft, 1711
breast, 544
central nervous system, 1872–1874
clotting factor deficiencies in, 106–107
diaphragmatic hernia in, 1712–1714
esophageal cysts in, 1083
hand, 1961
heart, 751–794
hepatic cysts in, 1373
inguinal hernia in, 1604, 1606f, 1743–1744
large intestine, 1263
lobar emphysema, 1714, 1714f
lymphedema in, 1001
Meckel’s diverticulum in, 1246–1248
nevi in, 527
pulmonary airway malformation in, 1715, 1715f
small intestine obstruction in, 1228
talipes equinovarus in, 1918
torticollis in, 1712
vagina, 1745
Congenital hemangiomas, 627
Congestive heart failure, 2051
Conjoined tendon, 1600
Connective tissue disorders, 276–279
in Ehlers-Danlos syndrome, 855, 970
inguinal hernia in, 1606r
in Marfan’s syndrome, 278, 855, 970
in pseudoxanthoma elasticum, 970–971
Conn’s syndrome, 1685
Consent, informed, 2062–2064
autonomy principle in, 2064
for children and adolescents, 2063–2064
elements of, 2063, 2063f
Corticosteroid therapy (Cont.):
in immunosuppression, 359t, 359–360, 361t
adverse effects of, 360, 360r
in inflammatory bowel disease, 1238, 1281
intra-articular, 1905
intralesional, in hypertrophic scars and keloids, 293
postoperative, 427
in Cushing’s syndrome, 1691
in septic shock, 148, 176r, 177
in spinal cord injury, 219, 1847
stress dose in, 427
in thrombocytopenic purpura, idiopathic, 1526.
in traumatic brain injury, 1839–1840
wound healing in, 283–284
Corticotrophin-releasing hormone
in Cushing’s syndrome diagnosis, 1690
description of, 1683
Cortisol
description of, 36, 1682, 1683f, 1684
functions of, 1684
serum levels of, 1682
in adrenocortical insufficiency, 1698
diurnal variation in, 1684, 1684f
in shock response, 135
Cost-utility analysis (CUA), 2098
Coudé catheter, 1768f
Cournaud, Claude, 1347
Counts, surgical, 410–411
Cowden disease
breast cancer risk in, 561
description of, 322t, 325, 1643r
diagnostic criteria on, 325
Cowden’s syndrome, 1291
Cow-worker Observation Reporting System,
16t
COX-2 in stomach cancer, 1141
Cox-Maze IV procedure, 838–839, 839f
Craniofacial clefts
lip and palate. See Cleft lip and palate
Craniofacial disorders
in Le Fort fractures, 628–629
Craniofacial microsomnia, 1990
Craniosphenygmoma, 1858–1859
Crawford classification of, thoracoabdominal aortic aneurysms, 869, 870f
C-reactive protein
in acute pancreatitis, 1445
description of, 34
in liver, 1354
Creatinine
age-related changes in, 2052
clearance ratio to calcium, in hyperparathyroidism, 1669, 1669r
serum levels in kidney transplantation, 373
Cremaster muscle, 1602
CREST syndrome, 1058
Crew resource management training, 2175
Crick, Francis H. C., 479–480, 482r
Cricothyroidopiglottopexy, 642

Cricopharyngeal muscle
anatomy of, 1010f, 1010–1012
motility disorders of, 1052–1053
myotomy of, 1054
Cricopharyngotomty, endoscopic, 1055, 1055f
Cricothyroidotomy, 185f
in children, 244
description of, 185
Crigler-Najjar syndrome, 1354
Criminal nerve of Grassi, 1102
Crigler-Najjar syndrome, 1354

Cryoprecipitate, 120
Cruzan, Nancy, 2065
Cruzdron, 321r
Cyst(s)
aneurysmal bone, 1860
Bartholin’s glands, 1791
biliary, 1417, 1739–1740. See also Biliary tract, cysts of branchial cleft, 649, 649f, 1711
breast, 552, 554, 572f
bronchogenic, 735, 1083, 1716
cutaneous, 527
dermoid, 527, 1304, 1859
enteric, 735, 1083
epidermoid, 527
epidermoid, retrorectal, 1304
esophagus, 735, 1082–1083
foregut duplication, 1716
ganglion, of hand and wrist, 1953f, 1953–1954
liver, 1373–1374. See also Liver, cysts of mediastinal, 728, 735
mesenteric, 1560, 1560f, 1733–1734
mucous, of hand and wrist, 1954
omental, 1558–1559
ovarian. See Ovaries, cysts of pancreas, 1500f, 1500–1502
pericardial, 735
popliteal artery adventitia, 973
in pulmonary airway malformation, 1145–1146
proinflammatory, 42–47, 138–140, 140f
in microbial invasion response, 160
in shock, 136, 138–140
in wound healing, 272–273
Cytology
cervical, 1789
peritoneal, in stomach cancer, 1145–1146
Cytomegalovirus infections, 123, 363
Cytotoxic T-lymphocyte antigen 4, 345

Da Vinci robotic system, 453, 455, 465, 467–468, 2174
Dabigatran, 113
Dabrafenib, 343f
Dacron graft, 5
Dalrymple’s sign in Graves’ disease, 1635
Damage control resuscitation strategy
description of, 117, 119
in hemorrhagic shock, 144
Damage control surgery
description of, 215–217, 216f–217f
in liver trauma, 216, 216f, 226
Damage-associated molecular pattern molecules
in cellular injury, 28
definition of, 28
extracellular matrix molecules as, 30–31
heat-shock proteins as, 30, 31f
heme as, 31–32
high-mobility group protein B1, 28–30
as ligands for pattern recognition receptors, 32–33
mitochondrial, 30
in pyropoiesis, 41
S100 proteins as, 31
in shock, 132, 137, 138f, 138f, 145
types of, 29f
Damus-Kaye-Stansel anastomosis
in hypoplastic left-heart syndrome, 775
in Taussig-Bing syndrome, 784
DaNang lung, 132
Dance’s sign, 1731
Daptomycin, 167
Dasatinib, 343
DCC in colorectal cancer, 1289
f
Death
description of, 106
in aortic dissection, 880
in traumatic brain injury, 1841
Diagnosis of, 364, 1842
organ donation process in, 364–365, 365f
cardiac death, organ donation in, 365–366
leading causes of, 398f
pronouncing of, 2072
Death domains, 51
Death receptor pathway, 315
Death-inducing signaling complex (DISC), 315
f
DeBakey, Michael E., 5, 6f
DeBakey classification of aortic dissection, 876, 877f
Decision-making capacity, in elderly, 2050
Decompressive craniectomy, for traumatic brain injury, 1841
Decubitus ulcers, 291, 423. See also Pressure ulcers
Deep brain stimulation, 1869f, 1869–1871
expanding indications of, 1871
obsessive-compulsive disorder treated with, 1871
Deep inferior epigastric perforator flap, 2010, 2010f
Deep venous thrombosis in bariatric surgery patients, 1178
postoperative, 2118
Defecation
accommodation response in, 1264
evacuation studies of, 1266
incontinence in. See Incontinence, fecal
neuropsychologic testing in, 1266
obstructed, 1268–1269
sampling reflex in, 1264
Defecography, 1266
Defibrillators, implantable, 837
Deficit accumulation model, of frailty, 2047
Degenerative disorders
of hand and wrist joints, 1945–1947
of spine, scoliosis in, 1903–1904
of thoracic aorta, mediastinal, cystic, 855
nonspecific, 855
wallerian, 281
Dehydration in children, 1707
Dehydronephrostectomy, 1684
Delayed breast reconstruction, 2009
Delayed hysterectomy, 1807
Delirium
definition of, 2047
in elderly, 2047, 2048f
postoperative, 2040, 2047, 2049
risk factors for, 2047, 2049
Delivery, 1806
brachial plexus palsy from injuries in, 1915
cesarean, 1806, 1806f
ex utero intrapartum therapy (EXIT) procedure in, 1754, 1754f
perineal lacerations in, 1805
Delorme procedure in rectal prolapse, 1306
Deltoperital flap, for cheek reconstruction, 1981f
Democratic leadership, 13
Dendritic cells, 55–56
Dendrocytes, 516
Dendritic leadership, 13
Dentate line, 1261
11-Deoxycorticosterone, 1682, 1683f
Depolarizing neuromuscular blockers, 2031–2032, 2032r
Depression
preoperative assessment of, in elderly, 2050–2051
risk factors for, 2050–2051
Dermal papillae, 1975
Dermatan sulfate, 1628
Dermatome, 275
Dermatitis
perianal, 1319–1320
of vulva, 1791
Dermatofibroma, 528
Dermatofibrosarcoma protuberans, 535, 535f, 1572, 1588–1589
Dermis
anatomy and histology of, 516, 1975
fibers of, 516
replacement products, 524
Dermoid tumors
of central nervous system, 1859
cutaneous, 527
retrorectal, 1304
Desarda repair, of inguinal hernia, 1611–1612, 1612f
Desflurane, 2030, 2030f, 2035
Desmoid tumors
of abdominal wall, 1556
of chest wall, 722
disorders associated with, 722
in familial adenomatous polyposis, 1292
genetic factors in, 1588
locations of, 1588
Desmoplastic fibroma, 1912–1913
Desmopressin acetate
in cardiopulmonary bypass, 114
in von Willebrand’s disease, 107
Desquamation
in pseudomembranous colitis, 1309
postgastrectomy, 1157
chronic, 1269
ulcers in, 291, 298
type III, in chronic pancreatitis, 1462
type 2, 1462
type 1, 2036
pancreatogenic, 1462
pancreas transplantation in
in glucagonoma, 1482
in von Willebrand’s disease, 107
in cardiopulmonary bypass, 114
locations of, 1588
genetic factors in, 1588
disorders associated with, 722
of chest wall, 722
of abdominal wall, 1556
prenatal ultrasonography in, 1712,
lung-to-head ratio in, 1712
honeymoon period in, 1713
chest radiography in, 1712, 1712f
insulin signaling pathway in, 1177, 1200, 1203f,
2036
microarray technology in, 497, 498f
polymerase chain reaction technique in, 495–496
Southern blot hybridization in, 494,
494f
cloning of, 480, 493–494
double-helical structure of, 480, 481f
double-strand breaks in, 326
and heredity, 480–481
loss of heterozygosity, colorectal cancer in, 1290
microsatellite instability
colorectal cancer in, 1290
description of, 325
mismatch repair, 324–325
in hereditary nonpolyposis colon cancer, 1292
next-generation sequencing, 497–499
radiation damage of, 346–347
recombinant technology, 480, 493–494, 494f
replication of, 479, 481, 482f
transfection with, 499–500, 501f, 504
in transgenic mice, 500–501, 501f
DNA tumor viruses, 328
Do not intubate orders, 2036
Do not resuscitate orders, 2036
Dobutamine
in cardiogenic shock, 149
in sepsis, 147
in stress echocardiography, 804
Docetaxel, in lung cancer, 700r
Docosahexaenoic acid, 48
Dog bites, 519–520, 520f
Domperidone, 1115
Duval procedure in chronic pancreatitis, 1473, 1473f
Dysgerminoma, ovarian, 1746
Dysoxia, 137
Dyspepsia, 1115
Dysphagia, 681
barium studies in, 1021
in esophageal cancer, 1069
functional grades of, 1073, 1073f
preoperative nutritional status in, 1073
in esophageal motility disorders
description of, 1051
long esophageal myotomy in, 1060–1062
esophagoscopy in, 1018
in gastroesophageal reflux disease, 1031
and hiatal hernia, 1047
indications for surgery in, 1038
outcome of surgery in, 1042, 1045
postoperative recurrence of, 1045
in hiatal hernia, 1047
in hypertensive lower esophageal sphincter, 1058
oropharyngeal, 1031
in esophagus, 1079
in Schatzki’s ring, 1050
Dysplasia
colonic
high-grade, 1295, 1298
in ulcerative colitis, 1283
fibromuscular, 972–973
of carotid artery, 916–917, 917f, 973
of renal artery, 935, 936f, 973
endovascular therapy in, 939
fibrous, of chest wall, 722
hip, Pavlik harness for, 1918
lung, 669
stomach, 1143
Dysplastic nevus syndrome, 530
Dyspnea
in cardiac disorders, 801–802
in aortic valve insufficiency, 827
in aortic valve stenosis, 824
in dying patients, 2069
Dysrhythmia, 1872
Dystonia, deep brain stimulation in,
Eaton-Lambert syndrome, 1867
EBM. See Evidence-based medicine
Ebstein’s anomaly, 777–780
one-and-a-half ventricle repair in, 779–780
Starnes procedure in, 779
Eccrine glands, 515
Echinocandins, 714
Echinococcosis, hydatid disease in,
Echocardiography, 804, 806
in aortic stenosis, 757, 825
in aortic valve insufficiency, 827
in atrial septal defect, 753
cardiopulmonary reserve evaluations using, 1072
dobutamine stress technique, 804
in double-outlet right ventricle, 783
in mitral stenosis, 819
speckle-tracking, 806
in thoracic aortic aneurysm, 858–859
in total anomalous pulmonary venous connection, 766
transesophageal, 804, 880
cardiac output in, 444
transhoracic, 804
in tricuspid valve disorders, 830
in truncus arteriosus, 764
Eck fistula, 1368
Ectasia, ductal, 552–553
Ectopic conditions
of adrenal gland, 1681
of endometrial glands, 1663, 1675,
1676f
in adenomyosis, 1795
in endometriosis, 1802
of parathyroid gland, persistent
hyperparathyroidism in, 1678,
1678f
in pregnancy, 1805
differential diagnosis in, 1335
testing for, 1789
of testis, 1744
of thyroid gland, 735, 1626–1627
Eculizumab, 359t, 363
Edema
in burns, 1958
in cerebral, in liver failure, 381
of leg, in chronic venous insufficiency, 983, 983f
lymphedema, 1001–1003. See also
Lymphedema
pulmonary
in mitral stenosis, 818
postexpansion, 738
of vocal cord, 625
Edoxaban, 989
Edrophonium, 1050
Edwin Smith Papryrus, 1967, 1968f
EESC. See Emergency and essential
surgical care
Efficacy, of drugs, 2029
Effort thrombosis, 995
Effusions, pleural
access and drainage of, 736–739, 737f
diagnostic work-up in, 736–739
differential diagnosis in, 738t, 738–739
etiologies of, 736, 737t
exudative, 737, 737t–738t, 738–739
in lung cancer, 688
malignant, 739, 739–740
transudative, 737, 737t–738t, 738–739
treatment algorithm on, 740, 740f
Ehlers-Danlos syndrome
aortic aneurysms in, 855
clinical features in, 278, 278t
description of, 276, 855, 970
inheritance of, 278t, 970
types of, 278t, 855, 970
Eicosanoids, 47–48, 49t
Eicosapentaenoic acid, 48
Eisenmenger’s syndrome, 787
Elastase, 1436, 1437f
Elastography, 1358
Elbow
dislocations of, 1885–1886
fractures of, in children, 1917
ulnar nerve compression at, 1867, 1944
Elderly
abdominal aortic aneurysms in, 2056
aging-related physiologic effects in,
243, 243t
ambulatory surgery centers and, 2160
aortic stenosis in, 2056
appendicitis in, 1338
benzodiazepines in, 2030
cancer surgery in, 2055
cardiovascular surgery in, 2055
coronary artery bypass grafting in, 2055
coronary revascularization in, 2056
counseling of, 2052
delirium in, 2047, 2048f
emergency surgery in, 2055
endovascular aortic surgery in, 2056
esophageal cancer in, 1072
falls in, 2047
frailty in
criteria for, 2046t, 2049
description of, 2045–2047
preoperative assessment of, 2049
surgical outcomes affected by, 2055
functional assessment in, 2048–2049
functional recovery in, 2054
geriatric syndromes that affect,
2045–2047
head trauma in, 243
hemorrhagic shock in, 141–142
malnutrition in, 2051
medical assessment of, 2050–2051
medication review in, 2052
minimally invasive surgery in, 474
mitral valve surgery in, 2056
nutritional assessment in, 2051
palliative care for, 2053
palliative surgery in, 2056
polypharmacy in, 2052
population growth of, 2045, 2055
postoperative complications in, 421
preoperative assessment in
activities of daily living, 2049
alcohol abuse, 2050
cardiovascular system, 2050–2051
checklist for, 2048, 2049t
cognitive assessment, 2049–2050, 2050r
congestive heart failure, 2051
decision-making capacity, 2050
depression, 2050–2051
frailty, 2049
functional assessment, 2048–2049
medical assessment, 2050–2051
medication review, 2052
nutritional assessment, 2051
patient counseling, 2052
psychosocial, 2052
pulmonary system, 2051
purpose of, 2047–2048
substance abuse, 2050
timed up and go test, 2049

Endometrium
ablation procedure in abnormal bleeding, 1797
anatomy of, 1785
biopsy of, 1791
cancer of, 1813–1815, 1814r
screening for, 331t
staging in, 1814r
hyperplasia of, 1795–1796
polyps of, 1795
sarcoma of, 1584
Endoneurium, 1848
Endoscopic endonasal approaches, 620, 622
Endoscopic retrograde cholangiopancreatography in chronic pancreatitis, 1457–1460, 1459f, 1463, 1463r
description of, 225, 422, 1400
in gallstone pancreatitis, 1447
pancreatitis after, 1441, 1447
small intestine perforation in, 1252
Endoscopy of anal canal, 1264
capsule, 1265
in Crohn’s disease, 1237, 1237f
in obscure gastrointestinal hemorrhage, 1251
of colon, 1264–1265. See also Colonoscopy
complications of, 416–417
coronary artery bypass surgery in, 813
energy sources in, 465–467, 466f
of esophagus, 1018
in hiatal hernia, 1018, 1047, 1048f
mucosal resection in, 1073–1074
in gastrointestinal hemorrhage, 1251
in gastrointestinal sarcomas, 1583
in hiatal hernia, 1018, 1047, 1048f
historical aspects of, 454
in inguinal hernia repair complications, 1620
intragastric balloon repair using, 1191
natural orifice transluminal endoscopic surgery. See Natural orifice transluminal endoscopic surgery
in pancreatitis, chronic, 1469–1471
parathyroidectomy in, 1674
percutaneous gastrostomy tube placement using, 1191
in portal hypertension and variceal bleeding, 1367
in sinusitis, 618, 618f
and nasal polyps, 618, 619f
skull base surgery, 620–621
small intestine perforation in, 1252
of stomach, 1115f
in cancer
  gastrectomy in, 1148–1149
ultrasonography in, 1116, 1145
indications for, 1115
ultrasonography in, 1116, 1145
surgical access to, 461–462
thyroidectomy in, 1659f, 1661
transanal microsurgery scope, 470–471, 472f
ultrasonography in
  biliary disorders, 1401
  in esophageal cancer, 1070, 1073
  in gastric disorders, 1116, 1145
  in cancer, 1116, 1145
  in pancreatitis, 1457–1458, 1458f, 1463
  video systems in, 463–465
Endoscopy, 2039, 2113
Endnostomy
in small intestine neoplasms, 1243
Endotension, 928
Endothelins, 59, 59f
Endotension
in portal hypertension and variceal bleeding, 1094
placement using, 1191
See also Stent placement
virtual reality training for, 2173–2174
End-systolic pressure-volume line, 437
End-tidal carbon dioxide monitoring, 2034
Enhanced recovery after surgery in abdominal surgery, 2123
in bariatric surgery, 1179–1180, 2123
in colorectal surgery, 2122–2123
components of, 2115, 2115f
in esophagectomy, 2123
evidence-based medicine and, 2113
in gastrectomy, 2123
goal of, 2113
in hepatopancreaticobiliary surgery, 2123
history of, 2113
intraoperative considerations
fluid management, 2118–2119, 2119f
hypothermia, 2117–2118, 2118r
minimally invasive surgery, 2117
surgical site infections, 2117
venous thromboembolism, 2118
metabolic stress response to surgery, 2116f, 2116–2117
mobilization in, 2122
multidisciplinary team involved in, 2114f
overview of, 2113
pain management, 2119–2121, 2120t
pathways, 2039, 2113
phases of, 2115f
postoperative period
hypothermia, 2117–2118
ileus prevention, 2121–2122
mobilization in, 2122
nasogastric tubes in, 2121
nausea and vomiting, 2121, 2121f
preoperative evaluation and optimization benefits of, 2113–2114
(carbohydrate loading, 2117
exercise, 2115–2116
fasting, 2117
nutrition, 2115
patient education in, 2114–2115
prehabilitation, 2115–2116
setting expectations in, 2114–2115
smoking cessation, 2116
program for, 2124, 2125f
in surgical specialties, 2123–2124
Enhanced recovery pathway, 2039, 2113
Enhancers, 485
Entamoeba histolytica infections
anorectal, 1320
liver abscess in, 1369–1370
Enteral nutrition, 68–72, 2115. See also Nutrition, enteral
Enteric nervous system, 1227, 1829
Enteritis, radiation acute, 1245–1246
chronic, 1245–1246
diagnosis of, 1245–1246, 1246f
outcomes in, 1246
treatment and prevention of, 1246
Enterochromaffin-like cells of stomach carcinoid tumors of, 1151
description of, 1103, 1107
Enteroclysis
description of, 1230–1231
in Meckel’s diverticulum, 1247
in radiation enteritis, 1246
in small intestine neoplasms, 1243
vessel access to, 461–462
Enterococcal infections, 165t–167t
Enteroctotis
in Hirschsprung’s disease, 1734
necrotizing, 1727–1730
Bell stages in, 1729
intestinal manifestations in, 1729
in older infants, 1730
pathogenesis in, 1729
pneumatosis intestinalis in, 1729,
1729f
risk factors for, 1728
short bowel syndrome in, 1730–1731
spontaneous perforation, 1730

treatment of, 1729–1730
Enterocolitis
necrotizing, 1010f
thoracic, 1012, 1012f
vascular, 1013f, 1013–1014, 1014f

Epidermoid tumors
Epidermoid cysts, 527
Epidermodysplasia verruciformis, 526
Epidermis, 513–514, 527, 1975
Epidermal growth factor receptors, 313,
Epidermal growth factor, 277
Epidermal appendages, 515–516

Epidemiologic transition of disease in
population, 1490

Enteroenterostomy, 1182
Enteroendocrine cells of small intestine,

Enzymes, pancreatic, 1436–1437, 1437f,
Enzyme-linked immunosorbent spot
(ELISPOT) analysis, 505
Environmental factors in colorectal cancer

Epithelioid cells, pancreas, 1438
Epithelial cells of stomach
description of, 1102–1103, 1105f, 1105f
surface, 1103, 1105, 1109
restitution of, 1109
Epithelialization in wound healing,
275–276, 276f
Epithelial-mesenchymal transition (EMT),
in tumor invasion, 317–318
Epithelioid cells, 716
Epithelioid sarcoma, 1956
Epstein-Barr virus infections
ophthalmology in, 622
in transplant recipients
description of, 363
posttransplant lymphoproliferative
disorder and, 363–364
tumors associated with
nasopharynx, 645
posttransplant lymphoproliferative
disorder and, 328f
stomach, 1141
Epworth Sleepiness Scale, 1177
Equivocality of foot deformity in cerebral
palsy, 1915
ERAS. See Enhanced recovery after
surgery
Erb’s palsy, 1848
ERCC2 gene, 321f
ERCC3 gene, 321f
ERCC4 gene, 321f
ERCC5 gene, 321f
ERCP. See Endoscopic retrograde
cholangiopancreatography
Erectile dysfunction, 1775–1776
Erection of penis, 1760
in priapism, 1769–1770
prostatectomy affecting, 1775
Erlotinib, 343f, 1490
ERP. See Enhanced recovery pathway
Error-enabled approach, 2178
Errors, medical
in communication issues, 401–402
definition of, 397
disclosure of, 413–415, 2074
high-profile, 400–401
Institute of Medicine report on, 397, 2074
legal issues in, 413–415, 1414–
in wrong-site surgery, 411–412
near misses and, 400t
retained surgical items in, 409–410, 410f
risk management in, 413–415
in surgical specimen labeling, 403, 404f
systems change in, 400

Ertapenem, 167t
Erythropoiesis, 524
Erythema chronicum migrans, 1974
Erythema nodosum, 1281
Erythrocytes. See Red blood cell(s)
Erythromycin, 167t
in delayed gastric emptying, 1115f
in preoperative bowel preparation, 1279
Erythropoiesis, 528
ESCAPE trial, 441
Escharotomy, 259, 1958
Escherichia coli
antibiotics for, 165f–167f
in DNA recombinant technology,
493–494, 494f
Esmolol, 222
Esophageal Doppler, 2118
Esophageal speech after laryngectomy, 643
Esophagectomy
in caustic injury of esophagus, 1086
enhanced recovery after surgery in, 2123
esophageal cancer uses of, 1074
in cervical portion, 640
Ivor Lewis en bloc technique, 1076
salvage procedure, 1076–1077
transhiatal approach, 1074–1075, 1077
minimally invasive, 1074–1075,
1075f
open, 1075
transhiatal approach, 2177
minimally invasive three-field,
1075–1076
minimally invasive two-field,
1075–1076
open three-field, 1076
Ivor-Lewis
cancer of esophagus, 1076
complications of, 421
reconstructive surgery after, 1089–1090
thoracic duct injury and chylothorax in,
742f
vagal sparing, colon interposition in,
1090–1091
Esophageal antireflux surgery in, 1038
endoscopy in, 1018, 1019f
in gastroesophageal reflux, 1033f,
1033–1035
Esophagogastroduodenoscopy, 1115, 1115f
Esophagogastroduodenoscopy, 1115, 1115f
in hiatal hernia, 1018, 1047, 1048f
mucosal resection in, 1073–1074
Esophagus, 1009–1091
achalasia of, 1056–1057, 1057f. See also
Achalasia
anatomy of, 1009–1015, 1010f
abdominal, 1012
lymphatic, 1014f, 1014–1015
muscles in, 1012–1013
nerves in, 1014, 1014f
normal areas of narrowing in, 1010,
1010f
thoracic, 1012, 1012f
vascular, 1013f, 1013–1014, 1014f

Index
Esophagus (Cont.):

- atresia of, 1717–1721
- anatomic varieties of, 1717, 1717f
- clinical presentation in, 1717f/1718
- complications of surgery in, 1720
- etiology and pathologic presentation in, 1717–1718
- initial management of, 1718
- outcome in, 1721
- postoperative management in, 1720
- in preterm infant, 1718–1719
- primary surgical correction of, 1719–1720
- Barrett’s. See Barrett’s esophagus
- benign tumors of, 1082–1083
- burns of, in caustic ingestion, 1085, 1086f
- cancer of, 1068–1079
  - in Barrett’s esophagus, 1035, 1036f, 1038, 1068
  - in cervical portion, 640, 1070, 1072f
    - in manifstations, 1068–1069
    - endoscopic mucosal resection in, 1073–1074
    - epidemiology of, 307f, 308f, 310f, 1068, 1069f
    - esophagectomy in, 1074. See also Esophagectomy, esophageal cancer uses of esophagram of, 469f
    - in gastroesophageal reflux disease, 1035
    - general approach to, 1070
    - locations of, 1070, 1072f
    - management algorithm for, 1079, 1080f
    - palliative treatment in, 1070, 1072f
    - preoperative evaluation in, 1073
    - risk factors for, 1068
    - staging of, 1069–1070, 1071f
      - treatment decisions based on, 1070
      - surgical treatment of, 1074
      - cardiopulmonary reserve evaluations before, 1072–1073
      - factors affecting decision for, 1069, 1072f
    - caustic and corrosive injury of, 1010, 1085–1088
      - in children, 1085, 1721
      - grades of, 1086, 1086f, 1721
      - management algorithm on, 1087f
      - columnar-lined, in Barrett’s esophagus, 1018, 1035, 1068
      - corkscrew deformity of, 1057, 1058f
cyst of congenital, 1083
- description of, 735, 1082–1083
diverticulum of. See Diverticulum, esophagus
- in end-stage motor disorders, 1068
evaluation of, 1018–1031
- functional, 1019–1028
- structural, 1018–1019
- fistula of
  - description of, 1088
  - with trachea, 665, 1717–1721. See also Tracheoesophageal fistula
- foreign body in, 1010
  - in children, 1716–1717
  - in gastroesophageal reflux disease, 1031–1045
- hiatal hernia of, 1045–1049
- leiomyoma of, 1082, 1082f
- Mallory-Weiss tears of, 1085
- manometry of. See Manometry, esophageal motility of barium studies of, 1018–1019, 1021f
disorders of, 1052–1060
  - end-stage, 1068
  - of esophageal body and lower esophageal sphincter, 1055–1060
  - long esophageal motility in, 1060f, 1060–1062, 1061f–1062f
  - nonspecific, 1058
  - pharyngoesophageal, 1052–1055
  - primary, 1055–1058, 1056f
  - in scleroderma, 1050f, 1050–1052
  - secondary, 1056f, 1058
  - manometry in assessment of, 1022, 1055
  - in swallowing mechanism, 1015–1017
  - video- and cineradiography of, 1028, 1029f
  - nutcracker, 1057
  - in partial esophageal resection, 1089, 1089f
  - perforation of, 1083–1085
  - diagnosis of, 1083–1084, 1084f
  - management in, 1083–1085, 1084f–1085f
  - pH monitoring of, 1017, 1028, 1029f
  - composite score in, 1030f
  - with gastric pH monitoring, 1030f
  - in gastroesophageal reflux, 1028–1030, 1030f, 1034, 1034f
  - and respiratory disorders, 1035–1038
  - normal values in, 1029, 1029r
  - physiology of, 1015–1017
  - pseudodiverticulosis of, 1057
  - reconstructive surgery of, 1089–1090
  - in caustic and corrosive injury, 1087f, 1087–1088, 1088f, 1721
  - composite of colon, jejunum, and stomach in, 1090
  - resection of. See Esophagectomy sarcoma of, 1079–1082, 1081f
  - polypoid, 1080
  - Schatzki’s ring of, 1049f, 1049–1050
  - scleroderma of, 1050f, 1050–1052, 1058 short, 1049
  - spasm of, diffuse and segmental, 1056–1057, 1058f
  - sphincter muscles of, 1012
    - in achalasia, 1056
    - in gastroesophageal reflux, 1017, 1032–1033, 1038
    - and antireflux surgery, 1039f
    - and hiatal hernia, 1033
    - permanently defective, 1032
    - Heller myotomy of, 1063–1065, 1064f
    - hypertensive, 1057–1058
    - manometry of, 1022
    - motility disorders of, 1055–1056
    - pressure regulation, 1017
    - in scleroderma, 1050
    - in swallowing mechanism, 1015–1017
- strictures of
  - in caustic ingestion, 1086–1087, 1721
  - stent placement in, 1087, 1087f
  - in gastroesophageal reflux, 1033f, 1035
  - indications for antireflux surgery in, 1038
  - outcome of surgery in, 1045
  - structural abnormalities of, 1018–1019
  - surgical complications of, 421
  - in total esophagectomy, 1089–1090
  - trauma of, 199, 222, 224
  - from caustic and corrosive substances, 1010, 1085–1088, 1721
  - fistula in, 1088
  - from foreign bodies, 1083, 1716–1717
  - perforation in, 1083–1085
  - varices of, 1368. See also Varices, gastroesophageal
- Essential Trauma Care project (EsTC), 2090, 2095
- Estlander flap, 634, 635f
- Esters, 2031
- Estrogen
  - breast cancer risk and, 555
  - antiestrogen therapy in, 597–598
  - in breast development and function, 544, 547–548
  - Estrogen receptor status in breast cancer, 578, 596
  - Ethanol, percutaneous injection in the hepatocellular carcinoma, 1380
- Ether, 2027
- Ethical decision making, 14
- Ethical issues, 2061–2066
  - in advance directives and living wills, 2064–2065
  - authorship, 2073r, 2073–2074
  - in conflicts of interest, 2069
  - definitions of terminology related to, 2061–2062
  - in disclosure of errors, 2074
  - in informed consent, 2062–2064
  - in palliative care, 2067–2068
  - in power of attorney for health care, 2064
  - in research, 2072–2074
  - in withdrawing and withholding life sustaining therapies, 2093–2095
  - Ethmoid sinus tumors, 644
  - Efomamide, 2030
  - Etoposide, 700f
- Eukaryotes
  - gene expression in, 483f
  - transcription in, 483
- European Carotid Surgery Trial, 911
- European Organization for Research and Treatment of Cancer (EORTC) protocol, 581, 1580
- Eustachian tubes, 1984
- Evacuation studies, rectal, 1266
- Evans’ syndrome, 487f
- EVAR-1 trial, 926
- Ewell’s syndrome, 2073
- Everolimus, 343f
- Evidence-based medicine
  - authorship, 2073
  - in advance directives and living wills, 2064–2065
  - in informed consent, 2062–2064
  - in palliative care, 2067–2068
  - in power of attorney for health care, 2064
enhanced recovery after surgery and, 2113
evidence
Center for Evidence Based Medicine
levels of, 2140, 2141r
GRADE system for, 2140–2142,
2148–2149
guidelines for synthesis of, 2142
tools for evaluating, 2140–2142
hierarchies of evidence, 2139f,
2139–2140
implications of, 2149
inconsistent use of, 2148
information searches, 2138–2139
internal consistency, 2148
medical reversal, 2150
meta-analysis, 2139
principles of, 2138
reproducibility crisis in, 2150
studies in, 2139
systematic review, 2139
validity
external, 2146
improvements in, 2149–2151
internal, 2144–2146
Evoked potentials, somatosensory and
brain stem, 448
Ewing’s sarcoma
chemotherapy in, 723, 724tr
description of, 721, 1567
Extrinsic pathway, of apoptosis, 40, 315
Extensor muscles of hand,
Extensively drug-resistant tuberculosis,
FADD-like interleukin-1 protease inhibitory proteins, 315
Fallopian tubes
anatomy of, 1786
benign neoplasms of, 1801–1802
caused by, 1815–1816
recurrernt, 1818–1819
risk factors for, 1815, 1816r
endometriosis of, 1802–1803
laparoscopic occlusion for sterilization,
1802
trauma of, 234
Fallot, Etienne Louis, 784
Fallot, tetralogy of
anatomy in, 784f, 784–785
classic presentation in, 784
operative repair in, 785–786
False positive, 2147
Family-centered care, 11
FANC gene, 321r
FANG™ (furin-knockdown and
GMCSF-augmented), 505
FASA. See Federated Ambulatory Surgery
Association
Fas-associated death domain, 51
Fascial planes of neck, deep, 649–650
Fasciitis, necrotizing
description of, 173, 175f, 288,
524–525
postoperative, 426
Fasciocutaneous flaps
description of, 1979, 1983r
in head and neck reconstructive surgery,
651
Nahai-Mathes classification of, 1982r
for posttraumatic extremity
reconstruction, 2007
Extracellular matrix, 30–31, 1972
Extracellular polymeric substance, 299,
1972, 1973f
Extracorporeal membrane oxygenation
in burns, 257
in congenital diaphragmatic hernia,
1713–1714
Extracorporeal perfusion, 807–808. See also
Cardiopulmonary bypass
Extracorporeal shock wave lithotripsy
description of, 467
in pancreatic duct stones, 1471, 1471f
Extravasation injuries, 521, 522f
Extremities
lower. See Lower extremity
trauma of fractures in, 235–237
vascular injuries in, 205, 206t, 207f,
235–237
trauma of tumors of
in osteosarcoma, 1911
in soft tissue sarcomas, 1574–1582
upper. See Upper extremity
Extrinsic pathway, of apoptosis, 40, 315
Eye
cancer of, 307r
inflammatory bowel disease involving,
cancer of, 307t
FAD folds, 315
FATT factors, clotted. See also specific
factor
deficiency of
congenital, 106–107
replacement in, 118t
description of, 104t–105t, 105t
labouratory tests of, 123
liver disease affecting, 111, 111r
liver synthesis of, 1353
vitamin K–dependent, 111r, 112, 123
FADD-like interleukin-1 protease
Exogenous factors, clotted. See also specific
factor
deficiency of
congenital, 106–107
replacement in, 118t
description of, 104t–105t, 105t
labouratory tests of, 123
liver disease affecting, 111, 111r
liver synthesis of, 1353
vitamin K–dependent, 111r, 112, 123
FADD-like interleukin-1 protease
Familial history, 776
Family history, 776
Familial mediterranean fever, 776
Family-based care, 11
Family-centered care, 11
FANC gene, 321r
FANG™ (furin-knockdown and
GMCSF-augmented), 505
FASA. See Federated Ambulatory Surgery
Association
Fas-associated death domain, 51
Fascial planes of neck, deep, 649–650
Fasciitis, necrotizing
description of, 173, 175f, 288,
524–525
postoperative, 426
Fasciocutaneous flaps
description of, 1979, 1983r
in head and neck reconstructive surgery,
651
Nahai-Mathes classification of, 1982r
for posttraumatic extremity
reconstruction, 2007
Fasciotomy, 237, 238f

Fasting
in elderly, 2053
metabolism in, 60–62, 61f
in extended starvation, 61, 61f
in short-term fast, 60f
preoperative, 2036, 2036f, 2053, 2117

Fat
in chyle, 742f
in enteral nutrition solutions for children, 1708
malabsorption of, in chronic pancreatitis, 1461f, 1461–1462
metabolism of
in fasting, 60f–61f, 61f
in injury, 62f, 63
in small intestine, 1224–1225, 1225f
in parenteral nutrition, 73
in stool, in chronic pancreatitis, 1463

Fatigue
in cardiac disorders, 802
in hyperparathyroidism, 1668

Fatty acids, 48
deficiency in parenteral nutrition, 73
in enteral nutrition, 70
metabolism of
in fasting, 62
in injury, 63f, 63–64
oxidation in, 64, 65f
in small intestine, 1224–1225, 1225f
short-chain, 1263
Fatty liver, 1281
FEC-75; 599
Feces. See Stool
Federaled Ambulatory Surgery Association, 2158
Federation of State Medical Boards, 17f
Felon, 1951, 1951f
Felty’s syndrome, 1523f, 1532

Femoral artery
in aortoiliac occlusive disease
and aortobifemoral bypass, 944–945
and axillofemoral bypass, 946
and femorofemoral bypass, 946f
and iliopopliteal bypass, 946f
and thoracoabdominal bypass, 946
injuries of, 235
occlusive disease of, 898
angioplasty in, 961f, 961–962
bypass grafting in, 966
chronic, 958, 959f
computed tomography angiography in, 901f, 959f
endarterectomy in, 966
stent graft in, 963–966
stent placement in, 963, 965f, 965–966
infrainguinal bypass grafting for, 967–968
surgical treatment of
amputation in, 966
bypass grafting in, 966
clinical results of, 968–969
complications of, 966–967
endarterectomy in, 966
endovascular interventions, 968–969
symptoms in, 958–959
Femur
fractures of, 1882f
in children, 1916
distal, 1890
hemorrhage in, 194–195
intercostal, 889–1890
in neck, 1889
in shaft, 1890
subcostal, 1890
vascular injury in, 207f
osteosarcoma of, 1911
rotational abnormalities of, 1918
slipped capital epiphysis of, 1918

Fentanyl
in children and infants, 1710
description of, 2031
in persistent pain, 2070f
Ferguson hemorrhoidectomy, 1311
α-Fetoprotein, 335
in hepatocellular carcinoma, 1364

Fetus
death of, in trauma, 240, 242
fluid and electrolyte balance in, 1707
interventions for disorders in, 1753–1754
ex utero intrapartum therapy (EXIT) procedure in, 1754, 1754f
monitoring of
in laparoscopy, 473
in trauma, 240
prenatal ultrasonography of
in diaphragmatic hernia, 1712, 1712f
in gastrochisis, 1742, 1742f
in ovarian cysts, 1746
wound healing in, 281–282

Fever
in liver trauma, 226
puerperal, 157, 271
in transfusion reactions, 121, 122r

FH gene, 321r
Fiber, 70
Fiberoptic pressure transducer, intraparenchymal, 1831
Fibrillation
atrial, 802, 838–839. See also Atrial fibrillation
ventricular, 802
Fibrillin
abnormalities in Marfan’s syndrome, 278, 855, 970
description of, 275
Fibrin
dissolution of, 106, 106f, 293
in peritoneal adhesions, 293, 294f
in topical hemostatic agents, 115
Fibrin glue in fistula, 1317
Fibrinogen, 118f
in chyle, 742f
deficiency of, 110, 118f
radioactive uptake in venous thromboembolism, 987
transfusion of, 121

Fibrinolysis
description of, 104f, 106, 106f, 207, 293
pathologic, 110
Fibrinopeptides A and B, 105
Fibroadenomas of breast
Fibrinopeptides, 105
Fibrinogen
in focal ischemia, 959

Fibrinogen
in focal ischemia, 959

Fibrinogen
in focal ischemia, 959

Fibrinopeptides
in focal ischemia, 959

Fibrinoporitin, 70
Fibroelastosis, endocardial, 755, 757–758
Fibroelastomas, 843
Fibrocystic disease of breast, 553
Fibroelastomas, 843
Fibroplasia, 283
Fibrosis
cystic
lung transplantation in, 720
meconium ileus in, 1726
pancreatic enzymatic, 1455
in chronic pancreatitis, 1455f, 1455–1456, 1456f
periductal, 552
pulmonary idiopathic, lung transplantation in, 720
trauma of, 226, 1419
ultrasonography of, 1398, 1398f, 1406, 1406f
Gallium nitrate, 1680t
Gallstones, 1401–1414
Gastroesophageal reflux (Cont.):
- physiologic, 1017
- preoperative evaluation in, 1038–1039
- principles of surgical therapy in, 1039–1040
- radiographic detection of, 1030
- respiratory disorders in, 1035–1038
- antireflux surgery in, 1037
- in aspiration of gastric contents, 1036–1037
- pH monitoring in, 1037
- reflux mechanisms in, 1036–1037
- Schatzki’s ring and, 1050
- scleroderma of esophagus and, 1050
- selection of procedure in, 1040
- symptoms in, 1031
- in hiatal hernia, 1047
- preoperative evaluation of, 1038–1039

Gastrohepatic ligament, 1100, 1346, 1347f, 1557

Gastrointestinal stromal tumors, 1585–1588
- colorectal, 1303–1304
- genetic factors in, 1585
- localized, 1585
- locally advanced, 1585–1586
- metastatic, 1585–1586
- multidisciplinary approach to, 1586, 1587f
- prognostic factors in, 1585
- small intestine, 1242, 1242r, 1242–1243, 1243f
- outcome in, 1245
- treatment in, 1245
- stomach, 1139, 1149–1151, 1151f

Gastrointestinal tract
- bacteria in, 159, 1263
- in necrotizing enterocolitis, 1728
- translocation of, 172
- diverticulum of. See Diverticulum
- embryonic development of, 1259–1260
- midgut in, 1259, 1263, 1725–1726
- small intestine in, 1221, 1222f
- endogenous microflora of, 159
- gastrointestinal disorders in, 1620
- hemorrhage from. See Hemorrhage, gastrointestinal
- Hirschsprung’s disease of, 1734–1735, 1735f
- in hyperparathyroidism, 1668
- intestinal duplications in, 1733
- intestinal perforation of
- in necrotizing enterocolitis, 1730
- spontaneous, 1730
- intestinal transplantation procedure
description of, 385–388, 386f–388f
- living donor in, 366
- in short bowel syndrome, 1255
- intussusception of. See Intussusception
- large intestine in, 1259–1324. See also Large intestine
- malrotation of, 1725–1726
- multivisceral transplantation, 385–388
- necrotizing enterocolitis of, 1727–1730
- obstruction of, 1228–1233
- double bubble sign in, 1724, 1724f
- in inguinal hernia, 1609
- in Meckel’s diverticulum, 1247f
- in neonates, 1723–1724
- postoperative, 421
- in inguinal hernia repair, 1620
- pediatric disorders of, 1722–1737
- peritoneal adhesions affecting, 293
- postoperative disorders of, 421–422
- ileus, 1233–1235
- in inguinal hernia repair, 1620
- small intestine obstruction, 1231
- sarcomas of, 1583–1584
- short bowel syndrome of, 1730–1731
- small intestine in, 1219–1225. See also Small intestine
- stomach in, 1099–1160. See also Stomach
- stromal tumors of, 1583–1584. See also Gastrointestinal stromal tumors
- trauma of, 228, 230
- damage control surgery in, 216
- in inguinal hernia repair, 1620
- wound healing in, 279–280, 280f, 280t
- volvulus of. See Volvulus
- wound healing in, 280f, 280t
- in anastomosis, 279–280
- compared to skin healing, 280t

Gastrojejunostomy
- in bile reflux gastritis, 1158
- in Crohn’s disease, 1240
- dumping syndrome in, 1157
- in laparoscopic Roux-en-Y gastric bypass, 1183, 1183f–1184f
- in peptic ulcer disease, 1128, 1128f
- Billroth technique, 1129, 1132f
- Roux-en-Y technique, 1129, 1133f
- Roux syndrome in, 1158–1159
- in stomach cancer, 1146
- volvulus of. See Volvulus

Gastroparesis, 1152
- in diabetes mellitus, 1152
- etiologies of, 1152f
- postoperative, 1157–1158

Gastropathy, hypertrophic, 1153f
- 1153–1154

Gastropathy, Collis
- in hiatal hernia, 1049
- in scleroderma of esophagus, 1050

Gastrochisis, 1553, 1740–1743
- anomalies associated with, 1742
- clinical presentation of, 1741–1742, 1742f
- fluid requirements in, 1707, 1742
- prenatal ultrasonography in, 1742, 1742f
- treatment of
- description of, 1742f, 1742–1743
- silo device in, 1742, 1742f

Gastroesplenic ligament, 1519, 1519f, 1557

Gastroschisis, 1155
- open technique, 72, 1155
- percutaneous endoscopic complications of, 72
- indications for, 72
- Janeway technique, 1155
- with jejunostomy, 72
- placement of, 72
- Stamm technique, 1155, 1156f
- Witzel technique, 1155

Gateway Technology, 493

Gatifloxacin, 287f

Gaucher’s disease, 1523r, 1530

GDNF in Hirschsprung’s disease, 1734

Gefitinib, 343r

Geisinger Health System, 1198–1199

Gelatin, in topical hemostatic agents, 115

Gelofusine in resuscitative fluids, 94t

Gemcitabine
- in pancreatic cancer, 1490, 1498
- in soft tissue sarcoma of extremity, 1578

Gene(s)
- analysis of animal models in, 500–504
- DNA microarray technology in, 497, 498f
- expression and regulation of, 481–485
- in breast cancer, 580
- in cancer initiation, 311
- in metastasis, 317
- in shock, 137, 141
- in human genome, 485–486
- manipulation of, 500–505
- metastasis-suppressor, 317
- mismatch repair, 324–325
- oncogenes, 310–311, 313–314, 320
- tumor suppressor in colorectal cancer, 1289–1290
- description of, 310–311, 490
- in hereditary cancer syndromes, 320–324
- in soft tissue sarcomas, 1570
- Gene array, 497
- Gene chips, 497, 498f
- Gene expression
- inflammatory, translation regulation of, 53–54
- schematic diagram of, 53f
- transcriptional regulation of, 52–53
- Gene knockout in mice, 501–503, 502f
- Gene therapy
- cancer applications of description of, 344f, 345, 491–492
- vector systems in, 345, 492
- in situ, 492
- in wound healing, 298–299
- General anesthesia
- airway management in, 2037, 2037f
- intraoperative management of, 2036–2039
- malignant hyperthermia caused by, 2040

GeneSearch Breast Lymph Node Assay, 585

Genetic code, 484, 484t

Genetic factors
- in aortic aneurysms, thoracic, 855–856, 860
- in biliary atresia, 1737
- in breast cancer
- expression of, 580
- health insurance concerns in, 560
- in hereditary syndromes, 320, 321f, 323–324, 558–561, 559f
- as predictive markers, 334–335
- prophylactic mastectomy in, 556, 558
- risk assessment in, 329
- in cancer, 320–324
- and gene therapy, 344f, 345, 491–492
- Knudson’s two-hit hypothesis on, 320, 324f
- somatic mutations in, 318, 318f
- surgical prevention in, 348–349
- and targeted therapy, 342, 343r
- in colorectal cancer, 1267, 1289–1290
CpG island methylation, 1290
in heterozygosity loss, 1290
in replication errors, 1290
in Crohn’s disease, 1235
in dermatofibrosarcoma protuberans, 1589
in Ehlers-Danlos syndrome, 278t, 970
in gastrointestinal stromal tumors, 1585
in Graves’ disease, 1634–1635
in Hirschsprung’s disease, 1734
in hyperparathyroidism, 1665–1666
in kidney cancer, 1773
in manipulation of, 500–505
in Marfan’s syndrome, 278, 970
in neuroblastoma, 1749
in neuroectodermal tumors, primitive, 725
in ovarian cancer
description of, 320, 321t, 323–324, 558–559
risk management in, 560–561
in pancreatic cancer
description of, 1485
in therapy, 1499
in pancreatitis, 1441, 1450–1451
in pheochromocytoma, 1694–1695
in polycystic liver disease, 1373
in polyposis, familial adenomatous, 324–325, 1267, 1292
in pseudoxanthoma elasticum, 970–971
in rhabdomyosarcoma, 1589
in small intestine neoplasms, 1242–1243
in soft tissue sarcoma, 1569–1570
in stomach cancer, 1141, 1142t, 1143
in thyroid, in solitary nodule, 1643
in thyroid cancer
description of, 1643t, 1644, 1645t, 1646, 1646f
follicular, 1650–1651
medullary, 1655, 1655f
papillary, 1648
in venous thromboembolism, 984–985
in Wilms’ tumor, 1747
Genetic instability, 481
Genetic testing for BRCA mutations, 559–560
Gengraf and Neoral, 362
Genetic testing for BRCA mutations, 559–560
Gentamicin
description of, 166t, 287t
in pelvic inflammatory disease, 1803t
Geriatric patients. See Elderly
Geriatric syndromes
definition of, 2045
delirium, 2047, 2048f
falls, 2047
frailty, 2045–2047, 2046f
Germ cell tumors
mediastinal, 734–735
incidence of, 728t
serum markers in, 728
signs and symptoms in, 729t
nonseminomatous, 734
ovarian, 1746, 1819
testicular, 1771–1772
Germ theory of Pasteur, 157
Germ line mutations, 318
Gestational trophoblastic disease, 1819–1820
Ghon complex, 710–711
Ghost writing, 2073
Ghrelin
in appetite regulation, 1110, 1439
description of, 37–38, 1109–1110, 1437
functions of, 1437t, 1439
serum levels of, 1110
in bariatric surgery, 1110, 1111f
in gastrectomy, 1110, 1110f
GIardia lamblia infections, anorectal, 1320
Giant cell arteritis, 856, 969
Giant cell tumors
of bone, 1913, 1913f, 1957
of tendon sheath in hand, 1954, 1954f
Giardia lamblia infections, anorectal, 1320
Gibbs-Donnan equilibrium equation, 83
Gigantomastia, 551
Gilbert syndrome, 1354
Gigli’s arc, 1932, 1932
Gilbert’s ligament, 1600
Gilles de la Tourette’s syndrome, 2070
Gimmick zone, 732
Gingeval, 2077–2107
in age-related changes in, 2052
in children, 1707
Glossectomy, 636, 639–640
Glossoptosis, 1993
Glottis
anatomy of, 631, 640
tumors of, 641–642
laryngeal preservation techniques in
treatment of, 642
laryngectomy in, 641f
Glucagon, 1436–1437
deficiency in chronic pancreatitis, 1462
functions of, 1437t
secretion of, 1438–1439
serum levels of, in glucagonoma, 1483
Glucagon-like peptide 2 (GLP-2), 1128, 1131t
Glucagonoma, 1482–1483
Glucocorticoid receptor, 36–37
Glucocorticoid response elements, 36
Glucocorticoids, 1683–1684, 1685t
functions of, 1684, 1685t
synthesis of, 36f
Gluconeogenesis
in injury, 62f
in starvation, 61f, 61–62
Glucose, 1223, 1224f
administration of
excessive, 65
in hyperkalemia, 95, 95t
in parenental nutrition, 74
blood levels of
in hyperglycemia. See Hyperglycemia
in hyponatremia, 86
in injury, 65
liver functions affecting, 1351
in pancreas transplantation, 374
in parenental nutrition, 73
postoperative, 427
in septic shock, 147
intolerance to, in parenental nutrition, 74
metabolism of
in fasting, 60
in injury, 65–66, 66f
osmolality of body fluids and, 84
tolerance tests, 1438
transport and signaling of, 66, 1223, 1224f
urine levels of, in parenental nutrition, 73
Glucose transporters (GLUTs), 1223, 1224f
Glucose-dependent insulintropic polypeptide, 1438
Glucose-6-phosphate, 65, 66f
Glucose-6-phosphate dehydrogenase deficiency, splenectomy in, 1522, 1524
Glucuronyl transferase, 1737
Glutamic-oxaloacetic transaminase, 1353
Glutamic-pyruvic transaminase, 1353
Glutamine, 69–70, 1224
Glycerol as fuel, 64
Glycogen in liver, 60, 61f, 61r in muscle, 60, 66
Glycogenolysis, 61
Glycopeptides, 166
Glycosaminoglycans, 274–275
Glycogenolysis, 61
Glycerol as fuel, 64
Glutamic-pyruvic transaminase, 1353
Glutamic-oxaloacetic transaminase, 1353
Glucuronyl transferase, 1737

Goiter
historical aspects of, 1625
intrathoracic, sternotomy in, 1661, 1661f mediastinal, 1661
nontoxic, 1641, 1641r, 1642f toxic diffuse, 1634–1637 multinodular, 1637–1638
Goldenhar’s complex, 1711
Goleman, Daniel, 12
Gonadal dysgenesis, mixed, 1746–1747
Gonadotropin, chorionic in pregnancy test, 1789
Gonadotropin-releasing hormone, 1522f, 1522t
examination of, 1931
Brunicardi_Index_Vol II.indd   40
Brunicardi_Index_Vol II.indd   40

Granuloma
in pelvic organ prolapse, 1807–1808
physical examination in, 1787–1789, 1789f
bimanual, 1788, 1789f
in ulcer syndromes, 1791, 1792r
urinary incontinence in, 1808–1809
Gynecologic surgery, minimally invasive, 1820–1822
Gynecomastia
breast cancer risk in, 600
clinical classification of, 549
description of, 2018
pathophysiology of, 549, 549r
refeeding, 549

H
Haemophilus ducreyi infections, chancre in, 1792t
Hageman factor (factor XII), 104–105, 118r
Hair follicles, 515–516
Hair cell leukemia, 1523r, 1527
Halsted, William, 18, 542, 2163
Halsted radical mastectomy, 542–543, 591
Halsted theory on lymphadenectomy in cancer, 338
Hamartoma
of bile duct, 1376
of breast, 553
Hamburg procedure in chronic pancreatitis, 1476, 1478f
Hand, 1925–1963
anatomy of, 1925–1931, 1928f
bite wounds of, 1950
bones in
anatomy of, 1925–1926, 1928f
fractures of, 1934–1935
burns of
congenital anomalies of, 1961
degenerative joint disorders of, 1945–1947
Dupuytren’s contracture of, 1947
examination of, 1931f–1932f, 1931–1932
fractures in, 1934–1935
Jahss maneuver in, 1934, 1934f
nonunion of, 1943
imaging of, 1932–1933
infections of, 1931, 1947–1952
diagnostic algorithm on, 1952f
joint space, 1949
necrotizing, 1949–1950
osteomyelitis in, 1949
palmar space, 1949
web space, 1948–1949, 1949f
mallet finger in, 1936
motions of
description of, 1926, 1927f, 1931–1932
stiffness in, 1943
muscles in, 1926–1929
nerves in
anatomy of, 1930–1931
compression of, 1943–1945
injuries of, 1936, 1943
neuroma in, 1943
reconstructive surgery of
in fingertip injuries, 1938–1939, 1940f–1942f
tendons in
anatomy of, 1929
examination of, 1931
inflammation of, 1950f, 1950–1951, 1951f
injuries of, 1935–1936
tenodesis maneuver in, 1935
zones in, 1935, 1935f
and pulley system, 1929, 1929f
tenosynovitis of flexor, 1950f, 1950–1951, 1951f
trauma of, 1933–1936
amputation and replantation in, 1938, 1940f–1942f
in bites, 1950
in burns, 1958–1960
of fingertips, 1938–1939, 1939f–1942f
in high pressure injection injuries, 1940f
local anesthesia in, 1933–1934, 1937f
reconstructive surgery in, 1939, 1940f–1942f
splints in, 1934–1935, 1935f
stiffness and limited mobility in, 1943
biopsy in, 1953
diagnosis of, 1953
melanoma, 530, 1955–1956
vascular supply of, 1929–1930, 1930f
tumors in, 1952–1958, 1953

Head and neck disorders, 613–654
See Hospital Consumer
See HATs.
Hasson technique in laparoscopy, 460, 1616
Hashimoto’s thyroiditis, 1634, 1640
Hartley-Dunhill procedure in
Harris-Benedict equation, 258
Hard palate, 1983–1984
Haptics, 2172
Hangman’s fracture, 1843, 1901
Haptics, 2172
Hard palate, 1983–1984
Harris-Benedict equation, 258
Hartley-Dunhill procedure in hyperthyroidism, 1637
Hartmann’s pouch, 1277
Hartmann’s procedure, 232, 1273–1274
Harvey-Bradshaw Index, 1238
Hashimoto’s thyroiditis, 1634, 1640
Hasson technique in laparoscopy, 460, 1616
HATs. See Histone acetyltransferases
HCAHPS. See Hospital Consumer
Assessment of Healthcare
Providers and Systems
Head and neck disorders, 613–654
benign, 613–627
of ears, 613–617, 645–646
in hemorrhage. See Hemorrhage, intracranial
of larynx, 624–626, 640–642
of lips, 630, 633–634
long-term management and rehabilitation in, 654
of nose and paranasal sinuses, 617–622, 643–644
of pharynx, 622–624
plastic and reconstructive surgery in. See Head and neck plastic and reconstructive surgery
tracheotomy in, 652–653
in trauma, 217–219, 627–629, 1833–1842. See also Head trauma
in tumors, 629–651
anatomy and histopathology in, 630f–631f, 630–631
epidemiology of, 630
etiology of, 630
follow-up care in, 654
intracranial, 1854–1859
lymph node metastasis in, 645f–646f, 646–648, 648f
secondary primary tumors in, 631–632
staging of, 632, 632f–634f
of unknown primary site, 643
of upper aerodigestive tract, 632–643
in vascular lesions, 626–627
Head and neck plastic and reconstructive surgery, 651–652
free flaps
description of, 651–652, 652f
in mandibular reconstruction, 651–652
of lips, 634, 635f
local flaps in, 651
regional flaps in, 651
in tumors, 634, 635f–636f
of buccal mucosa, 638
in floor of mouth, 637, 637f
free tissue transfer in, 651–652, 652f
of larynx, 642, 642f
of lip, 634, 635f
of temporal bone, 645–646
of tongue, 634–636, 636f, 639–640
Head trauma, 217–219, 627–629, 1833–1842
brain death in, 1842
in children, 244, 1752
classification of, 1836
closed injury in, 1836–1837
contusion in, 1836–1837
consanguinity in, 1837, 1837f
diffuse axonal injury in, 1837
in elderly, 243
epidural hematoma in
burr hole placement for
decompression in, 217, 219f
description of, 197f, 1837–1838
evaluation of
description of, 196–197, 197f
in initial assessment, 1836
facial fractures in, 219, 219f, 628f–629f, 628–629f
Glasgow Coma Scale in
in children, 1752
description of, 217, 1830f, 1833–1834, 1836
in initial assessment, 1835f, 1836
hemorrhage in
description of, 217, 1839
in scalp injuries, 188, 194, 1834
intracranial, 217–219
intracranial pressure increase in,
217–218, 447–448
lacerations in
description of, 627–628, 628f
of scalp, 188, 194, 1834
maxillofacial, 219, 219f
medical management of, 1839
neurologic monitoring in, 447–449
penetrating, 197, 1837
scalp injury in, hemorrhage in, 188, 194, 1834–1835
seizures in, 1836
skull fractures in, 1835
soft tissue injuries in, 627–628
subdural hematoma in, 1838–1839
vascular injuries in, 1841–1842
wound closure in, 627–628, 628f
Headache, 1852
Healing process, 271–300. See also Wound healing
biofilm and, 299–300
in bones, 280–281
in burns, 259, 1958–1959
in cartilage, 281
in connective tissue diseases, 276–279
delayed, 282, 282f
epithelialization in, 275–276, 276f
excessive, 281–294
factors affecting, 283f, 283–289
fetal, 281–282
in fractures, 280–281
in gastrointestinal tract, 279–280
growth factors in, 276, 298
historical aspects of, 271–272
impaired, 282, 282f
incisional hernia in disorders of, 1554
inflammatory response in
description of, 272–273, 273f–274f
fetal, 282
matrix synthesis in, 273f, 274t, 274–275
maturation and remodeling in, 273f, 275
in nerves, 281
oxygen therapy in, 299
phases of, 272–276, 273f
by primary intention, 282, 282f
proliferation in, 273f, 273–274, 274f
by secondary intention, 276, 282, 282f
tendon, 281
by tertiary intention, 282, 282f
Health insurance concerns in hereditary breast cancer risk, 560
Health Insurance Portability and Accountability Act (HIPAA), 560
Healthcare
costs associated with, 11–12
transparency in, 412
Healthcare Hashtag Project, 2190
Healthcare workers for global surgery, 2095f
Health-related quality of life, 2064
Heart
acquired disorders of, 801–844
in aortic coarctation, 761–764
in aortic stenosis, 755–759
in aortopulmonary window, 769
arrhythmias of, 837–839. See also Arrhythmias
assessment of disorders, 801–807
on disability, 803, 803f
in noncardiac surgery, 804
atrial septal defect of, 751–755
atrioventricular canal defects of, 789–790
cardiomyopathy of. See Cardiomyopathy
cateterization. See Catheterization, cardiac
congenital disorders of, 751–794
in cor triatriatum, 768–769
in coronary artery disease, 808–814
double-outlet right ventricle of, 783–784
in ductus arteriosus patency, 759–761
Heart (Cont.):
in Ebstein’s anomaly, 777–780
hypoplastic left-heart syndrome of, 777–778
in interrupted aortic arch, 790
mechanical circulatory support in disorders of, 835–837
murmurs of. See Murmurs
myocardial disorders of. See Myocardium
neoplasms of, 841–846
metastatic, 843–844
myxomas, 842–843, 843f
pericardial disorders of, 839–841
physical examination in, 803
signs and symptoms in disorders of, 801
in Taussig-Bing syndrome, 784
in tetralogy of Fallot, 784–786
in total anomalous pulmonary venous connections, 765–768
total artificial, 837
transplantation of, 388–391
bicaval anastomosis, 389, 390
complications in, 390
evaluation of, 389
historical aspects of, 357, 388–389
in hypoplastic left-heart syndrome, 776–777
indications of, 389
with lung transplantation, 388–391, 720
organ preservation methods and time in, 365
postoperative care in, 389–390
surgical procedure in, 389
ventricular assist devices as bridge to, 835–837
in transposition of great arteries, 780–783
trauma of, 223
blunt, 149, 193, 223
penetrating, 223, 223f
in truncus arteriosus, 764–765
valvar disorders of, 770–773, 814–817. See also Valvar heart disease
ventricular septal defect of, 786–788
Heart failure, 831–837
in coronary artery disease, 809
epidemiology of, 831
etiology of, 831
mechanical circulatory support in, 835–837
in mitral insufficiency, 821
pathophysiology in, 831
pleural effusions in, 736, 737f–738f
pulmonary artery catheterization in, 441
signs and symptoms in, 802
total artificial heart in, 837
in tricuspid atresia, 771
ventricular assist devices in, 835
in ventricular septal defect, 787
Heart Rhythm Society, 838
Heart sounds
in atrial septal defect, 753
in mitral stenosis, 819
Heartburn
in gastroesophageal reflux disease, 1031, 1031f, 1037
hiatal hernia and, 1047
outcome of surgery in, 1042, 1045
postoperative recurrence of, 1045
Heart-lung transplantation, 391
Heat applications for local hemostasis, 114–115
Heat-shock proteins, 30, 31f, 36f, 315
Heineke-Mikulicz pyloroplasty in Crohn’s disease, 1239f, 1240
in peptic ulcer disease, 1129, 1129f
Heinz bodies, 1520
Helicobacter pylori infections
atrophic gastritis caused by, 1138
gastric cancer in
description of, 326, 114f
risk for, 1139
gastric lymphoma in, 1118, 1149
peptic ulcer disease in, 1115, 1141f
diagnosis of, 1123–1124, 1124f, 1125
pathophysiology in, 1117–1121, 1118f–1119f, 1121f–1122f
treatment of, 1124f, 1125, 1127t
for
f
epithelial antigen test in, 1117
in peptic ulcer disease, 1123
urea breath test in, 1117, 1117f
Heller, Ernst, 1065
Heller myotomy, 1063–1065, 1064f
laparoscopic, 1065, 1066f–1067f
modified, 1065
outcome assessment in, 1068
Hemangioblastoma, 1858
Hemangioma
in children, 626–627, 1996, 1997f
of head and neck, 626–627
infantile, 1996
of liver, 1375, 1375f
natural history of, 1996
Hemangiopericytoma, 1913
Hematocrit, 1268
Hematologic disorders, postoperative, 423–424
Hematoma
aortic intramural, 876f, 879
auricular, 627
in chest trauma, 199, 200f
duodenal, 229
epidural, 197f, 1837–1838
burr hole placement for
decompression in, 217, 219f
in skull fracture, 1835
in inguinal hernia repair, 1621
rectus sheath, 1556
in renal trauma, 1765f
subdural, 196, 1838–1839
acute, 1838
chronic, 1838f, 1838–1839
Hematopoiesis, 1522
Heme, 31–32
Heme oxygenase-1, 1356–1357
Heme oxygenase-2, 1356–1357
Heme oxygenase, 1356–1357
Hematoma
auricular, 627
in chest trauma, 199, 200f
duodenal, 229
epidural, 197f, 1837–1838
burr hole placement for
decompression in, 217, 219f
in skull fracture, 1835
in inguinal hernia repair, 1621
rectus sheath, 1556
in renal trauma, 1765f
subdural, 196, 1838–1839
acute, 1838
chronic, 1838f, 1838–1839
Hematopoiesis, 1522
Heme, 31–32
Heme oxygenase, 1356–1357
Heme oxygenase-1, 1356–1357
Heme oxygenase-2, 1356
Hemihypoplasia
left, 1383–1385
right, 1383
Hemisphrectomy, 1869
Hemophilia, 227
Hemodynamic monitoring, 433–449
of arterial blood pressure, 434–435
of cardiac output and related parameters, 433–434, 436–442
normal range of values in, 439f
pulmonary artery catheterization for, 438
in valvular heart disease, 815f
Hemoglobin, 117
in anemia, 117, 1525
oxygen saturation
arterial, 446
tissue, 444–445
polymerized human, 117
sickle cell, 1525
splenectomy in disorders of, 1525
in thalassemia, 1525
in trauma resuscitation, optimal level
of, 237
Hemoglobin A1c, 2036
Hemoglobin-haptoglobin complexes, 31
Hemolytic anemia
autoimmune
splenectomy in, 1523, 1524–1525
warm-antibody, 1522, 1524–1525
description of, 1354
splenectomy in
description of, 1523, 1524–1525
outcome of, 1538
Hemolytic uremic syndrome, 109
Hemoperitoneum, massive, 189
Hemophilia, 106–107
Hemoptysis
in aspergillosis, 712–713
in bronchiectasis, 709–710
in coccidioidomycosis, 715, 717
in lung abscess, 707–708
massive, 717–719
anatomy in, 715, 717
etiologies of, 717–718, 718t
f
treatment of, 718t, 718–719, 719t
urgent surgery in, 719, 719t
in tuberculosis, 711
Hemorrhage
in abdominal trauma, 203f, 204, 211–212
blood loss source in, 142, 144, 194
blood pressure in, 144
and mortality risk in, 142, 142f
in persistent hypotension, 195
bloody vicious cycle in, 207, 215, 216f
in chest trauma, 222, 224
compensatory mechanisms in, 141
damage control surgery in, 215, 216f
excessive perioperative, 124–125
fibrinogen replacement for, 121
gastrointestinal, 1250–1251, 1267–1268
angiography in, 1265
causes of, 421t, 421–422
in diverticula
acquired, 1250
description of, 1286, 1288
Meckel’s, 1247
in gastric varices, 1153
initial evaluation and management of, 1267–1268
in Mallory-Weiss tear, 1085, 1155
massive, 1152–1153
obscure, 1250–1251, 1252f
occult
in colorectal cancer, 1293–1294
description of, 1266, 1268
ovar, 1250–1251
in peptic ulcer disease, 1124–1125, 1131–1133
Hemorrhoidectomy, 1310–1313
Hemorrhoids, 1310–1313
  - closed submucosal, 1311, 1312
  - Doppler-guided artery ligation, 1312
  - in Crohn's disease, 1285
  - in children, 1751, 1751f
  - external, 1310
  - internal, 1310
  - postpartum, 1310
  - in preoperative evaluation, 123–124
  - in prevention, 208–209, 993, 2053
  - right, 1383, 1384f
  - stapled, 1311–1312
  - therapy, 905, 905f
Hepatic artery
  - anatomy of
    - general, 1348, 1348f, 1433
    - variations in, 1348, 1396, 1433
  - chemoembolization in hepatocellular carcinoma using, 1381
  - thrombosis of, in liver transplantation, 383, 385
Hepatic duct
  - anatomy of, 1350, 1351f, 1394
  - variations in, 1350, 1351f, 1396
Hepatic resection, in liver cancer, 1379–1380
Hepatic veins
  - anatomy of, 1349–1350, 1350f
  - Budd-Chiari syndrome in obstruction of, 1368–1369
  - pressure measurement in, 1366
  - thrombosis of, 1368–1369
Hepaticojunostomy
  - in cholangiocarcinoma, 1424
  - description of, 1414, 1416f
Hepatitis
  - in kidney transplant recipients, 369
  - neonatal, 1738
Hepatitis A
  - description of, 1371
  - liver transplantation in, 381
Hepatitis B
  - description of, 178, 1371
  - liver transplantation in, 381
  - tumors associated with, 328, 328t
  - vaccination, 178
Hepatitis C, 1371–1372
  - chronic liver disease caused by, 1364
  - in kidney transplant recipients, 369
  - liver transplantation for, 380
  - in transplant recipients, 123, 369
  - tumors associated with, 328, 328t
Hepatitis E, 381
Hepatoblastoma, 1751
Hepatocellular carcinoma, 328, 328t, 1376–1377, 1377f, 1379–1381
  - chemoembolization in, 1381
  - in children, 1751, 1751f, 1751t
  - staging of, 1751, 1751t
  - treatment of, 1751
  - in cirrhosis, 1364, 1376
  - ethanol ablation in, 1380
  - hepatectomy in, 1386
  - hepatic resection in, 1379–1380
  - liver transplantation in, 381, 1377, 1380, 1751
Hepaticojunostomy, 1381–1386
  - devices and procedures in, 1382t, 1382–1383
  - historical aspects of, 1346
  - ischemic preconditioning in, 1385
  - laparoscopic, 1386–1387
  - left, 1383–1385, 1384f
  - in liver cancer, from colorectal cancer metastasis, 1386
  - in living donor kidney transplantation, 366, 367f
  - nomenclature on, 1381–1382, 1382f, 1382t
  - portal vein embolization before, 1385–1386
  - Pringle maneuver in, 1385
Hemorrhoid(s), 1310–1313
Hemorrhoidectomy, 1310–1313
  - closed submucosal, 1311, 1312f
  - complications in, 1312–1313
  - stapled, 1311–1312
Hemostasis, 103–125
  - biology of, 103–106, 104f
  - clotting factors in, 104f, 105, 105f.
  - See also Factors, clotting
  - disorders of, 106–115
  - acquired, 108–115
  - in congenital factor deficiencies, 106–107
  - fibrinolysis in, 104f, 106, 106f
  - laboratory tests of, 105, 123–124
  - local
    - in liver trauma, 226
    - mechanical procedures in, 114
  - thermal agents in, 114–115
  - topical agents in, 115
  - platelet function in
    - congenital disorders of, 107–108
    - description of, 103–104, 104f–105f
    - qualitative disorders of, 108r, 110
    - quantitative disorders of, 108r, 108–109
  - preoperative evaluation of, 123–124
  - primary, 103
  - thromboelastography in, 124, 124f
  - vasoconstriction in, 103, 104f
Hemostatic sheaths in endovascular therapy, 905, 905f
Hepatoblastoma, 1751
  - nomenclature on, 1381–1382, 1382t
  - in living donor kidney transplantation, 1386
  - in liver cancer, from colorectal cancer metastasis, 1386
  - laparoscopic, 1386–1387
  - historical aspects of, 1346
Hepatocellular carcinoma, 328, 328t
  - MRI in, 1381
  - nomenclature on, 1381–1382, 1382t
  - in living donor kidney transplantation, 1386
  - in living donor liver transplantation, 1386
  - in liver cancer, from colorectal cancer metastasis, 1386
  - laparoscopic, 1386–1387
  - historical aspects of, 1346
Hepatocellular carcinoma, 328, 328t
  - MRI in, 1381
  - nomenclature on, 1381–1382, 1382t
  - in living donor kidney transplantation, 1386
  - in living donor liver transplantation, 1386
  - in liver cancer, from colorectal cancer metastasis, 1386
  - laparoscopic, 1386–1387
  - historical aspects of, 1346
Hepatocellular carcinoma, 328, 328t
  - MRI in, 1381
  - nomenclature on, 1381–1382, 1382t
  - in living donor kidney transplantation, 1386
  - in living donor liver transplantation, 1386
  - in liver cancer, from colorectal cancer metastasis, 1386
  - laparoscopic, 1386–1387
  - historical aspects of, 1346
Hepatocellular carcinoma, 328, 328t
  - MRI in, 1381
  - nomenclature on, 1381–1382, 1382t
  - in living donor kidney transplantation, 1386
  - in living donor liver transplantation, 1386
  - in liver cancer, from colorectal cancer metastasis, 1386
  - laparoscopic, 1386–1387
  - historical aspects of, 1346
Hepatocellular carcinoma, 328, 328t
  - MRI in, 1381
  - nomenclature on, 1381–1382, 1382t
  - in living donor kidney transplantation, 1386
  - in living donor liver transplantation, 1386
  - in liver cancer, from colorectal cancer metastasis, 1386
  - laparoscopic, 1386–1387
  - historical aspects of, 1346

HOPDs. See Hospital outpatient departments
Horizontal gastroplasty, 1174f
Hormones
adrenal, 38, 1682–1685
breast cancer risk and, 555–556, 577, 579
in breast development and function, 544, 547f, 547–548
in cancer therapy, 342
gastric, 1109–1110
gut, 1227
in hypothalamic-pituitary-adrenal axis, 35f, 35–37
pancreatic, 1437f, 1437–1439
parathyroid. See Parathyroid hormone in postmenopausal therapy, 556
in shock response, 135–136
in stress response, 35–39
of thyroid, 1629–1633
Horner’s syndrome, 680, 698
Hospice care
in head and neck tumors, 654
pain management in, 2067
Hospital Compare, 17r
Hospital Consumer Assessment of Healthcare Providers and Systems, 16f, 413, 414f
Hospital outpatient departments
ambulatory surgery centers versus, 2157, 2160
description of, 2153
history of, 2153
ownership of, 2159
Hospitals, unplanned admission to, 2160
Host defense mechanisms
description of, 159–160
skin in, 159
small intestine in, 1226
spleen in, 1521
Hounsfield units, 807
Houston valves of rectum, 1261
Human umbilical vein, for infrainguinal bypass, 967
Human papillomavirus infections
anorectal, 1320
condyloma acuminata in, 526, 1320, 1320,
respiratory recurrent, 625
tumors associated with cervical, 1811
description of, 328, 328r
of head and neck, 630
of skin and subcutaneous tissue, 526
vaginal, 1810
vulvar, 1809
Human umbilical vein, for infrainguinal bypass, 967
Humerus
fractures of, 1884–1885
distal, 1885
proximal, 1884
in shaft, 1884–1885
osteosarcoma of, 1911
Humoral immunity, 342
Hungry bone syndrome, 99
Hunt-Hess grading system for subarachnoid hemorrhage, 1852, 1852r
Hurewitz, Mike, 409, 410r
Hürthle cell carcinoma of thyroid, 1651
Hyaluronic acid
in fetal wound healing, 282
in septic shock, 148, 176r, 177
Hydrochloric acid secretion, gastric,
Hydrocortisone
in hypothalamic-pituitary-adrenal axis,
Hydrocephalus
Hydrocele
Hydatid disease, 1370–1371
Hyaluronic acid
insulin resistance as cause of, 2116
in parenteral nutrition, 74
in septic shock, 147
Hyperinsulinemic hypoglycemia, 1204
Hyperkalemia
drug-induced, 89
etiology of, 89
signs and symptoms in, 89, 89r
treatment of, 95, 95r
Hyperkeratosis, 1001
Hypernatremia, 87f, 88–89
cancer, 99
etiologies of, 88
hypervolemic, 88
hypovolemic, 88, 94
postoperative, 427
signs and symptoms in, 88
in postmenopausal therapy, 95–96
Hyperparathyroidism, 1665–1681
bone disease in, 1663, 1667f,
1667–1668, 1670
in carcinoma of parathyroid, 1665, 1677
familial, 1677
gastrointestinal disorders in, 1668
hypercalcemia in crisis in, 1679
description of, 90, 1668–1669, 1677
hyperparathyroidism, 1685–1688
Hypermagnesemia
etiologies of, 91
signs and symptoms in, 89, 91
in shock response, 148, 176r, 177
Hydrochloric acid secretion, gastric,
Hydrocortisone, 2070r
Hydrocortisone
in adenocortical insufficiency, 1698
in Cushing’s syndrome, 1691
in hypercalcemia, 1680r
in septic shock, 148, 176r, 177
Hydroxyflurid acid, 521
Hydroxyflurid acid, 521
Hydromorphone, 2070r
Hydronephrosis, prenatal, 1777
Hydrops, gallbladder, 1404
Hydroxocobalamin, 256
Hydroydeoxycholic acids, 48f, 49r
Hydroxyethyl starch solutions, 94
21-Hydroxylase deficiency, 1693, 1746
Hydroxypropyl, 1974
5-Hydroxytryptamine, 50
Hygroma, cystic
description of, 649
ex utero intrapartum therapy (EXIT)
procedure in treatment of, 1754
Hymen, imperforate, 1745
in hyperparathyroidism, 1665–1681
in hyperparathyroidism, 1665–1681
Hyperparathyroidism, 1665–1681
bone disease in, 1663, 1667f,
1667–1668, 1670
in carcinoma of parathyroid, 1665, 1677
familial, 1677
gastrointestinal disorders in, 1668
hypercalcemia in crisis in, 1679
description of, 90, 1668–1669, 1677
disorders associated with, 1668
pancreateitis in, 1452
neonatal, 1677
neuropsychiatric disorders in, 1668
normocalcemic, 1676–1677
parathyroidectomy in description of, 1670–1676, 1675f
operative procedure in, 1672–1674,
1675f
preoperative localization tests in,
1671–1672, 1672r
in secondary hyperparathyroidism,
1679–1680
in tertiary hyperparathyroidism,
1679–1680
in primary, 1665–1678
clinical manifestations in, 1667–1668
diagnostic studies in, 1669–1670
differential diagnosis in, 1668–1669
etiology of, 1665–1667
geneic factors in, 1665–1666
treatment of, 1670–1676
renal disease in, 1667
secondary, 1679–1680
tertiary, 1680
urinary calculi in, 1667
algorithm on diagnosis and management of, 1138f
chronic, 1109
differential diagnosis in, 1137
in Helicobacter pylori infections, 1119
in Zollinger-Ellison syndrome, 1137
Hyperglycemia
in burns, 258
in critical illness, 39
in head trauma, 1839
hypotension in, 88
insulin resistance as cause of, 2116
in parenteral nutrition, 74
in shock, 147
Hyperinsulinemic hypoglycemia, 1204
Hyperkalemia
drug-induced, 89
etiology of, 89
signs and symptoms in, 89, 89r
treatment of, 95, 95r
Hyperkeratosis, 1001
Hyperlipidemia, pancreatitis in, 1441,
1452
Hypernatremia
etiologies of, 91
signs and symptoms in, 89, 91
in cancer, 99
etiologies of, 88
hypervolemic, 88
hypovolemic, 88, 94
postoperative, 427
signs and symptoms in, 88
in cancer, 99
gastrointestinal disorders from, 1124
in persistent pain, 2070
Idarucizumab, 113
Ifofamide
in lung cancer, 700
in soft tissue sarcoma of extremity, 1578
Ileal pouch anal anastomosis
description of, 1272–1273, 1273f, 1278
in familial adenomatous polyposis, 1292
pouchitis in, 1278
in ulcerative colitis, 1283
Ileitis, backwash
description of, 1236
in inflammatory bowel disease, 1280
Ileocolic artery, superior, 1260
Ileitis, Crohn’s disease of, 1284–1285
anatomy of, 1220
aneurysm of, with abdominal aortic aneurysm, 924
occlusive disease of, 898, 941–950
clinical results of treatment in, 901f, 949–950
endovascular therapy in, 948–949
trauma of, 212, 213f
in laparoscopy, 1821, 1821f
in inguinal hernia repair, 1621
transposition procedure for repair in, 215f
Iliac vein trauma, emergent abdominal exploration in, 212, 213f
Iliacus muscle, 1784
Iliococcygeus muscle, 1262, 1784
Iliofemoral bypass in aortoiliac occlusive disease, 946, 946f
Iliohypogastric nerve
anatomy of, 1602, 1604f
injury in inguinal hernia repair, 1619
Iliouninguinal nerve
anatomy of, 1601–1602, 1604f
injury in inguinal hernia repair, 1619
Iliosacral sign in appendicitis, 1332
Iliopubic tract, 1600, 1602f
Imaging systems in minimally invasive surgery, 463–465, 464f
Imatinib, 342, 343f
Imidazoles, 201
Imipenem-cilastatin, 166f
Immunocompromised hosts
infection and AIDS, 390
histoplasmosis in, 715, 716f
inflammatory bowel disease in, 1280
spleen function in, 1521–1522
small intestine function in, 1226
monocytes in, 54–55f
danger signaling in, 137, 138f
autophagy in, 40f
resistance to, 1586–1588, 1587f
Immunoglobulins, 160
Immunofluorescence, 496
Immunocompromised hosts
coccidiodomycosis in, 717
gastrointestinal disorders in, 1233
histoplasmosis in, 715, 716f
in HIV infection and AIDS. See HIV infection and AIDS perianal sepsis in, 1316
Immunofluorescence, 496
Immunoglubulins, 160
Immunohistochemistry, 496
Immunonutrition supplements, 2115
Immunoprecipitation, 496–497, 497f
Immunosuppression
corticosteroids in, 359, 359–360
in inflammatory bowel disease, 1281
in transplantation, 358–362
agents used in, 359f, 359–362
complications of, 360f, 363–364
gastrointestinal disorders in, 1324
heart, 390
historical aspects of, 356–357
intestine, 385–386
of kidney, 362, 369
of liver, 380, 385f
of pancreas, 374
islet cell, 379
Immunotherapy
adoptive transfer in, 344
antigen-specific, 344
description of, 491
nonspecific, 342, 344
Immunotherapy
adoptive transfer in, 344
antigen-specific, 344
description of, 491
nonspecific, 342, 344
Impedance, electrical
in cardiology, 443
in esophageal studies, 1022, 1028f
in plethysmography, 986–987
Imperforate anus, 1735–1737, 1736f
Impetigo, 524
Impingement syndromes in shoulder, 1896–1897
Implant-based reconstruction, 2009–2010
In situ cancer, 315
In situ simulation, 217f
Incidentaloma, adrenal, 1695–1697
diagnosis in, 1695–1696
differential diagnosis in, 1695f, 1695–1696
treatment in, 1696–1697, 1697f
Incision(s), abdominal, 1550–1553, 1552f
carcinoma in, 1554–1555
laparoscopic repair of, 1555
mesh repair of, 1555
in laparoscopy, 459–460, 1553
Incisional biopsy
in cancer diagnosis, 332–333
in soft tissue sarcomas, 1572
Incontinence
fecal, 1269–1270
in anal sphincter injury, 1322f
anatomic causes of, 126f
endoanal ultrasonography in, 1269, 1269f
neurogenic causes of, 1269
sacral nerve stimulation in, 1323
surgical repairs in, 1322f, 1322–1323
stress urinary, 1808–1809
Incretin, 1438
Indigestion, 1115
Indigestion, 1115
Indium-labeled autologous platelet scanning, 1541
Individualized medicine
in cancers, 349–350
genomic, 505–506
Indomethacin in ductus arteriosus patency, 215
Indolomethacin in ductus arteriosus patentcy, 760
Inducible nitric oxide synthase, 1356
Infantile hemangiomas, 626, 1996
Infants. See Children and infants
Infarction
myocardial. See Myocardium, infarction of
omentum, 1557–1558
Infections, 157–179
abdominal, 171–172. See also Abdominal infections
anorectal, 1313–1317, 1315f–1316f
aortic aneurysm in, 856
in aortoiliac reconstruction, 947
of appendix vermiform, 1331–1338
bacterial. See Bacterial infections from biologic warfare agents, 178–179
in bite wounds, 519–520, 520f
of bladder, 425, 1761
blood-borne, 177–178
of breast, 550
catheter-related. See Catheterization, infections associated with
Infections (Cont.):
of central nervous system, 1867–1868
of ears, 613–617
empyema in, 740–741
fungal, 162, 163t. See also Fungal
infections
gynecologic
pelvic inflammatory disease in,
1803–1804
screening for, 1789–1790
ulcer syndromes in, 1791, 1792r
of hand, 1931, 1947–1952
historical aspects of, 157–159
host defense mechanisms in, 159–160
skin in, 159
small intestine in, 1226
spleen in, 1521
inflammatory response to systemic
inflammatory response syndrome
in, 160, 161f
of liver, 172, 1369–1372
mediastinitis in, 735–736
microbiology of, 161–163
osteomyelitis in, 1867–1868
of pancreas, 172–173
in parenteral nutrition, 74
pathogenesis of, 159–161
peritonitis in, 171
pharyngitis in, 622
postoperative
description of, 174–175, 406–407
surgical site. See Surgical site
infections
postsplenectomy
antibiotics in, 1542
description of, 228
overwhelming, 1541–1542
vaccinations in prevention of, 1532,
1542
preoperative optimization during,
2126–2127
prevention and treatment of, 163–169
antibiotics in, 164, 165t–167t, 168
general principles in, 163
source control in, 163–164
pulmonary, 706–708
pyelonephritis in
description of, 1761
emphysematous, 1770
retroperitoneal, 1561
sepsis in. See Sepsis
sinusitis in, 617–622
of skin, 173–174, 524–527
of soft tissues
description of, 173–174, 175f–177f
necrotizing, 173, 175f, 524–526, 1316
of spleen
description of, 1529–1530
rupture in, 1530
of subcutaneous tissue, 524–527
of surgical site. See Surgical site
infections
transfusion-related, 123, 423, 423t
in transplant recipients, 363–364
in kidney transplantation, 369–370
in liver transplantation, 385
in pancreas transplantation, 377
of urinary tract, 1761–1762
viral, 162–163, 163t. See also Viral
infections
of wounds, 286–289, 295
Inferior gluteal artery perforator, 2010,
2011f–2012f
Inferior pedicle reduction mammaplasty,
2020f
Inferior vena caval filter, 1179
Inflamasomes, 28, 30, 39
Inflammatory bowel disease, 1279–1286
cholangitis in, primary sclerosing, 1281,
1417
colorectal cancer in
risk for, 1282, 1285, 1289
screening for, 1295t
Crohn’s disease, 1235–1240, 1279.
See also Crohn’s disease
drug therapy in, 1281–1282
epidemiology of, 1279
extraintestinal manifestations of, 1281
determinate colitis, 1279
nutrition in, 1282
ulcerative colitis, 1279. See also Colitis,
ulcerative
Inflammatory carcinoma of breast
description of, 601–602, 602f
noninflammatory cancer versus,
601–602, 602t
Inflammatory response
acute phase proteins in, 1354
anti-inflammatory response and,
cholinergic pathways in, 35
cell-mediated, 56–58
central nervous system in, 34–39, 36f
C-type lectin receptors in, 33
“damage-associated” molecules in, 27
dendritic cells in, 55–56
description of, 27–28
lectin-like receptors in, 33
macrophage migration-inhibiting factor
affecting, 37–38
macrophages in, 54–55
mediators of, 42–50
in shock, 137–141
neutrophils in, 54
in pancreatitis
acute, 1443
chronic, 1460
platelets in, 56
receptor for advanced glycation end
products in, 29, 33–34
in shock, 137–141
systemic
cellular injury in, 28–34
overview of, 27–28
trajectories of, 29f
systemic syndrome in, 35, 160, 161t
in cardiopulmonary bypass, 809
definition of, 29
diagnostic criteria on, 160–161, 161t,
426, 426t
mortality risk in, 426, 426t
postoperative, 426
vagus nerve in, 35
in wound healing
description of, 272–273, 273f–274f
fetal, 282
Infliximab
in Crohn’s disease, 1238, 1282
in ulcerative colitis, 1282
Informed consent, 2062–2064. See also
Consent, informed
Infringedinal bypass grafts
adjunctive modifications for, 967–968
autogenous vein for, 967
cryopreserved grafts, 967
femoropopliteal occlusive disease
reated with, 967–968
human umbilical vein, 967
prosthetic conduits, 967–968
Infraopliteal occlusive disease
classification of, 952
dendovascular therapy in, 960, 960t, 962
Infundibulopelvic ligaments, 1786
Inguinal canal, 1600, 1601f
Inguinal hernia, 1599–1621
age-related differences in, 1600r, 1604
anatomy in, 1600–1603, 1601f
anterior perspective, 1600, 1604f
circle of death in, 1603
ligaments in, 1600, 1601f
nerves in, 1601–1602, 1603f–1604f
posterior perspective, 1600, 1603f
triangle of doom and triangle of pain
in, 1603, 1605f
anesthesia in, 1620
in children, 1604, 1743–1744. See also
Children and infants, inguinal
hernia in
classification of, 1602r
in collagen disorders, 1605
complications in, 1618–1621, 1619f
in connective tissue disorders, 1606f
conservative management in, 1608
diagnosis of, 1606–1608
history-taking in, 1606
imaging in, 1607, 1608f
physical examination in, 1607, 1607f
differential diagnosis of, 1607, 1607f
direct
in children, 1743
classification of, 1600
diagnosis of, 1607
elective repair of, 1608
emergent repair of, 1608
endoscopic repair in, transabdominal
preperitoneal, 1613
historical treatment of, 1599–1600
incarceration of, 1608
in children, 1744
emergent repair in, 1608
indirect
classification of, 1600
diagnosis of, 1607
laparoscopic repair in
in children, 1744
complications of, 1620
intrapерitoneal onlay mesh procedure
in, 1616
patient position and room setup for,
1614f
totally extraperitoneal procedure,
1613–1616, 1615f
transabdominal preperitoneal, 1615f
mesh repair in. See Mesh repair, in
inguinal hernia
open repair of, 1609–1613
antibiotic prophylaxis in, 1609
Bassini technique, 1610, 1611f
Desarda technique, 1611–1612, 1612f
incisions in, 1609f
Lichtenstein technique, 1612–1613,
1613f, 1621
McVay technique, 1610–1611, 1611f
plug and patch technique, 1612–1613
Intracellular fluid, 83, 84f–85f
composition of, 83–84
osmolality of, 84
Intracerebral hemorrhage, 1839
Intracranial hemorrhage. See Hemorrhage, intracranial
Intracranial pressure
increase in, 447–448, 1831–1832, 1833f
brain herniation in, 1831–1832, 1833f
in craniostenosis, 1992
in head trauma, 217–218, 447–448
pressure-volume curve in, 1832f
monitoring of, 447–448, 1831–1832
intracranial fiberoptic pressure transducer in, 1831
in traumatic brain injury, 1840–1841
ventriculostomy in, 448, 1831
Intracranial tumors
description of, 1854–1859
metastatic, stereotactic radiosurgery in, 1872
Intraductal papillary mucinous neoplasm (IPMN), 1502, 1502f–1504f
Intraepithelial neoplasia
metastatic, stereotactic radiosurgery in, 1854–1859
description of, 1854–1859
Iodine
dietary
deficiency of, 1641, 1641f, 1650
sources of, 1629
metabolism of, thyroid function in, 1629
radioactive
and fibrinogen uptake in venous thromboembolism, 987
in hyperthyroidism therapy, 1637
in thyroid cancer therapy, 1651–1654, 1653f
in thyroid imaging, 1634, 1634f
in hyperthyroidism, 1634, 1635f
Ion channels, 488
Iprilimumab in cancer immunotherapy, 345
IPMN, 1502, 1502f–1504f, 1504
Irritable bowel syndrome, 1269
Iron
Iron metabolism of, thyroid function in, 1634
J
Jahss maneuver in hand fractures, 1934, 1934f
Janeway gastrostomy, 1155
Janeway lesions, 1155
Janus kinase, 47, 50, 51f
Jaundice, 1353–1354, 1737–1740
in biliary atresia, 1737–1739
in children and infants, 1737–1740
in choledochal cysts, 1739–1740
in inspissated bile syndrome, 1738
intrahepatic causes of, 1354
obstructive, 1399
in pancreatic cancer, 1487, 1489
physiologic, 1737
posthepatic causes of, 1354
prehepatic causes of, 1354
Jefferson fracture, 1843, 1900
Jejunostomy, 72
Jefferson
anatomy of, 1220, 1220f
atresia of, 1725f
diverticula of, 1248–1249, 1249f, 1250
enteral nutrition access via, 71, 71f, 72
as esophageal substitute, 1089, 1089f, 1090f
neoplasms of
description of, 1243, 1243f
treatment of, 1244
perforation of, 1251–1252
pneumatosis intestinalis in, 1253f
sarcoma of, 1584
short bowel syndrome in resection of, 1254
Jod-Basedow hyperthyroidism, 1637–1638
Johnson, Robert Wood, 271
Johnson classification of peptic ulcer disease, 1122, 1123f
Joint Commission protocol for prevention of wrong-site surgery, 411–412
Joints
arthrodesis of, 1906
in hand and wrist, 1945
spinal, 1866
arthroplasty of, 1906–1910. See also Arthroplasty
injection techniques, 1905
Journal clubs, 2190, 2190f
Journals, 2074
Jugular vein
as access for pulmonary artery catheterization, 438
oxygen saturation in, 448–449
trauma of, 220
Justice principle, 2061
K
Kaiser system, 9
Kallikrein, 49, 106
Kanavel’s sign, 1950
Kaposi’s sarcoma, 527, 535
Kappa coefficient, 2148
Kapandriti flap, 634, 635f
Kasai portoenterostomy, 1738f, 1738–1739
Kawasaki syndrome, 971
Kawashima procedure, 784
Kayexalate in hyperkalemia, 95, 95t
Kelling-Madlener procedure, 1137f
Keppra, 1839

**Kidneys**

- **Ki67 antigen**, 578
- **Ketone bodies**, 61–62
- **Ketogenesis**, 64–65
- **Ketoconazole**, 715
- **Ketamine**
  - administration of, 2031
  - analgesic uses of, 2120
  - anesthetic uses of, 2030–2031
  - benzodiazepines and, 2031
- **Ketoconazole**, 715
- **Ketogenesis**, 64–65
- **Ketorolac**, 2031
- **Ketorolac**, 2031
- **Kiev, Geoffrey**, 543
- **Ki67 antigen**, 578
- **Kidneys**
  - acute failure of
    - enteral nutrition formulas for, 70–71
    - fluid and electrolyte balance in, 99
    - postoperative, 422, 422r
  - anatomy of, 1759–1761
  - calculi in, 1667
  - cancer of, 1772–1773
    - epidemiology of, 307r, 308f, 310f
    - genetic factors in, 1773
    - nephrectomy in, 1772–1773,
      - 1773f
  - contrast agents affecting, 417, 859
    - cysts of, 1773f
    - enteral nutrition formulas in failure of, 70–71
  - in hyperparathyroidism, 1665–1666
  - ischemia of, in aortic aneurysm repair, 869
  - monitoring of, 447
    - postoperative disorders of, 422–423
    - preoperative evaluation of, in aortic aneurysm repair, 861
  - pylonephritis of, 1761
  - transplantation of, 368–374
    - cardiovascular disorders of, 369
    - complications in, 369–370, 373
    - contraindications to, 368
    - en bloc grafts, 372, 372f
    - historical aspects of, 356–357, 368
    - immunologic evaluation of, 370
    - immunosuppressive therapy in, 362, 369
    - indications for, 368
    - infections, 369–370
    - intraoperative care in, 373
    - living donor in, 357, 366, 368
    - operative technique in, 367f
    - medical evaluation in, 369–370
    - multiple renal arteries grafts,
      - 372, 372f
    - organ preservation methods and time in, 365
    - pancreas transplantation after, 374, 378
    - postoperative care in, 373
    - preoperative care, 373
    - preoperative evaluation in, 368–370
    - psychosocial evaluation of, 370
    - recipient surgical procedure in, 367f,
      - 370–371, 371f
    - results in, 373–374
    - simultaneous with pancreas transplantation, 374, 378
    - surgical evaluation of, 370
    - survival rates in, 373–374
    - tumors, 369
    - urologic evaluation of, 370
    - vascular evaluation of, 370
    - trauma of, 233f, 233–234, 1765
      - in children, 1753
    - grades of injury in, 1765, 1765f
    - indications for surgery in, 1765
    - nonoperative treatment of, 234, 1765
    - vascular injuries in, 233, 1765r
    - tubular necrosis of, acute, 422–423
    - Wilms’ tumor of, 1747f, 1747–1748
    - kinetic resistance in tumors, 346
    - King, Dr. Martin Luther, 7, 8f
    - King, Josie, 405, 410f
    - Kininogen, 49
    - Kinins, 49
    - KIT gene, 322
    - Kirkpatrick four-level scale, 2178, 2178f
    - Köhler’s syndrome, 1971, 1971f
    - Klippel-Trenaunay syndrome, 2000
    - Klumpke’s palsy, 1848
    - Knee
      - arthroplasty of, 1908, 1908f–1910f
      - alignment of components, 1908
      - background of, 1908
      - bearing surfaces in, 1908
      - osteolysis in, 1910
      - surgical approach, 1908
      - collateral ligaments of, 1898
      - contracture of, in cerebral palsy, 1915
      - cruciate ligaments of, 1898–1899, 1899f
      - dislocation of, 1890–1891
      - fractures of, 1897–1899
      - ligament injuries of, 1898–1899
      - meniscus tears of, 1897–1898, 1899f
      - posterior ligament injury, 1899
      - Knockout mouse technology, 502f
      - embryonic lethality in, 503
      - target vector in, 501–503
      - Knowledge, medical, 4r
      - Knudson’s two-hit hypothesis, 320, 324f
      - Koch, Robert, 158
      - Kocher, Emil Theodor, 5, 6f, 1625
      - Kocher maneuver, 212
      - Kock pouch ileostomy, 1278
      - in familial adenomatous polyposis, 1292
      - in ulcerative colitis, 1283
      - Korotkoff sounds, 434
      - Kouchoukos’s technique in aortic aneurysm repair, 863
      - KRAS, 493
    - KTP (potassium thionyl phosphate) laser, 466
    - Kulchitsky cells in tracheobronchial tree, 669, 669f
    - Kupffer cells, 1355
    - **L**
      - Labeling of surgical specimens, frequency of errors in, 403, 404f
      - Labia majora and minora, 1785
      - Labyrinthitis, 615
      - Lacerations
        - in head and neck injuries
          - closure in, 627–628, 628f
        - of scalp, 188, 194, 1834
        - of kidney, 1765f
        - of liver, 225
        - in obstetric injuries, 1805
        - of small intestine, 229
        - of vagina in children, 1745
      - Lactate
        - in gluconeogenesis, 61, 61f, 65
        - serum levels of
          - in burns, 254
          - in shock, 142, 143f, 152
      - Lactation
        - breast changes in, 545, 545f, 548f
        - mastitis in, 550
      - Lactic acidosis, 92
      - Lactose digestion and absorption, 1223, 1224f
      - Lacunar ligament anatomy, 1600
      - Ladd procedure in malrotation, 1726, 1727f
      - Ladd’s bands, 1725
      - Lambert-Eaton syndrome, 1292
      - Lamina propria, 1226, 1226f
      - Laminotomy, 1903
      - **Lancet Commission on Global Surgery**, 2086, 2086f
      - Langerhans cell histiocytosis, 722
      - Langerhans cells, 515, 710, 722
      - Langer’s lines, 1969, 1970f
      - Laparoscopic single-site surgery, 453
      - Laparoscopic coagulation shears, 467, 467f
      - Laparoscopy
        - with abdominal lift device, 456
        - access for, 459f, 459–460
        - adjustable gastric banding
          - background on, 1187
          - comparisons, 1199
          - complications of, 1188–1189, 1204
          - history of, 1168–1169
          - illustration of, 1174f
          - indications for, 1173
          - Lap-Band, 1188f
          - mechanism of, 1176f
          - patient selection for, 1187
          - short-term outcomes of, 1192
          - technique for, 1187f–1189f,
            - 1187–1188
          - weight loss after, 1175
        - adrenalec...
Laparoscopy (Cont.):
cholecystectomy. See Cholecystectomy.
laparoscopic
colon resection in
in colorectal cancer, 1302–1303
description of, 1270
common bile duct exploration in,
1413–1414, 1415f
complications of, 456
bladder trauma in, 1621, 1821
in inguinal hernia repair, 1620–1621
in Crohn's disease, 1240, 1283
in elderly, 474
diabetic response to, 456–457
esophagectomy in, in cancer of,
esophagus, 1075f, 1075–1076
Fundamentals of Laparoscopic Surgery, 2168, 2179
gastric
in cancer diagnosis, 1145–1146
description of, 1159–1160
in gynecologic disorders, 1800–1801
hand-assisted access for, 461, 461f
description of, 1270
splenectomy in, 1534–1535
Hasson technique, 460, 1616
hialtal hernia repair, 1048
historical aspects of, 454
incisional hernia repair, 1555
incisions
description of, 459–460, 1553
single, 1535, 1537
inguinal hernia repair in children, 1744
complications of, 1620–1621
description of, 1613–1616
liver resection in, 1386–1387
nephrectomy in, in kidney cancer, 1773
Nissen fundoplication in, 1040,
1040f–1042f
ovarian cancer assessments, 1818f
in pancreatic cancer, 1488f, 1488–1489
physiology and pathophysiology in,
455–457
pneumoperitoneum in, 455–456
positioning of patients in, 458, 459f
in pregnancy, 472–473
appendectomy in, 472–473
cholecystectomy in, 1405
Roux-en-Y gastric bypass
complications of, 1183–1185
configuration of, 1180f
description of, 1175
tentoenterostomy of, 1182f, 1184f
gastric pouch in, 1181, 1182f
gastrojejunostomy in, 1183,
1183f–1184f
glycemic control after, 1177, 1178f
hematemeses after, 1185
marginal ulcers secondary to, 1185
mechanism of, 1176f
overview of, 1180–1181
patient selection for, 1180–1183
port scheme for, 1181f
small bowel obstruction in, 1184
technique for, 1180f–1184f,
1181–1183
single-incision. See Single-incision
laparoscopic surgery
skill training in, 2167–2168
sleeve gastrectomy
background on, 1185–1186
complications of, 1186–1187, 1206
glycemic control after, 1178f
history of, 1168
mechanism of action, 1176f
patient selection for, 1185
port scheme for, 1185f
present-day status of, 1169
revision for, 1206
technique for, 1185–1186, 1186f
in small intestine obstruction, 1232
splenectomy in, 1533–1534,
1534f–1536f
team approach to, 457–458
telescopes in, 462, 464f
three-dimensional, 464–465
ultrasongraphy in, 1413
venous thromboembolism in, 456
Veress needle in, 459f, 459–460
video systems in, 463
virtual reality training of, 2171f,
2171–2172, 2172f, 2179
Lapatinib, 343f
LaPlace's law, 854
large cell carcinoma of lung, 672
large intestine, 1259–1324
anatomy of, 1260–1263
clinical evaluation of, 1264–1270
coitus of, 1309
colon in. See Colon
congenital anomalies of, 1263
diverticular disease of, 1286–1288
embryonic development of, 1259–1260
in immunocompromised patient, 1324
inflammatory bowel disease of,
1279–1286
in laparoscopy, 1821
megacolon of, 1308
neoplasms of, 1288–1303
physiology of, 1263–1264
pseudo-obstruction of, 1308–1309
sarcoma of, 1584
trauma of, 1321–1322
volvulus of, 1306–1308
Laryngeal mask airway, 2037
Laryngeal nerves
anatomy of
general, 1628f–1630f
1628–1629, 1630f
injury of, in surgery
description of, 419
in thyroidectomy, 1658, 1662
lung cancer involving, 681, 727, 729f
in thyroidectomy, 1658, 1659f, 1660
Laryngectomy
in esophageal cancer, 640
in hypopharyngeal cancer, 640
in laryngeal cancer, 641f
in pectoralis flap reconstruction in,
640
in rehabilitation in, 642
in supracricoid technique, 642
Laryngopharyngectomy, 640
Larynx
cancer of, 640–642
anatomy in, 631, 640
diagnosis of, 640
diastolic view of, 641f
epidemiology of, 307f
lymph node involvement in, 641
lymph node involvement in, 641
lymph node involvement in, 641
lymph node involvement in, 641
lymph node involvement in, 641
lymph node involvement in, 641
lymph node involvement in, 641
preservation of larynx in, 642–643
rehabilitation in, 643
staging of, 641
disorders of, 624–626
fracture, 198f, 222
granulomas, 625, 625f
recurrent papillomavirus infections of,
625
voice disorders, 624
Laser therapy
in atherectomy, 964–965
description of, 465–467
in hemangiosomas, 627
in prostate hyperplasia, benign, 1764
in transmyocardial revascularization,
813–814
in varicose veins, 996
in vocal cord cancer, 642
Latarjet nerve, 1102, 1104f
Lateral ligaments of rectum, 1261
Lateral malleolar defect, 1261
Latissimus dorsi flaps
in breast reconstruction, 593–594, 1982f
in chest wall reconstruction, 2012
in head and neck reconstruction, 651
Latissimus dorsi myocutaneous flap, 1979,
1982f
Lavage, peritoneal
in abdominal trauma, 202, 202f, 202r
in pregnancy, 240
LD50, 2029
Le Fort fractures, 628–629, 629f
Leadership in surgery
clinical outcomes affected by, 11–12
on core competencies, 4t
diversity in, 12
ethical decision making, 14
fundamental principles of, 4–11
healthcare cost affected by, 11–12
interpersonal and communication skills,
4t
introduction of, 3
medical knowledge in, 4t
mentoring of, 18
modeling of, for medical students,
18–19
objective structured clinical examination,
14–16
outcome measurements for, 19t,
19–20
patient care skills, 4t
practice-based learning and
professionalism, 4t
recruitment, 11
time management, 10
training programs in, 13
training modules in, 13
vision of, 5–7
wellness, 11
willingness, 7–10
Laparoscopic Surgery, 2168, 2179
Index
Leapfrog Group, on standards of health
care quality and safety, 408
Liver, cancer of (Cont.):
hepatic resection in, 1379–1380
liver transplantation in, 381–382, 1377, 1380, 1751
Milan criteria on, 381, 1380
in pancreatic cancer metastasis, 1488f, 1498–1499
radiofrequency ablation in, 1380
risk factors for, 1376
treatment algorithm on, 1376, 1378f
caudate lobe of, 1347f
cirrhosis of, 1362–1365. See also Cirrhosis
clotting factors synthesized in, 1353
cogaulopathy in disorders of transplantation of liver in, 381–382
cysts of
in Caroli’s disease, 1374
congenital, 1373
in hydatid disease, 1370
laparoscopic resection in, 1386
in polycystic disease, 1373–1374
cytokine response by, 1354
end-stage disease of
in Budd-Chiari syndrome, 1369
in cirrhosis, 1362, 1365
MELD score in, 381, 474, 1365, 1380
eneteral nutrition in failure of, 71
fatty infiltration of, in inflammatory bowel disease, 1281
function tests of, 1352–1353
glycogen in, 60
hemangiomata of, 1375, 1375f
hepatocellular injury of, 1353–1354
historical aspects of surgery on, 1346
hyperplasia of, focal nodular, 1375f, 1376
incidental mass of
diagnostic algorithm on, 1372f
differential diagnosis in, 1373, 1373f
infections of, 1369–1372
inflammatory bowel disease involving, 1281
ischemia of
in aortic aneurysm repair, 869
in preconditioning for surgery, 1385
liposarcoma metastatic to, 1573f
molecular signaling pathways in acute phase reaction, 1354–1355
heme oxygenase, 1356–1357
lipopolysaccharide signaling, 1355
nitric oxide, 1355–1356
toll-like receptors, 1357
nitric oxide in, 1355–1356
physiology of, 1351–1354
portal hypertension of, 1365–1367
quadrate lobe of, 1347
radiologic evaluation of, 1357–1360
radionuclide scans of, 1399
resection techniques, 1381–1386.
See also Hepatocarcinoma
soft tissue sarcoma metastatic to, 1581
transplantation of, 379–385, 1368
in acetaminophen hepatotoxicity, 381, 1362, 1362f
in acute failure, 381, 1362
in biliary atresia, 383, 1738
in Budd-Chiari syndrome, 1369
in children, 383, 1738
in hepatocellular carcinoma, 1751
split-liver procedure in, 383
in cholangiocarcinoma, 381, 1377
in cirrhosis, 380–381, 1368
complications in, 385
contraindications to, 382
donor in, 380
in hepatocellular carcinoma, 381, 1377, 1380, 1751
historical aspects of, 379–380
immunosuppressive therapy in, 380, 385
indications for, 380t, 380–381, 1362t, 1368, 1380
indicators of graft function in, 384–385
with living transplantation, 386, 387f
living donor
description of, 366, 367f, 380
operative technique in, 383, 384f–385f
Milan criteria on, 381
orthotopic, 382
piggyback technique in, 382–383
in polycystic liver disease, 1374
postoperative care in, 383
primary nonfunction of graft in, 385
in primary sclerosing cholangitis, 380, 1418
split-liver procedure, 383
surgical procedure in, 382f–383f, 382–383
survival rates in, 383
in varical bleeding, 381, 1368
venovenous bypass in, 382
waiting time for, 381, 383
trauma of
in children, 1752, 1753f, 1754f
grading scale on, 1753, 1753f
treatment options, 1379f
ultrasound of, 1357–1358
Live-tweeting conferences, 2190, 2191f
Living donor transplantation procedures
kidney
description of, 357, 366, 368
operative technique in, 367f
liver
description of, 366, 380
operative technique in, 383, 384f–385f
pancreas, 377–378
Living wills, 2064
Lobectomy
hepatic
left, 1383–1385, 1384f
right, 1383, 1384f
pulmonary, functional status assessment in, 691–692, 694f
temporal lobe, in epilepsy, 1869
video-assisted thoracic surgery, 702f
Lobular carcinoma of breast in situ
histopathology in, 564, 564f
pathology in, 553–554
pleomorphic, 554
management of, 580–582
invasive, 566f, 566–567
Local anesthetics
analgesic uses of, 2120f
description of, 2034
history of, 2027–2028
lidocaine, 2031
Local flaps, 1977
Locoregional lymphadenectomy, 1576
Loeys-Dietz syndrome, 855
Long esophageal myotomy, 1060f, 1060–1062, 1061f–1062f
Loop ileostomy, 1273f, 1276–1277
Low molecular weight heparin, for venous thromboembolism prevention, 2053
Lower extremity
arterial occlusive disease of, 950–969
acute bypass graft thrombectomy in, 956
clinical manifestations of, 951f, 954–955
complications of treatment in, 956–958, 957f
description of, 950, 952, 954
embolectomy in, 956
endovascular therapy in, 955–956
etiology of, 952, 954
reperfusion syndrome in, 957
chronic arterectomy in, 964–966
clinical manifestations of, 958f, 958–959, 959f
endovascular therapy in, 959–960, 965
in femoropopliteal disease, 966
surgical treatment in, 966
classification of, 951–952, 954f–955f
computed tomography angiography in, 951, 951f
diagnostic evaluation in, 950–951
differential diagnosis of, 951
epidemiology of, 950
intermittent claudication in, 951, 953–954f
pain in, 898
physical examination in, 899
symptoms in, 898
compartment syndrome of
fasciotomy in, 237, 238f
in ischemia treatment complications, 956–958, 957f, 957f
in trauma, 237
edema of, after surgical treatment of limb ischemia, 967
lymphedema of, 1001, 1001f, 1003
sarcoma, 2008f
trauma, 2005
ulcers of
ischemic, 958, 958f, 959f
neuropathic, 958f, 959f
veins in, 981–982
Loxiglumide, 1113
LUMBAR association, 1996
Lumbar spine
intervertebral disc herniation
description of, 1864, 1865f, 1865t, 1903
level and symptoms in, 1865f
surgical treatment in, 1903
spondylolisthesis of, 1862
stability of, 1861–1862
stenosis of, 1903
trauma of, 1843, 1844f
fractures in, 1902–1904
orthotic devices in, 1847
Lumbral muscles, 1929
Luminal tamponade, in variceal hemorrhage, 1367
Lumpectomy. See also Breast conservation surgery
in breast cancer, 580–581
description of, 591, 600
Lunate bone
anatomy of, 1926, 1928f
dislocation of, 1934
Lund and Browder chart, 252
Lung(s), 667–720
abscess of
chest radiograph in, 707, 707f
clinical features and diagnosis, 706–708
etiologies of, 706, 706t
primary, 706, 706f
secondary, 706, 706f
surgical drainage of, 708, 708t
acute injury of
description of, 419–420
transfusion-related, 121, 122t
anatomy of
alveolar spaces in, 668
general, 667–668, 717
lymphatic, 667–668, 668f–669f
segmental, 667, 668f
tracheobronchial tree in, 668, 669f
biopsy of, 685
blood supply of
anatomy of, 717
in sequestration, 1714f, 1715
computed tomography of
in cancer assessment, 683, 685, 687
screening, 675
in solitary nodule, 677–678, 678f
tumors of, 691–692
in children, 1752
in elderly, 243
end-stage disease of
lung transplantation in, 720, 720f
lung volume reduction surgery in, 719
foreign body in, in children, 1716–1717
function testing of
in aortic aneurysm repair, 861
in lung cancer, 691–693, 694f
in hemoptysis, massive, 717–719
histology of, 668–669, 669f
hyperplasia of, 669–670, 672
infections of, 706–708
metastatic lesions, 678–680
in pneumothorax, spontaneous, 705–706
soft tissue sarcoma metastasis to, 1573, 1581, 1581f
solitary nodule in, 676–678
calcification of, 677–678
corona radiata sign in, 678, 678f
differential diagnosis in, 677
imaging in, 677–678, 678f
management algorithm on, 686f
surgical approaches to, 701–703
transplantation of
complications in, 391
evaluation of, 390
with heart transplantation, 388–391, 720
historical aspects of, 389
indications for, 390, 720
living donor in, 366, 390
organ preservation methods and time in, 365
postoperative care in, 391
rejection in, 720
surgical procedure in, 390–391
survival rates in, 720, 720f
trauma of, 224
volume reduction surgery of, 719
Lung cancer, 680–701
chemotherapy in. See Chemotherapy, lung cancer treated with
diagnosis, 684f
drug resistance, 346
early-stage, 695
epidemiology of, 306–308, 307f, 308f–309f, 310f, 673–675
etiologies of, 673–675
functional status assessment in, 691–693, 694f
histology of, 670–672
lymph node involvement in
anatomy of, 667–668, 668f–669f
assessment of, 687–688, 688f
in locoregional advanced disease, 697–698
staging of, 687–691, 690–691t
management, 683–693, 685t
metastatic
description of, 683, 686–687
staging of, 689–691, 690–691t
osteopathopathy in, hypertrophic, 681, 682f
paraneoplastic syndromes in, 681t
preinvasive lesions in, 669–670
risk factors for, 675, 675f–676f
screening for, 331t, 332, 675–676
signs and symptoms in, 680t, 680–683
in metastasis, 683
nonpulmonary thoracic, 680t, 680–681
pulmonary, 680, 680t
small cell. See Small cell carcinoma of lung
staging of
description of, 689–691, 690–691t
management affected by, 693–701
survival rates in
in chemotherapy, 700f
description of, 693
treatment in, 693–701
in early-stage disease, 693, 695–697
in locoregional advanced disease, 697–698
radiofrequency ablation, 695
in stage IV disease, 698
stereotactic body radiotherapy, 695–696
Lupus anticoagulant in antiphospholipid syndrome, 112
Luteinizing hormone, 547–548
Lymph nodes
in breast cancer
axillary lymph nodes in. See Axillary lymph nodes, in breast cancer
description of, 576f–577f
historical aspects of surgery, 542
internal mammary lymph nodes in, 587
mapping of, 338f, 338–339
sentinel node dissection in, 338–339, 590–591
staging of, 575
cervical. See Cervical lymph nodes in colorectal cancer
colon carcinoma involving, 1299
description of, 338, 1296, 1296f
rectal carcinoma involving, 1301
staging of, 1295, 1297f
description of, 333, 338–339
in esophageal cancer
dissection of, 1076
staging of, 1069–1070, 1071t
in gallbladder cancer, 1422
Halsted theory on dissection of, 338
in head and neck tumors, 646f, 646–648, 648f
anatomy and levels in, 645f, 646–647
of larynx, 641
of lips, 634
of oral cavity, 634
of oropharynx, 639
staging of involvement, 633–643
in lung cancer
anatomy of, 667–668, 668f
assessment of, 687–688, 688f
in locoregional advanced disease, 697–698
staging of, 687–691, 690–691t
mapping of, 338f, 338–339
mediastinal. See Mediastinal lymph nodes
in melanoma, 530–533, 532f–533f
in pancreatic cancer, 1485–1487
sentinel. See Sentinel lymph nodes
in soft tissue sarcomas, 1573, 1576
in squamous cell carcinoma, 529–530
in stomach cancer
dissection of, 1101, 1147
staging of, 1145f
in testicular cancer, 1772
in thyroid cancer, 1649f
dissection of, 1661, 1662f
medullary, 1656
in vaginal cancer, 1811
in vulvar cancer
anatomy of, 1809, 1809f
staging of, 1809–1810, 1810f
Lymphadenitis, cervical, 1710
Lymphadenopathy, cervical, 1710
Lymphangiography
imaging of, 1002f–1003f
in lymphedema, 1002
Lymphangiomas, 649, 1711–1712
Lymphangiosarcoma, 1568
Lymphatic malformations, 1711
Lymphatic malformations, 1711
Lymphatic malformations, 1711–1712, 1999, 1999f
Lymphatic system
of anal canal, 1304, 1304f
of colon, 1261
of esophagus, 1014f, 1014–1015
of head and neck, 645f, 646–647
of liver, 1350–1351
of lungs, 667–668, 668f
of pancreas, 1433, 1434f
of stomach, 1100–1101, 1103
of thyroid, 1629, 1631f
antibiotics for, 1003
chronic, lymphangiosarcoma in, 1568, 1569f
clinical diagnosis of, 1001f, 1001–1002
Lymphedema (Cont.):
congenital, 1001
hyperkeratosis in, 1001, 1001f
intermittent pneumatic compression
therapy for, 1003
management of, 1003
in modified radical mastectomy, 593
pathophysiology in, 1001
praecox, 1001
primary, 1001
radiologic diagnosis of, 1002, 1002f–1003f
reconstructive surgery in, 1003
secondary, 1001
Lymphedema praecox, 2015
Lymphocyte, 55, 515
B
in chyle, 742r
in small intestine immune function, 1226
in transplant rejection, 358
T
allorecognition by, 358
description of, 55, 55f
in small intestine immune function, 1226
in transplant rejection, 358
in wound healing, 272–273
Lymphoepithelial cysts, of pancreas, 1505
Lymphogranuloma venereum, 1791, 1792
Lymphoma
of appendix, 1339
of bone, 1914
of breast, 602
of central nervous system, 1858
colorectal, 1303
epidemiology of, 307r, 308f, 310r
in HIV infection and AIDS, 1858
Hodgkin’s. See Hodgkin’s lymphoma
mediastinal
biopsy in, 729
differentiated from thymoma, 731
incidence of, 727, 728r, 734
signs and symptoms in, 729r
of pancreas, 1505
of small intestine
description of, 1242r, 1243
outcome in, 1245
treatment in, 1245
splenectomy in, 1523r, 1527–1528
of stomach
description of, 1139, 1149
in Helicobacter pylori infections, 1118
of thyroid, 1657–1658
in transplant recipients, 364
Lymphopenia, 40
Lymphoproliferative disorder
mediastinal, 727
posttransplant, 363–364
Lymphoscan
description of, 590, 1002f
in lymphedema, 1002
Lymphosarcoma
Amsterdam criteria on, 324, 325r
description of, 324–325, 1292–1293, 1295r
endometrial cancer in, 1815
Lysyl oxidase, 1974
M
M cells, 1226
MACIS scale on papillary carcinoma of thyroid, 1648
Mackenrodt’s ligament, 1787
Macroadenomas, pituitary, 1857
Macrautophagy, 40
Macrocysts, breast, 552–553
Macrodactyly, 1961
Macrolides, 167
Macromastia, 2018
Machado-Joseph disease, 1889
Magnetic resonance imaging
absorption of, 1225
interferon-gamma stimulation of, 47
in microbial invasion response, 160
in wound healing, 272, 273f–274f, 274r
Macrography antigen 1 (Mac-1), 57r
Macrophage migration-inhibiting factor, 37–38
MADH4 gene in juvenile polyposis, 322r
Mafenide acetate, 257
Maffucci’s syndrome, 1911, 1957
Magnesium
absorption of, 1225
dietary intake of, 91
in parenteral nutrition, 94r
serum levels of
in cancer patients, 99
description of, 91, 96
in hypermagnesemia, 89r, 91, 95–96
in hynomagnesemia. See Hypomagnesemia
in kidney transplantation, 373
supplemental, in children and infants, 1707
Magnetic resonance angiography
in aortic aneurysm, thoracic, 859
in aortic dissection, 880, 884
in carotid artery occlusive disease, 909–910
in celiac artery compression, 931f
description of, 902, 902f
in lower extremity arterial occlusive disease, 951
in mesenteric ischemia, 931, 931f–932f
in renal artery stenosis, 936f, 937, 937f
Magnetic resonance imaging
of abdominal wall tumors, 1557
in aortic aneurysm, 859
in aortic dissection, 880
in biliary disorders, 1399
in breast cancer, 560, 572–574, 574f–575f
in carotid artery occlusive disease, 910
in celiac artery compression, 931f
in chest wall tumors, 721
in colorectal disorders, 1265
contrast-enhanced digital mammography versus, 571
of hand and wrist
description of, 1933, 1933f
in tumors, 1953
of heart
in cardiomyopathy, 832
description of, 806
in inguinal hernia, 1607–1608
of liver
description of, 1359–1360, 1399
in incidental mass, 1373
in lung cancer
description of, 683
in distant metastasis, 689
in mediastinal tumors, 727–728
in mesenteric ischemia, 931, 931f–932f
in minimally invasive surgery, 455
in neurologic examination, 1829–1830
in pancreatitis, chronic, 1457–1459, 1459f
of parathyroid gland, in preoperative localization studies, 1671, 1672r
in pheochromocytoma, 1694
in renal artery stenosis, 936f, 937, 937f
in soft tissue sarcomas
in initial assessment, 1571, 1571f
in surveillance for recurrence, 1580–1581
in spine trauma, 1845
of spleen, 1540
of stomach
in cancer, 1145
description of, 1115–1116
in thoracic aortic aneurysm, 859
of thyroid, 1634
Malabsorption
in pancreatitis, 1461f, 1461–1462, 1469
in short bowel syndrome, 1255
Malignant fibrous histiocytoma, 535
Malignant hyperthermia, 428, 2040
Mallampati classification, 2034, 2034f
Malleolar fractures
bimalleolar, 1893
lateral, 1893
medial, 1893
posterior, 1893
Mallet finger, 1936
Mallory-Weiss tear, 1083, 1085, 1155
Malnutrition
in elderly, 2051
refeeding syndrome in, 98–99
respiratory quotient in, 420
severe, 2051
wound healing in, 284, 285f
Malocclusion, 2002
Malperfusion syndrome
description of, 876, 880–881
endovascular treatment in, 884
open repair in, 885
Malpractice, 413, 414r
Malrotation, intestinal
description of, 1725–1726
Ladd procedure in, 1726, 1727f
Mammalian target of rapamycin, 40
Mammalian target of rapamycin, 40
Mammagram test in breast cancer, 334, 580
Mammary duct fistulas, 551
Mammary lymph nodes
in breast cancer, 587
description of, 546–547
Mammary milk line, 544f
Mammography
breast cancer applications of diagnosis, 567–571, 568f–570f
risk assessment, 329
screening
description of, 557, 567–571
recommendations on, 570–571
in excisional biopsy with needle localization, 588, 589f
Mammoplasty
augmentation, 2018, 2021f
inferior pedicle reduction, 2020f
Mandible

distraction osteogenesis of, 1995
fractures of, 628, 628f, 2002f, 2002–2003
in head and neck tumors
of alveolar mucosa, 637, 638f
of buccal mucosa, 638
of oropharynx, 639
reconstructive surgery in, 651–652
resection technique in, 637, 638f
of retromolar trigone, 638
reconstructive surgery of fibula osteocutaneous flap in, 651–652
in tumors, 651–652
Mandibulotomy
in alveolar mucosa, 637
in retromolar trigone tumors, 638
Mandibulotomy in alveolar mucosa tumors, 638f
Mannitol, 1832
Mannose-binding lectins, 49
Manoufluorography, 1028
Manometry
anorectal
high pressure zone in, 1266
rectoanal inhibitory reflex in, 1266
resting and squeeze pressures in, 1266
esophageal
high-resolution, 1022, 1023f–1027f
normal values in, 1022, 1022f, 1032f
in nutcracker esophagus, 1057
preoperative
in antireflux surgery, 1038
in long esophageal myotomy, 1060
pressure profile in, 1021f
in primary motility disorders, 1055, 1056f
in scleroderma, 1050, 1050f
in spasm, diffuse and segmental, 1057
transit scintigraphy, 1028
Mapping, lymphatic
of breast, 338f, 338–339
description of, 338f, 338–339
Marfan’s syndrome
aortic abnormalities in aneurysms, 855, 857, 860, 863
description of, 278, 855, 970
genetic factors in, 278, 970
Marginal artery of Drummond, 1260
Marginal ulcers, 1185
Marjolin’s ulcer, 290, 1955
Masaoka staging system in thymoma, 731, 731f
Massachusetts General Hospital, 2027
Massage, lymphatic, in lymphedema, 1003
Mast cells, 56, 516
Mastalgia, cyclical, 552
Mastectomy, 592–593
in advanced local-regional disease, 587
in ductal carcinoma in situ, 580–581
in early invasive disease, 582–583
Halsted radical procedure, 542–543, 591
historical aspects of, 541–543
local-regional recurrence in, 587
modified radical procedure, 542–543, 592–593
auxiliary lymph node dissection in, 592
Patey technique, 591–593, 593f
in pregnancy, 600
nipple-areolar sparing, 591
partial procedure, 591. See also Breast conservation surgery
prophylactic, 556, 558, 560
radiation therapy and, 543, 594
reconstruction after, 583–584
reconstructive surgery after, 593–594
skin-sparing technique, 591, 593
survival rates in, 543
Mastitis
epidermic puerperal, 550
nonepidemic (sporadic) puerperal, 550
periductal
description of, 552–555
recurrent, 550
Mastoiditis, 615
Maxopexy, 2018
Mathes-Nahai classification, of muscular flaps, 1979, 1981f
Matrix metalloproteinases
description of, 313, 315–316
tissue inhibitors of, 315–316
Matrix synthesis in wound healing
description of, 273f, 274f, 274–275
fetal, 282
Maturation phase of wound healing, 273f, 275
MAUDE, 415
Maxillary sinus tumors, 644
Maxillofacial fractures, 2002
Maxillectomy, 638
Maxillary sinus tumors, 644
Mediastinum
anatomy of, 726–727, 727f
compartments of, 727f
contents of, 726–727
disorders associated with, 727, 728f
cysts of, 735
ectopic endocrine glands in, 735
infections of, 735–736
lymphoproliferative disorders of, 727
tumors of, 726–736
biopsy in, 729–730
in children, 727, 728f
diagnostic evaluation in, 729–730
history and physical examination in, 727, 729f
radionuclide scans in, 730f
types of, 727, 728f
Medical education
on minimally invasive surgery, 475
morbidity and mortality conferences in, 2074
resident training, 15f
Medical errors
in communication issues, 401–402
definition of, 397
disclosure of, 413–415, 2074
high-profile, 400–401
Institute of Medicine report on, 397, 2074
legal issues associated with
description of, 413–415, 414f
wrong-site surgery, 411–412
near misses and, 400f
retained surgical items as, 409–410, 410f
risk management in, 413–415
in surgical specimen labeling, 403, 404f
systems change in, 400
types of, 400f
wrong-site surgery in, 411
Medical knowledge, as core competency, 4f
Medical professionalism, online, 2193, 2193f
Medical reversal, 2150
Medical students, leadership modeling for, 18–19
Medication review, in elderly, 2052
MEDLINE, 2138
Medullary carcinoma
of breast, 565–566
of thyroid
description of, 1655–1656
in multiple endocrine neoplasia, 1677
Medulloblastoma, 1857
in fasting, 60–62, 61t
extended, 61, 61f
short-term, 60f
in injury, 62f, 63–64
iodine, thyroid function in, 1629
liver functions in, 1351
in shock, 136–137
in surgical patients
description of, 60–66
energy requirements for, 66–68, 68t
Metacarpal bones, 1926, 1928f
Metacarpophalangeal joint
anatomy of, 1926, 1928f
degenerative disorders of, 1945
fractures of, 1934
rheumatoid arthritis of, 1946
degenerative disorders of, 1945
anatomy of, 1926, 1928f

Metastasis
to adrenal gland, 1695
to axillary lymph nodes
in breast cancer. See Axillary lymph nodes, in breast cancer in unknown primary cancer, 600
to bone
description of, 490
in lung cancer, 683
to brain
description of, 1854–1855, 1855f
stereotactic radiosurgery in, 1872
of breast cancer. See Breast cancer, metastasis of
collorectal cancer. See Colorectal cancer, metastasis of
description of, 316–317, 317f, 490, 490f
distant, surgical management of, 339–340
descriptor, 317–318
of endometrial cancer, 1814t
of esophageal cancer
description of, 1069
staging of, 1069–1070, 1071t
of gallbladder cancer, 1422
doctrine of, 1482
of gastronoma, 1482
of gastrointestinal stromal tumors, 1585–1586
genetic factors in, 317
of hand tumors, 1958
of head and neck tumors
to cervical lymph nodes, 646f,
646–648, 648f
anatomy and levels of, 645f,
646–647
dissection techniques in, 646f,
647–648, 648f
skip metastasis in, 647
in hypopharyngeal cancer, 640
in lip cancer, 634
in oropharyngeal cancer, 639
in salivary gland tumors, 650
staging of, 632, 633–634t
to heart, 843–844
to liver, 1378–1379

of colorectal cancer, 1295, 1299,
1378–1379
hepatectomy in, 1386
of lung cancer
description of, 683, 686–687
staging of, 689–691, 690t–691t
to lungs, 678–680
lymphatic mapping in, 338f, 338–339
of melanoma, 530–531
of mesothelioma, 744
micrometastasis in, 337
of pancreatic cancer
to liver, 1488f, 1498–1499
staging of, 1485, 1487
of parathyroid carcinoma, 1677
seed and seed theory on, 317
sentinel lymph nodes in, 338–339,
See also Sentinel lymph nodes
to small intestine
description of, 1242
treatment for, 1245
of soft tissue sarcomas
description of, 1573, 1573f, 1576
surveillance for, 1580
to spine, 1860
staging of, 333–334
of stomach cancer, 1145t
of thyroid cancer
description of, 1649t
neck dissection in, 1661, 1662f
to thyroid gland, 1658
of vaginal cancer, 1811, 1811t
of vulvar cancer, 1810, 1810t
Metastasis-suppressor genes, 317
Metatarsal fractures, 1894–1895
Metatarsal-phalangeal joint fractures, 1895
Metencephalon, 1828
Methacholine, 1050
Methadone, 2070t
Methimazole, 1636
Methotrexate
in Crohn’s disease, 1238, 1281
in ulcerative colitis, 1281
Methylprednisolone
in Crohn’s disease, 1238, 1281
in ulcerative colitis, 1281
Methylprednisolone in spinal cord injuries, 219, 1847
in traumatic brain injuries, 1839
Meclopramide, 1115t
Metronidazole, 167t, 287t
in amebic liver abscess, 1370
in pelvic inflammatory disease, 1803
in peptic ulcer disease, 1127t
MEWS. See Modified Early Warning System
Micafungin, 163
Mice models in genetic research, 500–504
Millennium development goals in global surgery, 2088, 2088t
Milk-alkali syndrome, 1669
Milan criteria, for liver transplantation in hepatic cellular carcinoma, 381, 1380
Milk-Line enterostomy, 1728f
Milklineٍ, 149
Mineral(s)
dietary requirements for, 68
in parenteral nutrition, 73
small intestine absorption of, 1225
in wound healing, 286
Mineralocorticoids, 38f, 38, 1682
Minimal residual disease, 337
Minimally invasive adenocarcinoma, 671
Minimally invasive surgery, 453–476
access for, 459–460
extraperitoneal, 460–461
subcutaneous, 460–461, 461f
anesthesia in, 457
balloon inflation technique for, 457, 457f
in cancer, 473–474
in children, 472
in cirrhosis and portal hypertension, 474
description of, 2117
economics of, 474–475
education and training on, 475
in elderly, 474
ergy sources in, 465–467, 466f
extracavitary, 457, 457f
hand-assisted
access for, 461, 461f
description of, 1270
historical aspects of, 454–455
imaging systems in, 463–465, 464f
innovations in, 475–476
instrumentation in, 467
in laparoscopy. See Laparoscopy
in natural orifice transluminal endoscopic surgery. See Natural orifice transluminal endoscopic surgery
physiology and pathophysiology in, 455–472
port placement in, 462, 463f
positioning of patients in, 458, 459f
in pregnancy, 472–473
surgical suite for, 458, 458f
team approach in, 457–458
Minimally invasive surgery (Cont.):
telescopes in, 462, 464f
in thoracoscopy
description of, 459
in esophageal atresia and tracheoesophageal fistula, 1720
Minimum alveolar concentration, 2029
Minocycline, 167
Mirizzi’s syndrome, 1405, 1418
Mismatch repair genes, 324–325
Misoprostol, 1124
Mithramycin, 1680
Mitochondrial damage-associated molecular pattern molecules, 30
Mitochondrial DNA, 30
Mitochondrial transcription factor A, 30
Mitogen-activated protein kinases (MAPK)
Mitochondrial DNA, 30
Mitochondrial damage-associated molecular pattern molecules, 30
Mitochondrial transcription factor A, 30
Mitogen-activated protein kinases (MAPK) description of, 313
in insulin pathway, 488, 489f
in thyroid cancer, 1646
Mitomycin, 700
Mitosis, 481
Mitotane, 1692
Mitral valve disorders, 817–823
ACC/AHA guidelines, 820t
annuloplasty in, 822, 823f
atresia, 773
balloon valvuloplasty in, 821
commisurotomy in, 819, 821–822
in hypoplastic left-heart syndrome, 773–777
insufficiency and regurgitation atrial fibrillation in, 819
diagnostic studies in, 821
functional, 832–833
indications for surgery in, 821
ischemic cardiomyopathy in, 831–832
jet lesion in, 822
repair procedures in, 817
types of, 819
prolapse, 822–823
prosthetic replacement in description of, 816
operative technique in, 821
in stenosis, 818f
regurgitation, 819–821
repair procedures in anterior leaflet, 822
double-orifice, 822
type 2, 1695
type 2A
description of, 325, 1655, 1655t
Brunicardi_Index_Vol II.indd   60 22/02/19   11:26 AM
in systemic inflammatory response syndrome, 426, 426f
Morton, William, 2027
Mochowicz’s operation, 1306
Motilin, 1227
Motility antroduodenal, 1117
of esophagus. See Esophagus, motility of
of large intestine, 1263–1264
of small intestine. See Small intestine, motility of
of stomach. See Stomach, motility and emptying of
Motor function assessment
Glasgow Coma Scale in, 1829, 1830t
response patterns in, 1830f
scoring system in, 1829t
Motor vehicle accidents
description of, 2083–2084, 2085f
mechanisms and patterns of injury in, 1830f
pediatric trauma in, 244
in pregnancy, 240
seat belt injuries in, 244, 1753, 1753f, 1902
“Mount Fuji” sign, 1839, 1839f
Mourning process, 2069
Mouse models in genetic research chimeric, 503
description of, 500–504, 501f
gene knockout in, 501–503, 502f
transgenic mice in, 500–501
Mouth cancer. See Oral cavity cancer
Moxifloxacin, 287
MPL gene in familial essential thrombocytopenia, 322t
mRNA degradation control, 485
MSH2 gene, 322
MSH6 gene, 322
MUC1, 336
Mucinous carcinoma, 566
Mucocoele of appendix, 1339–1340
Mucopidermoid carcinoma, 651, 672
Mucormycosis, 714–715
Mucosa oral cavity
alveolar, 637, 638f
buccal, 638
small intestine, 1220, 1221f
stomach barrier function of, 1109, 1109f
description of, 1102, 1104f
Mucosa-associated lymphoid tissue, 1149
Mucous cyst, 1954
Müllerian inhibiting substance, 1746
Multicentricity of breast cancer, 564, 564f
Multidisciplinary approach to cancer therapy description of, 337
in head and neck tumors, 630
to colorectal surgery, 1279
Multifactor Leadership Questionnaire, 13–14, 19, 20
Multifocality of breast cancer, 564
Multimodal analgesia, 2119–2121
Multiple endocrine neoplasia type 1, 322
gastric carcinoid tumors in, 1151
gastrinoma in, 1136–1137, 1482
hyperparathyroidism in, 1666, 1674, 1677
pancreatic neoplasms in, 1480, 1482
type 2, 1695
type 2A
description of, 325, 1655, 1655t
hyperparathyroidism in, 1666, 1674, 1677
medullary thyroid carcinoma in, 1677
type 2B
clinical features in, 1655, 1655f
description of, 325, 1655, 1655r
Multiple organ dysfunction syndrome complement activation in, 140
neutrophils in, 140
postoperative, 426
in traumatic shock, 145
Multivisceral transplantation, 385–388
Mupirocin in burns, 257
Murmurs
in aortic insufficiency, 827
in aortic stenosis, 825
in atrial septal defect, 753
in ductus arteriosus patency, 760
in mitral insufficiency, 821
in mitral stenosis, 819
in tricuspid valve disorders, 830
in truncus arteriosus, 764
in valvular heart disease, 815r
Murphy’s sign, 1406
Murray, Joseph E., 1968, 1969f
Musculocutaneous flaps, 1977, 1979
Muscularis propria, 1221, 1226
Muscle flaps, 1977f
Murray, Joseph E., 1968, 1969
Musculoskeletal disorders
obesity and, 1200
orthopedic surgery in, 1880–1919
postoperative, 423
Mustard procedure, 780
Mutations, 318, 481
Mutator phenotypes in cancer, 319
*MUTYH* gene in adenomatous polyposis coli, 322r
Myasthenia gravis
description of, 1867
thymoma in, 731
Mycobacterial infections
in bronchiectasis, 709
description of, 162, 710–711, 712f
treatment of
algorithm for, 711, 712f
surgical, 711, 712r
tuberculosis in. See *Tuberculosis*
Mycophenolate mofetil, 359f, 361r
adverse effects of, 360f, 361
mechanism of action, 361
Mycotic infections, 711–715. See also
Fungal infections
MyD88-dependent pathway, 32
Myelencephalon, 1828
Myelofibrosis, splenectomy in, 1523r, 1529
Myeloid differentiation factor 88, 1355
Myeloid metaplasia, agnogenic,
splenectomy in, 1523r, 1529
Myeloid-derived suppressor cells, 55
Myeloma
multiple, 1914
epidemiology of, 307r, 310t
stem cells in, 318
solitary, 1914
Myelomeningocele
fetal surgery in, 1754
and spina bifida, 1872
Myelopathy, cervical spondylotic,
1863–1864, 1864f
Myeloproliferative disorders, 110
platelet dysfunction in, 110
splenectomy in, 1528–1529
*MYH*
in colorectal cancer, 1290
in familial adenomatous polyposis, 1292
Myocardium
infarction of
cardiogenic shock in, 148–150
in coronary artery disease, 809
in elderly, 2055
left ventricular aneurysm in, 833–834, 834f
left ventricular remodeling in, 831, 833
peripartum, 406–407
postoperative, 421
risk of noncardiac surgery in, 804
transfusion triggers in, 117
in trauma, 193
ischemia of
angina pectoris in, 801–802, 809
in coronary artery disease, 808–814
diagnostic studies in, 804
laser revascularization in, 814
postoperative, 421
radionuclide scans in, 806
silent, 802
transfusion triggers in, 117
protection in cardiopulmonary bypass, 808
viability assessment, 832
Mycobacterial flaps
in breast reconstruction, 593–594
in chest wall reconstruction, 726, 726f
in head and neck reconstruction, 651
Myofibroblasts, 1974
Myoglobinuria, 423
Myomectomy, 1798f
in hysterectomy, 1797
laparoscopic, 1797
Myometrium, 1785
Myopectineal orifice of Fruchaud, 1600, 1603f
Myotony, esophageal
cricopharyngeal, 1054–1055
failure of, 1067r
Heller, 1063–1065, 1064f
laparoscopic, 1065f, 1066f–1067f
long, 1060f, 1062, 1061f–1062f
open technique, 1065
outcome assessment in, 1067r, 1068
peroral, 472f, 1065
Myringotomy, 615, 615f, 622
Myxomas, cardiac, 842–843, 843f
N
NADPH oxidases, 39
Nafcilin, 165r
NAFLD. See Nonalcoholic fatty liver
disease
Nails, 516
paronychia of, 1951–1952, 1952f
subungal melanoma of, 1956
Naloxone, 2031, 2040
Naproxen
gastrointestinal disorders from, 1124r
in persistent pain, 2070r
Nasal disorders. See *Nose*
NASH. See Nonalcoholic steatohepatitis
Nasolabial molding, 1984
Nasenteric feeding tubes, 71, 71f
Nasogastric intubation
for enteral nutrition, 71, 71f
in small intestine obstruction, 1231
tracheoesophageal fistula in, 665
in trauma, 195
tubes for, 2121
Nasomaxillary buttress, 2002
Nasoorbitalethmoid fractures, 2005
Nasopharynx
anatomy of, 630–631, 640f, 644
cancer of, 632, 644–645
Nasotracheal intubation, 184
National Cancer Institute
description of, 555
soft tissue sarcoma classification
system, 1573
National Comprehensive Cancer Network,
1580
National Lung Screening Trial (NLST),
675–676
National Quality Forum
description of, 408–409
on “never” events in surgery, 409, 409f
National Surgical Adjuvant Breast and
Bowel Project trials, 334,
542–543, 595–596
on breast cancer prevention, 557–558
on sentinel node surgery in breast
cancer, 590
National Surgical Quality Improvement
Program/Project, 17r, 407–408,
2036, 2047, 2049
National Wilms Tumor Study Group,
1747–1748
Natural orifice transluminal endoscopic
surgery
access for, 459, 461–462
in appendicitis, 1336
cholecystectomy in, 471f
description of, 453, 462f, 470
historical aspects of, 454
instrumentation in, 467
team approach in, 458
Nausea, postoperative
prevention of, 2121, 2121f, 2127
traditional Chinese medicine for,
2127
Navicular fracture, 1894
NBS1 gene in Nijmegen breakage
syndrome, 322r
Nd:YAG (neodymium yttrium-aluminum
garnet) lasers, 466
Neck. See also Cervical spine; Head and
neck disorders
cervical lymph nodes in. See *Cervical
lymph nodes
masses in, 646–650
benign, 649
in cervical lymph node disorders,
646–648, 1710
in children, 1710–1712
in deep fascial planes, 649–650
diagnostic evaluation of, 646
in parapharyngeal space, 648–649
trauma of, 219–222
airway management in, 183–185,
185f
blunt, 197
evaluation of, 197–199
laryngeal fracture in, 198f
Neck, trauma of (Cont.):
- operative approaches and exposure in, 209f, 209–210
- penetrating, 198, 198f–199f
- vascular, 220f–221f, 220–222, 222f
- zones of injury of, 198, 199f

Necrosis, 40–41
- in colostomy, 1278
- differentiated from apoptosis, 314
- in ileostomy, 1277
- of pancreas
  - in acute pancreatitis, 1443, 1447f
  - description of, 172–173, 173f
  - renal tubular, 422–423
- in soft tissue infections, 173, 175f, 524–526

Necrotizing enterocolitis, 1727–1730
Necrotizing pancreatitis, 1442, 1443f

Needle
- in endovascular therapy, 904
- Veress, 459f, 459–460

Needlestick injuries, 178

Negative pressure wound dressings, 2015

Negligence, 400

Nelson’s syndrome, 1701

NEO Five-Factor Personality Inventory, 19f, 20

Neodymium yttrium-aluminum garnet (Nd:YAG) lasers, 466

Neoplasm
- in living donor kidney transplantation, 1707–1708
- in kidney cancer
  - description of, 1773, 1773f
  - laparoscopic technique, 1773
  - in living donor kidney transplantation, 367f, 369

Nephrocalcinosis in hyperparathyroidism, 1667

Nepholithiasis, 1667

Nephropathy from contrast agents, 417, 859

Nerve(s)
- of anorectum, 1263
- of breast, 546
- of colon, 1261
- of esophagus, 1014, 1014f
- of gallbladder, 1393
- of hand and wrist, 1930–1931

Nerve conduction studies, 1831, 1848

Neurogenic tumors, mediastinal
- Neurofibrosarcoma, mediastinal, 733
- Neurofibroma
- Neuroendocrine tumors
  - Neuroendocrine hyperplasia of lung
description of, 672
- Neuroendocrine carcinoma, 672–673, 693
- Neurofibromatosis, 733, 1570
  - type 1, 322f, 1570, 1861, 1867
  - type 2, 322f
- Neurofibromatosis-1, 1995
- Neurofibrosarcoma, mediastinal, 733
- Neurogenic claudication, 1864–1865, 1903
- description of, 131, 151f, 151–152
- hemodynamic response to, 135t
  - in trauma, 151, 193
- Neurogenic tumors, mediastinal
  - classification of, 733f
  - description of, 732, 733f
  - incidence of, 727, 728t

Neurologic disorders
- autoimmune and inflammatory, 1867
- emergencies in, 1831–1833
- in entrapment neuropathies
  - of median nerve, 1867, 1943–1945
  - of ulnar nerve, 1867, 1944
- examination in, 1829–1831. See also Neurologic examination
- fluid and electrolyte balance in, 98
- lower extremity ulcers in, compared to ischemic ulcers, 958f, 959f
- in lung cancer
description of, 680–681
- in paraneoplastic syndromes, 682
- monitoring in, 447–449
- postoperative, 418, 428
- surgery in, 1827–1874. See also Neurosurgery

Neurosurgery
- traumatic
description of, 281, 1847–1848
- in hand and wrist, 1936, 1943

Neurologic examination, 1829–1831
- in brain death, 1842
- diagnostic studies in, 1831
- motor tests in, 1829, 1829f, 1830f, 1830r
- sensory tests in, 1829
- in trauma
description of, 190, 192t
  - of spine, 1843–1844

Neurolytic therapy, in chronic pancreatitis, 1469

Neuroma
- acoustic, 1857, 1858f
description of, 528
- in hand, 1943

Neuromuscular blockers
- anesthetic uses of, 2031–2032, 2032r
- depolarizing, 2031–2032, 2032r
- nondepolarizing, 2031–2032, 2032r
- peripheral nerve stimulators for monitoring of, 2033

Neuronal nitric oxide synthase, 1356

Neuropathy
- entrapment
  - of median nerve, 1867, 1943–1945
  - of ulnar nerve, 1867, 1944
  - lower extremity ulcers in, compared to ischemic ulcers, 958f, 959f
  - traumatic, of common peroneal nerve, 1848–1849

Neuropsychiatric disorders in hyperparathyroidism, 1668

Neurosurgery, 1827–1874
- anatomy in, 1854
- in brain abscess, 1868
- in brain stem compression, 1833
- in cerebrovascular disorders, 1833, 1849–1853
- in congenital and developmental anomalies, 1872–1874
- in cranial infections, 1867–1868
- in head trauma, 1833–1842
- in intracranial pressure increase, 1831–1832
- in intracranial tumors, 1854–1859
- neurologic examination in, 1829–1831
  - of peripheral nerves
description of, 1866–1867
  - in trauma, 1847–1848
  - in seizures, 1833, 1869–1871
  - of spine, 1861–1866
  - in infections, 1867–1868
  - in trauma, 1842–1847
  - in tumors, 1859–1861
  - stereotactic radiosurgery techniques in, 1871–1872
- in subdural empyema, 1868
- in trauma, 1833–1849
- in trigeminal neuralgia, 1871

Neurotransin, 1227

Neurotension, 281, 1848

Neurotransmitters, 50
Index

Neutropenia, 1323–1324
Neutrophil(s)
  bone marrow mobilization of, 54
  endothelium interactions and, 56–57, 58f, 140
  functions of, 56
  in inflammatory response, 54
  in microbial invasion response, 160
  monocyte recruitment by, 54
  polymorphonuclear, 140, 160
  reduced number of, 1323
  in shock response, 140
  in wound healing, 272, 273f–274f
  Neutrophil extracellular traps, 54, 56
  “Never events” in surgery, 409t, 409–412
Nevi
  congenital, 527
  giant, 527
  dysplastic, 530
  melanoma risk in, 530
  melanocytic, 527
Nevoid basal cell carcinoma, 322r
New York Heart Association classification of heart disease, 803, 803r
New York State Study Group on coronary stenting and bypass surgery, 810
Newborn. See Neonates
Next-generation sequencing, 497–499, 506
NF1, 322r
NF2, 322r
NHEJ-mediated gene editing, 506–507
Nicoll, James, 2154, 2155
Nicotine, 2116
Night stick fractures, 1886
Nilotinib, 343r
Nipple
  accessory, 543
  ampulla of, 545
  anatomy of, 544–545
  discharge from, 553
  bilateral, 600
  bloody, 552
  unilateral, 599–600
  hidradenitis suppurativa of, 550
  inversion of
description of, 551
  treatment for, 555
  Paget’s disease of, 550, 565
  ptosis of, 2011f
  reconstruction of, 2011f
Nissen, Rudolph, 1040
Nissen fundoplication
  laparoscopic, 1040, 1040f–1042f
  outcome of, 1042–1043
  posterior partial fundoplication versus, 1044
  Toupet fundoplication versus, 1044
Nitrates, dietary, 1140
Nitric oxide
  description of, 58, 59f, 1355–1356
  in erection physiology, 1776
  in gastric relaxation, 1110
  in liver, 1355–1356
Nitric oxide synthases
description of, 1356
in septic shock, 146
Nitroblue tetrazolium reduction test, 289
Nitrogen
  dietary intake of, postoperative, 66
  urine levels of, in injury, 66
  Nitroglycerin, 237r
  Nitroprusside, 882
  Nitrous oxide
    advantages and disadvantages of, 2030r
    anesthesia uses of, 2027, 2029–2030, 2030r
    history of, 2027
  NLRP3 gene, 33
  N-methyl-d-aspartate receptor, 2031
_N-myc, 1749
Nocardiocysis, 708
Nodularity of breast, 552
Nodules
  mural, in hemangioblastoma, 1858
  solitary, of lungs, 676–678. See also Lung(s), solitary nodule in thyroid
  hyperthyroidism in, 1637
  solitary, 1641, 1643–1645
Nonalcoholic fatty liver disease, 1200, 1363, 1541
Nonalcoholic steatohepatitis, 1175, 1363, 1541
Nondepolarizing neuromuscular blockers, 2120
Noninferiority trials, 2147
Non-Hodgkin’s lymphoma, 1528
Noninferiority trials, 2147
Noninsulinoma hyperinsulinemia, 1481
Nonmaleficence principle
description of, 2061
0escription of
  in withdrawing or withholding of life-sustaining therapies, 2065
Nonopioid analgesics, 2031
Nonseminomatous germ cell tumors, 734
Non–small cell lung carcinoma, 670–672, 693, 698–699, 701
Nontumoral anti-inflammatory drugs, 2031, 2120r
Nonunion, of hand fractures, 1943
No-reflow phenomenon, 136
Norepinephrine
  description of, 177, 1684–1685, 1686f
  in pheochromocytoma, 1694
  in shock response, 137
  Normal accident theory, 399
  Normovolemia, 2118–2119
  North American Symptomatic Carotid Endarterectomy Trial, 911
  Northern blot hybridization, 494–495
  Nitritopyline in pain, 2071r
  Norwood procedure, 775–777
  Nose
    forehead flap for reconstruction of, 1980f
    fractures of, 219
    polyps of, 618, 619f, 620
    postoperative complications of, 418
    tumors of, 643–644
  Nosocomial infections, postoperative
    respiratory, 174–175, 419
    of surgical site, 174–175, 425
    of urinary tract, 425
  NOTECHS, 2177
  NOTES. See Natural orifice transluminal endoscopic surgery
  NOTSS, 2177
  NSQIP. see National Surgical Quality Improvement Program
  Nuclear factor-kB
  in acute pancreatitis, 1443
  inhibitor of, 53
  target genes, 32
  Nucleic acid amplification testing, 1789
  Nucleotide-binding oligomerization domain-like receptor family, 33
  Nucleus pulposus, 1861, 1903
  Number needed to harm, 2146
  Number needed to treat, 2146
  Nutracker esophagus, 1057
  Nutrition, 66–74
  age-related differences in requirements for, 1708r
  arginine in, 70
  in burn patients, 257–258
  in children, 1708r, 1708–1709
  enteral, 1708r, 1708–1709
  parenteral, 1709
  in chylous ascites, 1253
  colorectal cancer risk factors in, 1289
  in elderly, 2051
  energy requirements in, 66–68
  enteral, 68–72
  access for, 71r, 71–72
  in burns, 257–258
  in children, 1708r, 1708–1709
  complications of, 72
  early initiation of, 68–69
  formulas used in, 69–71
  in gastrointestinal trauma, 228
  indications for, 68, 426–427
  in pancreatitis, 1446
  parenteral nutrition versus, 68
  in short bowel syndrome, 1254
  in esophageal cancer, preoperative, 1073
  gastric cancer risk factors in, 1139
  glutamine in, 69–70
  in inflammatory bowel disease, 1282
  inflammatory response affected by, 69–70
  in intestinal fistulas, 1241
  omega-3 fatty acids in, 48, 70
  overfeeding in
description of, 68
  in parenteral nutrition, 74
  in pancreatitis
    acute, 1446
    chronic, 1455
  parenteral, 72–74
  access for, 74
  central or total, 73
  in children, 1709
  complications of, 73–74, 427
  contraindications to, 73
  electrolytes in, 93, 94t
  enteral nutrition versus, 68
  indications for, 73
  initiation of, 73–74
  in pancreatitis, 1446
  peripheral, 73
  in short bowel syndrome, 1254–1255
  postoperative, 420, 426–427
  preoperative
evaluation of, 2115
  in traditional Chinese medicine, 2126
  refeeding syndrome and, 98–99
  in ventilatory support, 420
  vitamin and mineral requirements in, 68
  wound healing affected by, 284, 285f
  Nyhus classification of inguinal hernia, 1600, 1602r

See also 
Urinary tract infections, 425
Urinary tract infection, 425
Obesity

asthma and, 1200

bariatric surgery for. See Bariatric surgery

behavioral therapy for, 1171

body mass index in, 1169–1170

breast cancer risk in, 555

cardiovascular disease in, 1199

causes of, 1170

in children, 1170

classification of, 1169

comorbid conditions with, 1199–1200

in Cushing’s syndrome, 1689

definition of, 1169

endotracheal intubation in, 2034

environmental factors, 1170

gastroesophageal reflux disease and, 1199–1200

genetic factors, 1170

hypoventilation syndrome of, 1177

inguinal hernia in, 1605

lifestyle interventions for, 1171–1172

medical management of, 1171–1173

medical problems associated with, 1170–1171

musculoskeletal disease and, 1200

nonalcoholic fatty liver disease and, 1200

obstructive sleep apnea as risk factor for, 1200, 2035

overview of, 1169–1170

pharmacotherapy for, 1172f, 1172–1173

phentermine for, 1172

prevalence of, 1170

psychiatric problems associated with, 1172–1173

risk factors for, 1171f

social problems associated with, 1170–1171

surgery for, 1167–1168

wound healing problems in, 284

Objective structured assessment of surgical skills, 2164, 2166

Objective structured assessment of technical skills, 15

Objective structured clinical examination, 14–16, 19t

Objectivity, 9

Oblique abdominal muscles external, 1549, 1550f–1551f

internal, 1549, 1551f

Obsessive-compulsive disorder, deep brain stimulation in, 1871

Obstructive shock, 150–151

Obstructive sleep apnea, 623

in bariatric surgery, 1177–1178

obesity as risk factor for, 1200, 2035

Obturator bypass in aortoiliac occlusive disease, 946

Obturator internus muscle, 1784

Obturator sign in, 1332

Occipital dislocation, 1900

Occupational exposure to blood-borne pathogens, 177–178

to chemicals, soft tissue sarcoma risk in, 1568

Octreotide in chylous ascites, 1253

in dumping syndrome, 1157

in gastrinoma scan, 1137

in intestinal fistulas, 1241

in pancreatic cancer, 1468

in portal hypertension and variceal bleeding, 1367

in visceral pain, 2071r

Octyl-cyanoacrylate tissue glues, 295

Oddi sphincter

anatomy of, 1395–1396, 1396f

cholecystokinin affecting, 1397

functions of, 1397

incompetent, acute pancreatitis in, 1440

sphincteroplasty in chronic pancreatitis, 1472–1473, 1473f

sphincterotomy in choledochothiasis, transduodenal, 1414

stenosis of, 1415

Odontoid process fractures, 1843, 1900–1901

Office-based surgical suites, 2160

Off-pump coronary artery bypass surgery, 812f, 812–813

Olfaxin, 287r

Ogilvie’s syndrome, 1308–1309

Ongren’s line, 642f, 644

Olaratumab, 1579

Older adults. See Elderly

Olecranon fractures, 1886

Olfactory groove meningeoma, 621f

Oligodendroglioma, 1856

Oligomenorrhea, 1795

Oligospermia, 1793

Omentum anatomy in, 637–638

in floor of mouth

anatomy in, 639

reconstructive surgery in, 637, 637f

in palate, 639

reconstructive surgery in, 637, 637f

in palate, 638

reconstructive surgery for, 638

in retromolar trigone, 639–639f

staging of, 639f–639f

in tongue, 634–636, 635f, 639

Organ culture. See Organ culture

Orchidectomy in testicular cancer, 1772

in testicular torsion, 1769

Orchitis description of, 1761–1762

ischemic, in inguinal hernia repair, 1620

Organ failure, in acute pancreatitis, 1448–1449

Organizational culture. See Culture, organizational

Oropharyngeal cancer, 638–640

Oral cavity cancer, 634–640

in alveolus/gingiva, 637

anatomy in, 630, 630f, 634, 635f–636f

in buccal mucosa, 638

epidemiology of, 308f, 310t

t etiology of, 630

in floor of mouth

anatomy in, 636f

reconstructive surgery in, 637, 637f

in palate, 638

reconstructive surgery for, 638

in retromolar trigone, 637–638

staging of, 639f–639f

in tongue, 634–636, 635f, 639

Orbicularis oris, 1900

Orbital apex syndrome, 2004

Orbital fissure syndrome, 2004

Orbital fractures blowout, 628, 2004

t description of, 219, 2004, 2004f

Orchidopexy, 1745

Orchiectomy in testicular cancer, 1772

in testicular torsion, 1769

in testicular torsion, 1769

in testicular torsion, 1769

Orchitis description of, 1761–1762

ischemic, in inguinal hernia repair, 1620

Organ failure, in acute pancreatitis, 1448–1449

Organizational culture. See Culture, organizational

Oropharyngeal cancer, 638–640

anatomy in, 631, 638–639

metastasis of, 639

rehabilitation in, 639–640

Oropharyngeal dysphagia, 1031

Orotracheal intubation, 185
Pain (Cont.):
multimodal analgesia for, 2119–2120
palliative care in
adjuvant medications in, 2071t
description of, 2067, 2069
in pancreatic cancer, 1489
in persistent pain, 2070r
step ladder approach to, 2069, 2069t
in pancreatic cancer, 1487, 1489
pancreatitis-related
acute, 1444
chronic
description of, 1460–1461, 1461f/treatment of, 1469, 1470f/pelvic, 1267
postoperative
delirium caused by, 2047
description of, 2039–2040
in inguinal hernia repair, 1619–1620
in thoracic surgery, 701f/scrotal, in testicular torsion, 1769
in shoulder impingement syndromes, 1896
sympathetic mediated, 1943
in thoracic aortic aneurysm, thoracic, 857–858
in abdominal aortic aneurysm, 857–858
in thoracic surgery, 701f, 704–705
in thyroiditis, 1639
traditional Chinese medicine for, 2127
in trigeminal neuralgia, 1871
Pain management
in palliative care, 2053
perioperative, 2119–2121, 2120r, 2127
Pallate
anatomy of, 1984
cleft. See Cleft lip and palate
tumors of, 638
Pallaplasty, 1985–1987
PALB2, 322r, 561
Palifosfamide, 1579
Palliative care, 2066–2068
communication skills in, 2067–2068, 2068r
core competencies in, 2053, 2054r
definition of, 2066
for elderly, 2053
general principles of, 2066–2067
in head and neck tumors, 654
health and healing, 2067
indications for, 2067, 2067t/pain management in
adjuvant medications in, 2071t
description of, 2053, 2067, 2069
in pancreatic cancer, 1489
in persistent pain, 2070r
step ladder approach to, 2069, 2069t
in pancreatic cancer, 1489–1490
pharmacotherapy principles in, 2069, 2069r/radiation therapy, in recurrent soft tissue sarcoma, 1582
surgical, 2066
Palliative surgery
definition of, 2056
for elderly, 2056
Palmar space infections, 1949
Palmaris longus tendon, 1926
Pulpitations, 802
Palsy
brachial plexus, 1915
cerebral, 1915
Pamidronate, 2071t/Pancoast, Henry, 698
Pancoast’s tumor
description of, 680
treatment of, 698, 699f/Pancolitis, 1280/Pancoast, Henry, 698
Palpitations, 802
Palmaris longus tendon, 1926
Palmar space infections, 1949
Palmaris longus tendon, 1926
Pulpitations, 802
Palsy
brachial plexus, 1915
cerebral, 1915
Pamidronate, 2071t/Panocast, Henry, 698
Pancoast’s tumor
description of, 680
treatment of, 698, 699f/Pancolitis, 1280/Pancoast, Henry, 698
Palpitations, 802
Palmaris longus tendon, 1926/Palmar space infections, 1949/Palmaris longus tendon, 1926/Pulpitations, 802
rejection in, 377
results of, 378
simultaneous with kidney
transplantation, 374, 378
survival rates in, 378
Y-graft reconstruction with, 375, 376f
trauma of, 1429
complications in, 230
with duodenal injuries, pyloric
exclusion procedure in, 230, 231f
emergent abdominal exploration in,
212
identification of, 229
Roux-en-Y pancreateicojejunostomy
in, 229–230, 230f
vasoactive intestinal peptide-secreting
tumors of, 1482
Pancreas surgery, traditional Chinese
medicine in, 2128
Pancreatin, 1437f, 1438
Pancrectomy
in chronic pancreatitis, 1472,
1474–1481
in pancreatic cancer, 1493
for pancreatic trauma, 229–230
in pseudocyst of pancreas, 1467
total, with islet cell auto-transplantation,
1480f, 1480–1481
Pancreatic arteries, 1433, 1435
Pancreatic duct
anatomy of, 1432
calculi in, extracorporeal shock wave
lithotripsy for, 1471, 1471f
distortion, 1457
obstruction of, 1457
sphincteroplasty of, 1472–1473, 1473f
stent placement in, 1471
Pancreatic intraepithelial neoplasia
(PanIN), 1485, 1486f
Pancreatic polypeptide
in chronic pancreatitis, 1462, 1463f
description of, 1436, 1439
functions of, 1437f, 1439
Pancreatic secretory trypsin inhibitor
(PSTI), 1436
Pancreatic stone protein (PSP), 1457
Pancreatic thread protein, 1457
Pancreaticoduodenal artery, 1433
Pancreaticoduodenal vein, 1433
Pancreatectomy in duodenal neoplasms, 1244
in pancreatic cancer, 1491–1493,
1492f–1493f
anastomosis techniques in, 1493,
1495–1497
complications of, 1497–1498
outcome and value of, 1498
recurrence, 1499
in pancreatic injuries, 229
traditional Chinese medicine after, 2128
Pancreateicojejunostomy
in chronic pancreatitis, 1473, 1473f
in pancreatic ascites, 1468, 1468f
in pancreatic cancer, 1494f–1496f
in pancreatic trauma, 229–230, 230f
Pancreatitis, 172–173, 1439–1481
acute, 1439–1449
in alcohol use and alcoholism, 1440
algorithm for, 1449f
APACHE-II score in, 1445
assessment of severity, 1444–1445
cholecystectomy for, 1449
colonization theory of, 1442f
complications of, 1445f, 1447–1448,
1450f
computed tomography in, 1443f,
1443–1444, 1446, 1447f
definition of, 1439
diabetes mellitus after, 1449
diagnosis of, 1444
drug-induced, 1441
etiologies of, 1440, 1440f
factors affecting severity of,
1444–1446
fluid therapy in, 1446
inflammatory mass in, 1468, 1468f
inflammatory, 1454
fibrosis in, 1455–1456, 1456f
Berne procedure for, 1478, 1479f
for, 1478, 1479f
burned out, 1460
chronic, 1450–1481
postoperative, 422
in infected pancreas, 1450
in pseudocyst of pancreas, 1467
autoimmune, 1454
antisecretory therapy in, 1469
ascites in, 1467f, 1467–1468
autoimmune, 1545
Beger procedure for, 1475, 1476f,
1479–1480
Berne procedure for, 1478, 1479f
burned out, 1460
calcific, 1452, 1456f
classification of, 1452, 1458f
complications of, 1465–1468
definition of, 1450
factors affecting severity of,
1444–1446
inflammatory, 1454
enzymatic therapy in, 1469, 1469f
epidemiology of, 1450
etiologies of, 1450, 1452f
fibrosis in, 1455f, 1455–1456, 1456f
Frey procedure for, 1475–1476,
1477f, 1478–1480
histology in, 1456, 1456f
inflammatory, 1454
inflammatory mass in, 1468, 1468f
laboratory tests in, 1462–1464, 1463f
lipase deficiency in
description of, 1461f, 1462
replacement therapy for, 1469f
obstructive
description of, 1453–1454
results of surgery in, 1471
pain associated with
description of, 1460–1461, 1461f
treatment of, 1468f, 1469
pancreatic cancer in, 1464–1465,
1465f, 1484
pancreatectoduodenectomy for,
1479–1480
prognosis and natural history in,
1463f, 1464f
pseudocyst in, 1465–1467
radiologic imaging in, 1457–1460,
1459f
resection procedures in, 1474–1481
signs and symptoms in, 1460–1461
sphincteroplasty in, 1472–1473, 1473f
stent placement in, 1471
survival rates in, 1464, 1464f
total pancreatectomy with islet
auto-transplantation for, 1480f,
1480–1481
tropical (nutritional), 1450, 1455
Whipple procedure for, 1475,
1479–1480
in gallstones, Ranson criteria on, 1445f
hereditary
description of, 1436, 1441, 1450–1451
pancreatic cancer in, 1484
in hyperparathyroidism, 1668
interstitial edematous, 1442, 1443f
multiple hit theory on, 1452, 1452f
necrotizing, 1442, 1443f
postoperative, 422
in transplanted pancreas, 377
Panreatostomy, 1472
Paneth cells, 1220–1221
Paneth cell tumor, 1220
Paracetamol, 1042–1043
Paracentesis
in chylous ascites, 1253
in chylothorax, 1253
in chylous pleural effusion, 1254
in gastrointestinal bleeding, 309f
in retroperitoneal bleeding, 309f
of pleural effusion, 1254f
of pneumothorax, 1255
Paraffin, 855
Paragonimiasis, 681–683
Paraneoplastic syndromes, 681
Paranasal sinuses
See Facial nerve,
Facial nerve, 732–733
of, 733–734
Paranasal sinusitis
See also Sinusitis
of, 681–683
Paranasal sinusitis
See also Sinusitis
of, 642f, 643–644
Paraneoplastic syndromes, 681f, 681–683
Parapharyngeal space masses, 648f,
648–649
Paraphimosis, 1770
Parasitic infections
See also Sinusitis
of carotid body, 919
familial, 323f
hereditary, 323f
Paraganglioma
of carotid body, 919
familial, 323f
hereditary, 323f
Paraganglionic tumors, mediastinal, 733f,
734
Paralysis, facial nerve. See Facial nerve,
paralysis of
Paranasal sinuses
fungal ball in, 620f
inflammatory disease of, 617–622.
See also Sinusitis
of, 642f, 643–644
Paraneoplastic syndromes, 681f, 681–683
Parapharyngeal space masses, 648f,
648–649
Paraphimosis, 1770
Parasitic infections
anorectal, 1320
cholangiohepatitis in, 1410
splenic cysts in, 1523, 1530
Index

Parasympathetic nervous system, 1829
of anorectum, 1263
of colon, 1261
of esophagus, 1014
of liver, 1350
of pancreas, 1435
of stomach, 1101, 1110
of thyroid gland, 1629

Parathyroid gland, 1663–1681
adenoa of, 728–729, 1663
hyperparathyroidism in, 1671, 1675
parathyroidecmy in, 1674–1676
anatomy and histology of, 1629, 1630f,
1633f, 1664
autotransplantation of, 1680
and calcium homeostasis, 1663–1665,
1665f–1666f

carcinoma of, hyperparathyroidism in,
1666, 1677
ectopic, 735, 1663, 1675, 1676f
persistent hyperparathyroidism in,
1678, 1678f
embryonic development of, 1663,
1663–1664
historical descriptions of, 1663
hyperparathyroidism, 1665–1681.

See also Hyperparathyroidism
hyperplasia of, 1675
hypoparathyroidism, 1681
physiology of, 1664–1665, 1665f–1666f
surgery of, 419

Parathyroid hormone
in calcium homeostasis, 1664,
165f–1666f
ectopic secretion in lung cancer, 681
in hyperparathyroidism
description of, 1669f, 1669t,
1669–1670
intraoperative quick assay of, 1674f
Parathyroidectomy, 1670–1676, 1675f
autotransplantation of tissue in, 1676,
1680
complications of, 419, 1680–1681
endoscopic, 1674
in familial hyperparathyroidism, 1677
historical aspects of, 1663
indications for, 1671, 1671t
minimally invasive, 1674
in neonatal hyperparathyroidism, 1677
operative procedure in, 1672–1674, 1675f
intraoperative identification
of parathyroid glands in,
1674–1676
standard bilateral exploration in,
1674–1676
postoperative care and follow-up in,
1677
preoperative localization tests in,
1671–1672, 1672r
radio-guided, 1674
in secondary hyperparathyroidism,
1679–1680
in tertiary hyperparathyroidism, 1680
Parathyromatosis, 1677–1678
Paratracheal lymph nodes
description of, 667
in lung cancer, 688, 688f
Parenteral nutrition, 72–74, 2115.

See also Nutrition, parenteral
Parietal cell vagotomy in peptic ulcer
disease, 1127f

Parietal cells of stomach
acid secretion of, 1106–1107, 1107f
description of, 1099, 1103, 1105,
1105f–1106f, 1106r
intrinsic factor secretion of, 1108

Parietal lobe, 1827
Parietal pleura, 736
Parkinson, Cyril Northcote, 10
Parkinson’s disease, 1870
Parkinson’s law, 10
Parkland formula, 254
Parks hemorrhoidectomy, 1311
Parona’s space, 1950
Paronychia, 1951–1952, 1952f
Parosteal osteosarcoma, 1912
Parotid gland tumors, 649f, 650
Parry-Romberg syndrome, 1901, 1995f

PARS. See Patient Advocacy Reporting
System
Partial agonist, 2029
Partial palatoplasty, 638
Passaro’s triangle, 1482, 1482f
Passive transport, 1222
Pastor, Louis, 157, 271
Patella
dislocation of, 1890–1891
fractures of, 1891
Patey modified radical mastectomy,
591–593, 593f
Pathogen-associated molecular patterns, 138
Patient Advocacy Reporting System, 16t
Patient care
skill in, as core competency, 4r
training on, 4r
Patient controlled analgesia, 2119
Patient education, in preoperative
evaluation, 2114–2115
Patient Health Questionnaire-2, 2050
Patient Safety Indicators, 406, 406r
Patient Self Determination Act (1991), 2064
Patient-centered care, 11
Patient-specific virtual reality simulations,
2174
Pattern recognition molecules, 34
Pattern recognition receptors
damage-associated molecular pattern
molecules as ligands for, 32–33
definition of, 32
heat shock protein binding to, 30
in shock, 132, 137, 139f
traumatic, 145
Pauchet procedure, 1137f
Pavlk harness, 1918
Pazopanib, 343f, 1579
PECAM-1. See Platelet-endothelial cell
adhesion molecule-1
Pectinate line, 1261
Pectineal ligament, 1600
Pectoralis muscles
flap, in head and neck reconstruction,
640, 651
preservation of, in modified radical
mastectomy, 592f
Pediatric disorders, 1705–1745.
See also Children and infants
Pedicle flaps, 1977
Pelvic examination, 1788, 1789f
Pelvic exenteration
in cervical cancer recurrence, 1813,
1814f
in rectal cancer recurrence, 1300

Pelvic floor
anatomy of
muscles in, 1262, 1784, 1784f
nerves in, 1786f, 1787
disorders of
prolapse of pelvic organs in,
1807–1808
stress urinary incontinence in,
1808–1809
in fecal incontinence, 1264
function testing of, 1266
Pelvic fractures
bladder trauma in, 1766
description of, 188, 204f, 204–205, 206f,
234–235, 188–1888
hemorrhage in, 184, 204,
234–235
and hemodynamic instability, 234,
235f
pelvic packing in, 234, 236f
open wound in, 234–235
in pregnancy, 242
Pelvic inflammatory disease
description of, 1803–1804
differential diagnosis of, 1335, 1803
of, 1803f, 1804
Pelvic organ prolapse, 1807–1808
Pelvic pain, 1267
Pelvic reconstruction, 2012
Pemberton’s sign, 1641, 1642f
Pendred’s syndrome, 1638
Penetrating injuries
abdominal, 200–204
diagnostic peritoneal lavage in, 202,
202f
in pregnancy, 242
chest, 199–200, 223
in children, 1751
description of, 196
hemorrhage and persistent hypotension
in, 193
removal of weapon in, 194, 194f
Penicillin
allergic reactions to, 168
description of, 165f, 287f
in lymphedema, 1003
in soft tissue infections, 174
Penile implant, 1776, 1776f
Penis
anatomy of, 1760
cancer of, 307f
curvature of, 1776–1777
erection of
description of, 1760
in priapism, 1769–1770
prostatectomy affecting, 1775
paraphimosis of, 1770
trauma of
fracture in, 1767
priapism in, 1769–1770
Pentaxins, 34
Pepsinogen, 1105, 1108
Peptic ulcer disease, 1121–1137
anti-inflammatory drugs as cause of,
1121–1123
hospitalization rates in, 1124f
medical treatment in, 1125
perforation in, 1134
risk factors in, 1123, 1124f
Blatchford score in, 1125, 1126f
clinical manifestations in, 1123
complications of, 1124–1125, 1131–1133
diagnosis of, 1123
drug therapy in, 1125
etiologies of, 1118f, 1122f
gastric acid secretion in, 1122, 1122f
gastric cancer versus, 1145
gastric outlet obstruction in, 1125
Helicobacter pylori infections, 1141f
diagnosis of, 1123–1124, 1124f, 1125
pathophysiology in, 1117–1121, 1118f–1119f, 1121f–1122f
treatment of, 1124f, 1125, 1127f
hemorrhage in description of, 1124–1125, 1131–1133
risk stratification in, 1125, 1126f
treatment algorithms on, 1135f
intractable or nonhealing, 1136, 1136f
Johnson classification of, 1122f, 1123f
obsilation in, 1125, 1136
pathophysiology in, 1121–1123
perforation in
description of, 1125, 1134–1136
surgical treatment of, 1134, 1136, 1137f
treatment algorithm on, 1135f
perforation in
in smoking, 1123
in stress, 1123
surgical treatment in, 1127–1131
antrectomy in, 1127f, 1129
in gastric outlet obstruction, 1136
gastric remnant cancer in, 1143
in hemorrhage, 1131–1133
highly selective or parietal cell
vagotomy in, 1127f, 1128f
in intractable or nonhealing ulcer,
1136, 1136f
procedures, 1130–1131, 1133r
pyloroplasty in, 1127f, 1128, 1129f–1131f
truncal vagotomy in, 1127f, 1128f,
1128–1129
in Zollinger-Ellison syndrome, 1136
Perforator flaps, 1979–1980, 1983f
Perforator vein surgery, subfascial
endoscopic, 1000, 1000f
Perianal space
fallopian tube, in shock, 153
gynecologic, 1787–1789, 1789f
in inguinal hernia, 1607, 1607f
of hand, 1931–1932
Pharyngitis, 622–624
Pharmacologic resistance in tumors, 346
Pharmacodynamics, 2028–2029
Pharmacokinetics, 2028–2029
Pharmacologic resistance in tumors, 346
Pharyngitis, 622–624
Pharyngo esophageal swallowing disorders,
1052–1055, 1053f
Pharynx, 622–624
anatomy of, 630–631
constictor muscles of, 1010f,
1010–1012
in swallowing mechanism, 1016
Phenocopy, 560
Phenotypes
in gene knockout mice, 503
and genotype correlations, 324–325
in transgenic mice, 501
Phenoxybenzamine, 1694
Phentermine, 1162
Phenytoine, 151
Phenylethanolamine N-methyltransferase, 38
Pheochromocytoma, 1693–1695
adrenalectomy in, 1695, 1698
diagnosis of, 1693–1694, 1694f
hereditary, 323, 1695
historical aspects of, 1681
malignant, 1695
in multiple endocrine neoplasia, 1656, 1695
thoracic, 734
Phlebitis, 1997
Phlegmasia alba dolens, 986, 986f
Phlegmasia cerulea dolens, 986, 992
Phosphate, 1669f, 1669–1670
Phosphodiestrase-5 inhibitors, 1776
Phosphoinositide-3-kinase pathway
in Cowden disease, 325
description of, 313, 488
Phospholipase A2, 47, 48f, 1436, 1437f
Phospholipids in bile, 1396
Phosphorus
dietary intake of, 90
serum levels of
in cancer, 99
in hyperphosphatemia, 90, 95, 99
in hypophosphatemia, 90–91, 95, 99
Phosphorylation
description of, 276, 488
oxidative, 137
Photocauterization, infrared, 1310
Photodynamic therapy
in cholangiocarcinoma, 1424–1425
description of, 466
Photoplethysmography, 435
PHOX2B gene, 322f
Phrenic nerve, 680–681
Phrenoesophageal membrane, 1012, 1013f
Phrenosplenic ligament, 1519
Phyllodes tumors, 600–601, 601f
Physical activity, frailty and, 2046f
Physical examination
in arterial disease, 899
in breast cancer diagnosis, 567, 567f
in cardiac disorders, 803
in gynecologic, 1787–1789, 1790f
of hand, 1931f–1932f, 1931–1932
in inguinal hernia, 1607, 1607f
in large intestine disorders, 1264
in mediastinal tumors, 727, 729f
in trauma, 195
Physical therapy in burns, 260
PICT. See Persistent inflammation,
immunosuppression, and
catabolism syndrome
Pigment gallstones
black, 1403
brown, 1403f–1404
Pigmentation, in chronic venous
insufficiency, 983, 983f
in hemorrhagic shock, 144
in massive transfusions, 120
in thrombocytopenia, 109, 423
Platelet count
description of, 103, 123
reduced, in thrombocytopenia, 108–109
Platelet factor 3, 104
Platelet factor 4, 104, 109
Platelet-activating factor, 59
Platelet-derived growth factor
in angiogenesis, 316
in diabetic foot ulcers, 298
in wound healing, 276, 277
therapeutic use of, 298
Platelet-endothelial cell adhesion
molecule-1, 57
Pleomorphic sarcoma and
myxofibrosarcoma, undifferentiated, 535
Plethysmography
description of, 983
in venous insufficiency, 997
in venous thromboembolism, 986–987
Pleural disorders, 736–744
access and drainage in, 736–739
anatomy in, 736
chyllothorax, 741–743
effusions, 736. See also Effusions, pleural
empyema, 740–741
tumors
fibrous, 744
mesothelioma, 743–744
Pleural drainage, 738
Pleural effusions
access and drainage of, 736–739, 737f
diagnostic work-up in, 736–739
differential diagnosis in, 738t, 738–739
etiologies of, 736, 737f
exudative, 737, 737t–738, 738–739
in lung cancer, 688
malignant, 739, 739–740
transudative, 737, 737t–738, 738–739
treatment algorithm on, 740, 740f
Pleural fluid analysis, 738–739
Pleurectomy with decortication in
mesothelioma, 744
Plicae circulares, 1220
Plug and patch technique, in inguinal
hernia repair, 1612–1613
MPS1 gene, 322
MPS2 gene, 322
Pneumatocele, traumatic, 224
Pneumatosis intestinalis
description of, 72, 1231f, 1253f
in necrotizing enterocolitis, 1729, 1729f
Pneumocephalus, 1839
Pneumocystis jiroveci, 364
Pneumocytes, 669
Pneumonecrosis
in lung cancer, 691–692, 694f
in mesothelioma, 744
Pneumonia
aspiration, 706
community-acquired, 708
necrotizing
description of, 708
pathogenesis, 706
nosocomial, 174, 419
pleural effusions in, 736, 737t–738t
postoperative, 419
ventilator-associated
in burns, 258
description of, 174, 419
Pneumonitis, postoperative, 419
Pneumoperitoneum
in bariatric surgery, 1180
in laparoscopy, 455–456
with carbon dioxide, 455, 455f
in pregnancy, 473
in necrotizing enterocolitis, 1729
Pneumothorax
in central venous catheterization, 415–416
in chest trauma
in children, 244
description of, 186, 186f, 199, 222
chest tubes in, 186, 186f, 705
in laparoscopy, 456
open, 186, 186f
postoperative, 419
spontaneous, 705–706
catamenial, 706
tension
description of, 185–186, 186f, 193
obstructive shock in, 150
POCD. See Postoperative cognitive dysfunction
POEM. See Peroral endoscopic myotomy
Point-of-care CT scan, 619, 619f
Poland’s syndrome, 544
Polidocanol, 996
Polgar, R. M., 199, 222
Polley’s syndrome, 544
Polos, S. J., 1529
Polymorphisms, single nucleotide, 485
Polymerase chain reaction technique, 495–496
Polymenorrhea, 1795
Polycythemia vera
description of, 110
splenectomy in, 1522, 1529
Polycystic kidney disease, 1373–1374
Polycystic disease, of liver, 1373–1374
Polycystic liver disease, 1373–1374
Polypectomy
in colorectal cancer, 1242
description of, 1241
independent, 1241
surgical shunts in, 1367–1368, 1368f
in cirrhosis, 1365, 1366
Port Placement, in minimally invasive surgery, 462, 463f
Port wine stains, 1996–1997
Porta hepatis, 1346
Portacaval shunt
in Budd-Chiari syndrome, 1369
in gastrosophageal varices, 1368
Portal vein hypertension, 1365–1367, 1366f
Budd-Chiari syndrome in, 1369
in cirrhosis, 1365, 1368
etiologies of, 1366f, 1366–1367
gastric varices in, isolated, 1153
gastrosophageal varices in, 1367
balloon-occluded retrograde transvenous obliteration, 1367
hemorrhage in, 381, 1367–1368
liver transplantation in, 381
surgical shunts in, 1367–1368, 1368f
transjugular intrahepatic portosystemic shunt in, 1367–1368
imaging in, 1366
liver transplantation in, 381, 1368
minimally invasive surgery in, 474
portal venous pressure measurement in, 1366
splenectomy in, 1523t, 1531–1532
Portal triad, 1557
Portal vein
anatomy of
general, 1348–1349, 1349f, 1365, 1366f
variations in, 1430, 1431f
embolization of, before liver resection, 1385–1386
imaging of, 1366
pressure in
elevated, 1366
measurement of, 1366
normal values for, 1365
with chemotherapy, 1580
complications of, 1578
external beam, 1577
intensity-modulated technique, 1578
metastatic, 1582
timing of, 1577
fractionation of
description of, 347
in hypertrophic scars and keloids, 293
in lung cancer, 697–698
in melanoma, 533–534
in mesothelioma, 744
in nasopharyngeal tumors, 645
in oropharyngeal tumors, 639
in pancreatic cancer, 1490–1491, 1498
in paranasal sinus tumors, 644
physical basis of, 346
planning and simulation of, 347
postoperative, 347
preoperative, 347
in prostate cancer, 1774
reassessment of, 347
in rectal cancer, 1299, 1301
in retroperitoneal sarcoma, 1583
in rhabdomyosarcoma
in children, 1750
description of, 1589
in salivary gland tumors, 650
sensitizing agents in, 347
side effects of, 348, 348f, 349r
skin damage in, 535
solitary thyroid nodule associated with, 1641
in spinal tumors, 1860
in squamous cell carcinoma, 530
in stomach cancer, 1147–1148
in temporal bone tumors, 1200
in cardiac disorders, 804
in atrial septal defect, 753
in aortic dissection, 880
in aortic aneurysm, 858, 858f
Radiodiode. See Iodine, radioactive
Radionuclide scans
in acute cholecystitis, 1399, 1406
biliary atresia, 1377
of biliary tract, 1399
in gastrointestinal, 1137, 1139f
in gastrointestinal hemorrhage, obscure, 1251
in lymphedema, 1002
in Meckel’s diverticulum, 1247, 1248f, 1733
in mediastinal tumors, 730r
of myocardin perfusion, 806
in pancreatic neoplasms, 1480, 1481f
of parathyroid gland
in preoperative localization studies, 1671, 1672f, 1673f
in radio-guided parathyroidectomy, 1674
of stomach, 1117
of thyroid
hot and cold lesions in, 1634, 1634f
in solitary nodule, 1644
in venous thromboembolism, 987
Radiosurgery, stereotactic
description of, 1871–1872
in liver tumors, 1381
Radioluinar joint synovitis, 1946
Radius fractures, 1886–1887
in head, 1886
with ulna fracture, 1886
Radon exposure, 675
RAGE. See Receptor for advanced glycation end products
Raloxifene in breast cancer prevention, 348, 558
Ramsay Hunt syndrome, 616
Randomization, 2144
Randomized controlled trials
all-or-none situation, 2146–2147
analysis of, 2144
challenges associated with applying to surgery, 2144–2147
description of, 2138–2140
of esophagus
description of, 1018–1019, 1021f
perforation, 1083, 1084f
of hand and wrist, 1932, 1932f
in large intestine disorders, 1265
in lung abscess, 707f, 708
in mediastinal tumors, 727
in mesenteric ischemia, 931
in neurologic examination, 1829
of small intestine neoplasms, 1243
obstruction, 1229
radiation enteritis, 1246, 1246f
in spine trauma, 1845
of spleen, 1540
of stomach, 1115
in thoracic aortic aneurysm, 858, 858f
Radioiodine. See Iodine, radioactive
Radiofrequency electrosurgery
Radioiodine. See Iodine, radioactive
Radiofrequency electrosurgery
Radioembolization in liver tumors, 1381
Radiculopathy, 1905
cervical, 1862–1863, 1865r
lumbar, 1864, 1865f, 1865r
Radioembolization in liver tumors, 1381
Radiofrequency electrosurgery
argon beam coagulation in, 465
capacitive coupling in, 465, 466f
in liver tumors, 1380
in varicose veins, 996
Radiography
in aortic aneurysm, 858, 858f
in aortic dissection, 880
in atrial septal defect, 753
in cardiac disorders, 804
in chest trauma, 199, 200f, 200r
in chest wall tumors, 721
in congenital diaphragmatic hernia, 1712, 1712f
of esophagus
description of, 1018–1019, 1021f
perforation, 1083, 1084f
of hand and wrist, 1932, 1932f
in large intestine disorders, 1265
in lung abscess, 707f, 708
in mediastinal tumors, 727
in mesenteric ischemia, 931
in neurologic examination, 1829
of small intestine neoplasms, 1243
obstruction, 1229
radiation enteritis, 1246, 1246f
in spine trauma, 1845
of spleen, 1540
of stomach, 1115
in thoracic aortic aneurysm, 858, 858f
Rasmussen’s aneurysm, 718
Rastelli operation
in Taussig-Bing syndrome, 784
in transposition of great arteries, 782
Raynaud’s disease, 1961
Raynaud’s phenomenon, 1961
Raynaud’s syndrome, 792, 1961
RBI gene, 320, 322r
RCTs. See Randomized controlled trials
Reactive oxygen species, 39
Reading frame, 484
Receptor for advanced glycation end products, 29, 33–34
Receptor-interacting protein, 51
Receptor-interacting protein kinase, 40
Reconstructive surgery. See Plastic and reconstructive surgery
RECQL4 gene, 322r
Recruitment, 11
Rectal arteries, 1260, 1262
Rectal veins, 1262
Rectangular advancement flaps, 1978, 1979f
Rectoepxy, 1306
Rectosigmoidectomy, 1306, 1307f
Rectum
anatomy of, 1260–1262, 1262f
lymphatic, 1262, 1269f
nerve supply in, 1262
vascular, 1262, 1263f
biopsy in Hirschsprung’s disease, 1734
cancer of. See also Colorectal cancer
diagnostic algorithm on, 1300f
direct digital ultrasonography in, 1296, 1298f
epidemiology of, 307f, 308f, 309, 309f, 310r
local therapy in, 1299
locally advanced, 1301
recurrent, 1298, 1302
staging of, 1296
total mesorectal excision of, 338, 1300
tumorigenesis in, 311
carcinoid tumors of, 1303
clinical evaluation of disorders, 1264–1270
digital examination of, 1788, 1789f
embryonic development of, 1259–1260
endoscopy of, 1264–1265
fistula of
description of, 1791
with vagina, 1285, 1317–1318
foreign body in, 1323
gastrointestinal stromal tumors of, 1303–1304
leiomyoma and leiomyosarcoma of, 1303
lipoma of, 1303
lymphoma of, 1303
multidisciplinary approach to disorders of, 1279
prolapose of, 1305–1306, 1306f
resection of
abdominoperineal, 1274
anastomosis in, 1274, 1275f–1276f
anterior, 1273
in carcinoma, 1299–1300
high, 1273
low, 1273
low extended, 1273
Index

Semmelweis, Ignaz Philipp, 157, 271
Senior faculty development, 21
Senning operation, 780, 781f
Sensory tests
  in hand examination, 1931
  in neurologic examination, 1829
Sentinel events, 400t
  high-profile, 401
  root causes of, 402f
  in wrong-site surgery, 412
Sentinel lymph nodes
  in breast cancer, 338–339, 590–591
  and breast conservation surgery, 591
  dissection, 589f, 589–590
  in melanoma, 531, 532f–533f
  in vulvar cancer, 1810
Sepsis, 175–177
  anorectal
    description of, 1313–1314
    in immunocompromised host, 1316
    apoptosis during, 40
    definition of, 29t
    diagnostic criteria on, 146
    disseminated intravascular coagulation in, 110
    guidelines on, 175–177
    lipopolysaccharide in, 1355
    lymphopenia and, 40
    mortality rate in, 146
    in necrotizing enterocolitis, 1729
    in parenteral nutrition, 74
    postoperative, 426
    postplenectomy, 228, 1541–1542
    preoperative optimization during, 2126–2127
  severe, 146, 161, 426
  shock in, See Shock, septic
  subareolar, 554–555, 555r
  systemic inflammatory response syndrome in, 160, 161t
  thrombocytopenia in, and bleeding, 125
  treatment of, 146, 146–148, 176t
Septal defects. See Atrial septal defect;
Ventricular septal defect
Septic shock, 29t
  causes of, 146, 146r
  complement activation in, 140
  definition of, 160–161
  description of, 131, 161, 426
  diagnosis of, 146
  hemodynamic response to, 135t
  pathophysiology in, 133
  resuscitation endpoints in, 152
  treatment of, 146, 146–148
  Sequential organ failure assessment (SOFA) score, 1444t
Sequestration
  platelet, in spleen, 109, 1522
  pulmonary, 1714f, 1715–1716
Seroma
  in inguinal hernia repair, 1621
  in mastectomy complications, 593
Serosa of small intestine, 1221
Serotonin, 50, 1439
Serpentine asbestos fibers, 744
Serrated polyposis syndrome, 1291
Serratia anterior, 2012
Sertoli cell tumors, 1819
Serum markers, of mediastinal tumors, 727–729
Setons, 1317
Severe sepsis, 29t
Sevoflurane, 2030, 2030v
Sex cord-stromal cell tumors, 1746, 1819
Sex steroids
  description of, 1684
  excess secretion of, 1692–1693
  synthesis of, 36f
Sexual abuse of child, 1745
Sexual development
  ambiguous genitalia in, 1746–1747
  embryology in, 1746
  Sexually transmitted diseases, 1320
  SGLT1, 1223, 1224f
  SHD, B, D gene, 323r
  Shen Nong Ben Cao, 2124
  Shimada classification of neuroblastoma, 1749
Shock, 131–153
  cardiogenic
    description of, 131, 148t, 148–150
    differential diagnosis of, 193
    hemodynamic response to, 135t
    in trauma, 149, 193
    cellular hyperperfusion in, 137
    classification of, 131–133, 132t
    compensated phase in, 133, 134f, 152
    danger signaling in, 137, 139f
    decompensated phase in, 133, 134f
    definitions of, 133, 148, 152
    hemodynamic response to, 135, 135t,
    148
    historical aspects of, 131–133
    hypovolemic/hemorrhagic, 141–145
    base deficit in, 142, 143f
    blood pressure in, 141, 142f, 144
    in children, 1752
    description of, 131
    diagnosis of, 141–143
    hemodynamic response to, 135t
    hypothermia in, 145, 209
    neutrophils in, 140
    pathophysiology in, 134f, 140
    resuscitation endpoints in, 152
    signs and symptoms in, 190–192, 192r
    source of blood loss in, 143–144
    thermal protection in, 209
    in trauma, 190–193, 192r, 209
    treatment of, 143–145
    volume of blood loss in, 141, 142r
    inflammatory and immune responses in, 137–141
    irreversible phase of, 133, 134f
    metabolic effects of, 136–137
    neurogenic
      description of, 131, 151t, 151–152
      hemodynamic response to, 135t
      in trauma, 151, 193
      obstructive, 131, 150t, 150–151
      overview of, 131
      pathophysiology of, 133f, 133–136, 134f
      afferent signals in, 134–135
      circulatory homeostasis in, 136
      effluent signals in, 135–136
      “vicious cycle” in, 133
      resuscitation endpoints in, 134, 152–153
      cellular, 152, 152r
      systemic, 152, 152r
      tissue-specific, 152, 152r
      septic
        causes of, 146, 146r
        complement activation in, 140
        definition of, 160–161
        description of, 131, 161, 426
        diagnosis of, 146
        hemodynamic response to, 135t
        pathophysiology in, 133
        resuscitation endpoints in, 152
        treatment of, 146r, 146–148
  in trauma
    cardiogenic, 149, 193
    in children, 1752
    description of, 132, 145, 209
    hemorrhagic, 142–143
    hypovolemic/hemorrhagic, 190–193, 192r, 209, 1752
    initial evaluation and management of, 190–193
    neurogenic, 151–152, 193
    obstructive, 150
    persistent hypotension in, 193–195
    resuscitation endpoints in, 152–153
    signs and symptoms in, 190–192, 192r
    vasodilatory, 146–148
    vasogenic, 131
  Shock lung, 132
  Shockwave lithotripsy, 1763
Short bowel syndrome
  intestinal lengthening procedure in, 1255
  intestinal transplantation in, 1255
  in radiation enteritis, 1246
  risk factors for, 1254, 1254r
  serial transverse enteroplasty in, 1254f, 1255
  Short esophagus, 1049
Shotgun sequencing, 498
Shotgun wounds, 196
  abdominal, 201
Shoulder, 1895–1897
  acromioclavicular joint in, 1897
  dislocations of, 1884
  impingement syndromes of, 1896–1897
  labrum tear, 1896, 1897f
  Shouldice repair, 1599, 1610, 1611f
Shunt
  in hypoplastic left-heart syndrome, 775–776, 776f
  left-to-right
    in aortopulmonary window, 769
    in atrial septal defect, 752–753
    in duc tus arteriosus patency, 760
  in portal hypertension and
gastroesophageal varices
  surgical shunts in, 1367–1368, 1368f
  transjugular intrahepatic portosystemic\ntosystemic shunt in, 1367–1368
  right-to-left
    in hypoplastic left-heart syndrome, 774
    in total anomalous pulmonary venous connection, 766
  transjugular intrahepatic portosystemic\nsystemic, 381
  in Budd-Chiari syndrome, 1369
  in gastric varices, isolated, 1153
  in gastroesophageal varices, 1367–1368
  in tricuspid atresia, 771–773, 772f
Sandhu Ramesh, 1967, 1968f
future approaches to surgical closure, 754–755
with partial anomalous pulmonary venous connection, 754
pathophysiology in, 753
treatment of, 754
Sinuses
paranasal
fungal ball in, 620f
inflammatory disease of, 617–622.
See also Sinusitis
tumors of, 642f, 643–644
thyroglossal duct and, 1625–1626
venous, 982
Sinusitis, 617–622
acute, 617–618
chronic, 618–620
diagnosis of, 618f–619f, 618–619
management of, 619
and nasal polyps, 618, 619f, 620
classification of, 617
fungal, 620f
history consistent with, 617
Sirolimus, 359f, 361f, 361–362
mechanism of action, 361
tsides effects of, 360f, 362
Sistrunk operation, 1626
Situational debriefing model, 403
Skene’s glands, 1785
Skills
description of, 2163
dreyfus model for development for,
2165f
Skills labs, 2163–2165
Skills training
description of, 2163–2165
virtual reality for, 2171
Skin, 513–536
anatomy and histology of, 513–517,
514f, 1966, 1975
dermis layer of, 516
epidermis layer of, 513–514
grafts of. See Skin grafts
in HIV infection and AIDS, 527
in host defense mechanisms, 159
hypodermis of, 517
epidermis of, 173–174, 524–527
inflammatory diseases of, 517–518
injuries of, 518–524
from caustic substances, 520–521
coagulative, 520
in intravenous fluid extravasation,
521, 522f
lequifactive, 520
mechanical, 519
pressure, 523f, 523–524
in radiation exposure, 518–519, 528,
535
temperature-related, 521–523
thermal, 521–523, 522f–523f
innervation of, 517
substitute replacement products for. See
Skin substitutes
ulcers of, 290, 523f, 523–524.
See also Ulcer(s), of skin
vasculature of, 516
wound healing in, compared to
gastrointestinal wound healing, 280f
Skin grafts, 295–297
alloigenic, 296
autologous, 296
in breast reconstruction, 594
in burns, 257, 259
donor sites in, 259
of hand, 1959
full-thickness, 296
in head and neck tumors, 651
split-thickness, 296
xenogeneic, 296
Skin incisions, 1968–1971, 1969f
Skin substitutes
in burns, 259–260
description of, 259–260, 298f, 324
in venous insufficiency, 1000, 1000f
Skin tags
anal and perianal, 1310
in Crohn’s disease, 1285
Skin tumors, 527–536, 1955
adjutant therapies for, 533–534
benign, 527–528
description of, 518
of ear, 645
epidemiology of, 307f, 308f–309f, 310f
in hand, 1955f, 1955–1956
malignant, 528–536
melanoma, 530–534. See also
Melanoma
palliative therapies for, 533–534
in sun exposure, 326
surgical management of, 532–533
in transplant recipients, 364, 369
Skull
craniosynostosis of, 1872–1873
fractures of, 1835, 1835f, 1839
osteomyelitis of, 1867–1868
Skull base surgery, endoscopic, 620–621
Sleep apnea, obstructive, 623
in bariatric surgery, 1959
in bariatric surgery, 1959
in bariatric surgery, 1959
Sleep-disordered breathing, 623
Sleeve gastrectomy, laparoscopic
background on, 1185–1186
complications of, 1186–1187, 1206
glycemic control after, 1178f
history of, 1168
mechanism of action, 1176f
patient selection for, 1185
port scheme for, 1185f
present-day status of, 1169
revision for, 1206
technique for, 1185–1186, 1186f
SLI. See Surgeon’s Leadership Inventory
Sling procedures, in stress urinary
continence, 1808–1809
Slipped capital femoral epiphysis, 1918
SMAD4, 489
SMAD pathway, 1974
Small cell carcinoma of lung
description of, 673
hypertrophic osteoarthropathy in, 681,
682f
paraneoplastic syndromes in, 681–683
Small intestine, 1219–1255
anatomy of, 1219–1220, 1220f
atresia of, 1724–1725
apple peel deformity in, 1725
Christmas tree deformity in, 1725,
1725f
duodenal, 1720, 1724–1725, 1725f
ileal, 1725f
jejunal, 1725f
Small cell carcinoma of lung
description of, 673
hypertrophic osteoarthropathy in, 681,
682f
paraneoplastic syndromes in, 681–683
Small intestine, 1219–1255
anatomy of, 1219–1220, 1220f
atresia of, 1724–1725
apple peel deformity in, 1725
Christmas tree deformity in, 1725,
1725f
duodenal, 1720, 1724–1725, 1725f
ileal, 1725f
jejunal, 1725f
Small intestine, 1219–1255
anatomy of, 1219–1220, 1220f
atresia of, 1724–1725
apple peel deformity in, 1725
Christmas tree deformity in, 1725,
1725f
duodenal, 1720, 1724–1725, 1725f
ileal, 1725f
jejunal, 1725f
barrier functions of, 1225–1226, 1226f
in chylous ascites, 1252–1253
Crohn’s disease of, 1235–1240,
1284–1285
digestion and absorption in, 1222–1225
carbohydrates, 65, 1223–1224, 1224f/
fat, 1224–1225, 1225f
fluids and electrolytes, 1222f,
1222–1223, 1223f
protein, 1224, 1224f
regulation of, 1223f
vitamins and minerals, 1225
diverticulum of, 1246–1250. See also
Diverticulum, small intestine
duodenum in. See Duodenum duplicatured, 1733
embryonic development of, 1221, 1222f
duodenal endocrine functions of, 1127–1128,
1227f
enteral nutrition access using, 71t,
71–72
fistulas of, 1240–1241
hemorrhage from, 1250–1251
histology of, 1220–1221, 1221f
ileum in. See Ileum
ileus and motility disorders of,
See Ileum
jejunum in. See Jejunum
intussusception of, 1253
jejenum in. See Jejunum
mesenteric ischemia of, 1250
motility of, 1226–1227, 1227f
ascending excitation and descending
inhibition of, 1227, 1227f
disorders of, 1233–1235
fasting pattern or interdigestive motor
inhibitory cycle in, 1226
fed or postprandial pattern in, 1226
regulation of, 1227
rhythmic segmentations in, 1226
neoplasms of, 1241–1245
clinical presentation in, 1243
diagnosis of, 1243–1244
epidemiology of, 307t
intestinal obstruction in, 1243
pathophysiology in, 1242–1243
prognosis and outcomes in, 1245
treatment of, 1244–1245
types of, 1242f, 1243
obstruction of, 1228–1233
clinical presentation in, 1229
closed loop, 1229
complete, 1229, 1230f, 1231
diagnosis of, 1229–1231, 1230f
double bubble sign in, 1724, 1724f/
in early postoperative period, 1231
epidemiology of, 1228–1229
etiologies of, 1228t, 1228–1229
in inguinal hernia, 1609
in laparoscopic Roux-en-Y gastric
bypass, 1184
in Meckel’s diverticulum, 1247, 1247f/
in neonates, 1723–1724, 1724f
in neoplasms, 1243
outcome in, 1232
partial, 1229, 1230f, 1231
pathophysiology in, 1229
postoperative, 421
prevention of, 1233
strangulated, 1229
treatment for
jejunum in.
See Jejunum
Bayesian statistics, 2148
P values, 2147–2148
significance level, 2147
type I error, 2147
type II error, 2147
Status epilepticus, 1833
Steatocarcinoma, 1459
Steinstrasse, 1763
Stellate cells, pancreatic, 1456
Stem cells
cancer of, 318
description of, 492f, 492–493
targeting vector, 503

Stent grafts
in abdominal aortic aneurysm, 919, 922–927, 923f
compared to open repair, 925–927
complications of, 924–925, 928
devices available for, 924, 926
endoleaks in, 924, 927f, 927f, 927f–927f
927–928
tendovascular aneurysm in, 928
in aortic dissection, 884
in aortoiliac occlusive disease, 949–950
description of, 906–907, 907f
in femoropopliteal occlusive disease, 968
in lower extremity arterial occlusive disease
complications of, 965–966
description of, 963–964
in thoracic aortic aneurysm
complications of, 875
description of, 865, 868, 872
endoleaks in, 875
Stent placement, 468–470, 469f
in aortic aneurysm, 865, 868, 872
complications of, 875
devices available for, 875
in aortic dissection, 884
in aortic injuries, 222
in aortoiliac occlusive disease, 947–949
biliary
in cholangiocarcinoma, 1424
in chronic pancreatitis, 1471
in pancreatic cancer, 1489–1490, 1490f
in carotid artery occlusive disease, 911–912, 914–917
complications of, 916
contraindications to, 914, 915f
embolic protection devices in, 915f, 915–916
derendarterectomy versus, 911–912
stents for, 915, 916f
technique in, 914–917, 915f
in coronary artery disease, 810
in esophagus
in cancer, 1074
in caustic injury, 1087, 1087f
indications for, 469, 905
in lower extremity arterial occlusive disease, 963, 965f
965–966
in mesenteric artery occlusive disease, 933–934
in pancreatic pseudocyst, 1467, 1467f
in renal artery occlusive disease
clinical results in, 939–940, 940f, 941f
technique in, 940
ureteral
before colorectal surgery, 1279
ureter trauma caused by, 1765
vascular stents in, 469–470, 905–906, 906f
balloon-expandable, 906, 906f
in carotid artery occlusive disease, 915, 916f
drug-eluting, 469, 906
self-expanding, 469f, 470, 906, 906f
Stereotactic body radiotherapy, 695–696
Stereotactic radiosurgery
in arteriovenous malformations, 1872
in intracranial metastasis, 1872
in liver tumors, 1381
in prostate cancer, 1872
in vestibular schwannoma, 1872
Sterilization procedures, tubal, 1802
Stemovacular displacement, 1883
Stentomy
in chest trauma, 210
in coronary artery bypass surgery, 811
in goiter, 1661, 1661f
in hyperparathyroidism, 1676
in mediastinal mass biopsy, 730
Sternum fractures, 224
Stevens-Johnson syndrome, 517–518
Stewart-Hamilton equation, 439, 443
Stewart-Treves syndrome, 535
STI571. See Imatinib
Stickler syndrome, 1993
STK11 in Peutz-Jeghers syndrome, 323t
Stoma procedures. See Ostomy procedures
Stomach, 1099–1160
in Peutz-Jeghers syndrome, 323t
vasculature of, 1100, 1103f
complications of, 875
in arteriovenous malformations, 1872
in carotid artery bypass surgery.
blood supply of, 1100, 1103f
bypass surgery. See Gastric bypass, Roux-en-Y
cancer of, 1139–1151
Gastric banding
in breast cancer, 1154
blood supply of, 1100, 1103f
bypass surgery. See Gastric bypass, Roux-en-Y
cancer of, 1139–1151
carcinogenesis in, 1141f
chemotherapy in, 1147–1148
clinical manifestations of, 1144–1145
diagnostic evaluation in, 1145–1146
differentiated from peptic ulcer disease, 1145
diffuse form, 1139
early, 1143, 1143f, 1144f
endoscopic ultrasound in, 1116, 1145
epidemiology of, 307r, 307–308,
310r, 311f, 1139, 1140f
etiology of, 1140–1143
fungating, 1144
gastrectomy in, 1146f, 1146–1147, 1147f
endoscopic, 1148–1149
reconstruction after, 1146, 1147f
survival rates in, 1147, 1147r
in gastrectomy remnant, 1143
in Helicobacter pylori infections
description of, 325, 1141f
risk for, 1139
hereditary diffuse, 325

Index
Stomach, cancer of (Cont.):
- histologic subtypes of, 1143t–1144t, 1144
- intestinal form, 1139
- lymphadenectomy in, 1101, 1147
- mortality rates in, 1140t
- pathology in, 1143–1144
- polypoid, 1144
- prognosis in, 1149
- radiation therapy in, 1147–1148
- risk factors for, 1140t, 1140–1143
- scirrhous, 1144
- screening for, 1149
- staging of, 1144, 1145t
- treatment, 1145–1151, 1147–1150
- ulcerative, 1144
- World Health Organization
  - classification of, 1144, 1144t
- carcinoid tumors of, 1151–1152
- composition of secretions, 87t
- description of, 1102–1103, 1105f
- open, 1106–1107
- peritoneal, 1107
- pyloric, 1107
- rectal, 1107
- rare, 1107
- rectosigmoid, 1107
- small bowel, 1107
- ultrasonic, 1107
- vapor, 1107
- vascular, 1107
- vesical, 1107
- vesicocolic, 1107
- vesicovaginal, 1107
- ventral, 1107
- ventral rectal, 1107
- vertical, 1107
- v-shaped, 1107
- wedge, 1107
- wide, 1107
- wing, 1107
- winged, 1107
- x-ray, 1107
- xenon, 1107
Supracolic injuries, 212
Supraglottic laryngectomy, 642
Supraglottis
anatomy of, 631, 640
tumors of, 641–642
Suprarelevator space
abscess of, 1316
description of, 1313
Supraspinatus tendon impingement, 1896
Surgeon
career development for, 20–21
leadership training for, 20
performance of, 16, 16r–17r
senior faculty development, 21
Surgeon’s Leadership Inventory, 19r
Surgeons Over Seas Assessment of Surgical Need (SOSAS) survey, 2082
Surgery
bowl preparation for, 2126
carbohydrate loading before, 2117
leadership in. See Leadership in surgery
metabolic stress response to, 2116f,
2116–2117
women in, 12
Surgical care
advanced care in resource poor areas,
2102–2103
cost-effectiveness of, 2098–2099
description of, 5–7
factors affecting utilization and outcome of,
2099f, 2099–2101
preventive role, 2098
Surgical Care Improvement Project,
406–407
Surgical Council on Resident Education
curriculum, 2187
description of, 2164
Surgical counts, for retained surgical item
prevention, 410–411
Surgical leadership. See Leadership in surgery
Surgical residency
bootcamps for, 2165
goals for, 2177
simulation in, 2165
Surgical site infections
description of, 164, 170f, 174–175, 286,
406–407, 424
incidence of, 407
prevention and treatment of, 164, 286,
287r–288r, 406–407, 424
prevention of, 2117
quality improvement efforts related to,
171, 406–407
rates, 169r
risk factors for, 169, 169r
Surgical skills
Fundamentals of Robotic Surgery,
2174–2175, 2175f, 2175r
laparoscopic, 2167–2168
open, 2166–2167
simulation training of, 2166–2168
Surgical training model, 2163
Surgeicenter, 2155–2156, 2156f
Surviving Sepsis Campaign, 2148
Sustainable development goals, 2088,
2088t
Suture techniques, 295
Swallowing
barium studies of, 1018–1019, 1021f
cineradiography of, 1028, 1029f
innervation of, 1016
motility disorders of, 1052–1060
peristalsis in, 1016
pharyngoesophageal phase of,
1052–1055, 1053f
physiology of, 1015f, 1015–1017
pressures in, 1015f, 1015–1016,
1016f
rehabilitation of, in laryngeal cancer,
643
sequence of events in, 1015, 1015f
videoradiography of, 1028, 1029f
Swan-Ganz catheters, 416
Sweat glands, 515
Swedish Obese Subjects study, 1193,
1197
Swenson, Orvar, 1734
Swenson procedure in Hirschsprung’s
disease, 1735, 1735f
Swimmer’s ear, 613
Symmastia, 544
Sympathetic nervous system
of anorectum, 1263
of colon, 1261
description of, 35, 1829
of liver, 1350
of pancreas, 1435
in shock response, 135
of stomach, 1102, 1110
of thyroid gland, 1629
Symphysiosis pubis, 1784–1785
in trauma, 192
Symphonitis, of distal radioulnar joint,
281
Synergistic effect, 2029
Synovial sarcoma, 1956
Synovitis, of distal radioulnar joint,
1946
SYNTAX trial, 810
Syphilis, 1320
of aortic aneurysm in, 856
genital ulcers in, 1791, 1792r
Syndactyly, 1961, 1962
Syndesmosis of ankle, 1893
Synergistic effect, 2029
Symphysis pubis, 1784–1785
in wound healing, 272–273
Temporal arteritis, 856, 969
Temporalis muscle, 2004
Temporalis muscle, 2004
in minimally invasive surgery, 457–458
in operating room, 401–402, 402f, 403r
Team-based health care, 2176r
TeamsRPEPS, 2178–2179
Technetium-99m radionuclide scans
in cardiac disorders, 806
in thyroid, 1634
Telecanthus, 2005
Teleconsultation, 475
Tele-education, 2085
Telehealth, 2106
Teledicine, 2085, 2106
Telementoring, 475, 2106
Telescopes in minimally invasive surgery,
462, 464f
Telesurgery, 468
Temporal arteritis, 856, 969
Temporal bone
fractures of
facial nerve injuries in, 628–629
longitudinal, 629
transverse, 629
tumors of
reconstructive surgery in, 645–646
resection techniques in, 644f,
645–646
Temporal lobe
anatomy of, 1828
anterior lobectomy in epilepsy, 1869
Temporalis muscle, 2004
Temsirolimus, 343r
Tendons
in hand, 1929. See also Hand, tendons in
healing of, 281
in rotator cuff, impingement of, 1895f,
1896
Tenecteplase, 990
Tenodesis maneuver in tendon injuries of
hand, 1935
Tenosynovitis of hand and wrist flexor, 1950f, 1950–1951, 1951r
Tension pneumocephalus, 1839
Tensor veli palatini, 1984
Teratocarcinoma, mediastinal, 735
Teratoma
central nervous system, 1859
cervical, ex utero intrapartum therapy
(EXIT) procedure in, 1754, 1754f
mediastinal, 735
ovarian, 1750
retroperitoneal, 1750
retrorectal, 1304
sacroccoygeal, 1750f, 1750–1751
thoracic, 1750

Terminal illness. See End-of-life care

Tertiary intention, 1975

Tessier classification, of craniofacial clefts,

Thalidomide, 189

Thalamus, 1828

Thalian, 1828

Thalassemia, 1522, 1525–1526

Thallium-201 scans of myocardial perfusion, 806

Thames, 1819

Therapeutics index, 2029

Therapeutically Applicable Research to Generate Effective Treatments, 319

Thermal agents for local hemostasis, 114–115

Thermal injuries of skin, 521–523, 523f–523f

Thermography, 439

Thermoregulation in children and infants, 1709

Therapeutic approach, 427–428

Thrombocytopenia, 108–109

Thrombosis in arterial thrombosis, percutaneous technique, 956

Thrombotic thrombocytopenic purpura, 259

Thrombotic thromboembolism in arterial thrombosis, 946

Thrombotic thrombocytopenic purpura

Thromboangiitis obliterans, 974–975, 1960

Thromboembolectomy, 1522, 1525–1526

Thrombolysis in myocardial infarction, 1529

Thrombotic thrombocytopenia, 108–109

Thrombomicroangiopathy, 108, 108f

Thrombomicroangiopathy

Thrombotic thrombocytopenic purpura

Thrombotic thromboembolism

Thrombolytic therapy

in lower extremity arterial occlusive disease, 955–956, 956f

in mesenteric ischemia, catheter-directed therapy in, 933

in venous thromboembolism, 987–988
Index

Tic douloureux, 1871
Ticarcillin-clavulanate, 165t
Tie-over bolster dressings, 1976
Tigecycline, 167t
Tillaux’s sign, 1560
Time management in surgery, 10f, 10–11
Time management matrix technique, 10, 10f
Time out protocol, 9
Timed up and go test, 2049
Tinel’s sign, 1944
TIPS. See Transjugular intrahepatic portosystemic shunt
Tissue engineering, 814
Tissue expansion, 1981, 2001f
Tissue factor pathway inhibitor, 106
Tissue heart valves, 816, 817f
Tissue plasminogen activator in arterial vasospasm, 237r
description of, 106, 315
in ischemic stroke, 1850
in venous thromboembolism, 990
TNFR-1, 42
TNFR-2, 42
TNFR-associated death domain (TRADD), 51
TNM staging of cancer
adrenocortical, 1691, 1692t
of breast, 575, 576t–577t
colorectal, 1295, 1297r
description of, 333
of head and neck, 632, 632t–634t
of lungs, 689–691, 690s–691t
of stomach, 1144, 1145t
of thyroid, 1648, 1649r
To Err Is Human: Building a Safer Health System, 8, 12
Tobacco use. See also Smoking
head and neck tumors caused by, 630, 639
Tobin Index, 420
Tobramycin, 1666, 287t
Toker cells, 515
Tolerance, 2029
Toll/IL-1R homology domain, 32
Toll-like receptors
description of, 32
ligand-binding domain of, 32
in liver, 1357
in shock, 137, 139f
signaling events, 32
TLR4, 37, 1357
Tongue
cancer of, 634–636, 635f, 639
strawberry, 622
Tongue-lip adhesion, 1995
Tonsil(s)
hypertrophy of, 622–624
infections of, 622–624
tumors of, 639
Tonsillar herniation, 1833f
Tonsillectomy, 623
Tonsillitis, 622–623
Topical hemostatic agents, 115
Topoisomerase II, 340
TORCH infections, 1737
Torsion
ovarian cyst, 1335, 1746
testicular, 1769
Torticollis, congenital, 1712
Total endoscopic coronary artery bypass surgery, 813
Total intravenous anesthetics, 2030
Total pancreatectomy with islet auto-transplantation, 1480f, 1480–1481
Toupet fundoplication
description of, 1041, 1044
esophageal myotomy with, 1064, 1066f–1067f
Tourniquet, for local hemostasis, 188
Toxic epidermal necrolysis, 517–518
TP53 gene, 323t
TPTD. See Transpulmonary thermodilution
Trabectedin, 1579
Trachea, 661–667
anatomy of, 661, 663f–664f
blood supply of, 661, 663f
C-rings, 661
fistulas of, 664–665
with esophagus, 665, 1717–1721.
See also Tracheoesophageal fistula
intubation of, 2037. See also Endotracheal intubation
neoplasms of, 665–667, 667f
stenosis of, postintubation, 663–664, 664f
trauma of, 185, 186f, 199, 222, 224
iatrogenic, 661, 664f
Trachelectomy, radical, 1813
Tracheobronchial lymph nodes, 667
Tracheobronchial tree
cell types in, 669, 669f
histology of, 668–669, 669f
salivary gland-type tumors of, 672
Tracheoesophageal fistula, 664–665, 1717–1721
anatomic varieties of, 1717, 1717f
in cancer of esophagus, 1069
clinical presentation in, 1717f, 1718
complications of surgery in, 1720
etiology and pathologic presentation in, 1717–1718
H-type, 1717, 1717f, 1720
initial management of, 1718
outcome in, 1721
postoperative management in, 1720
in preterm infant, 1718–1719
primary surgical correction of, 1719f, 1719–1720
recurrent, 1720
single-stage operation for closure of, 666f
in ventilatory support, 664–665
Tracheoesophageal puncture, for speech rehabilitation after laryngectomy, 643
Tracheoinnominate artery fistula emergency management of, 665, 665f
in tracheostomy complications, 417, 417f, 664–665
Tracheostomy
in burns, 258
complications of
tracheal stenosis in, 663–664, 664f
tracheoinnominate artery fistula in, 417, 417f, 664–665
in trauma, 185
in children, 244
Tracheotomy, 653–654
complications of, 653
indications for, 653
speech with, 653–654
Traditional Chinese medicine in appendicitis, 2127
in biliary disease, 2127–2128
bowel preparation for surgery, 2126
in colon surgery, 2127
history of, 2124, 2126
in intestinal obstruction, 2128
in pancreas surgery, 2128
postoperative ileus prevention using, 2127
preoperative nutritional optimization, 2126
preoperative optimization during sepsis and infection, 2126–2127
TRAM flap. See Transverse rectus abdominis musculocutaneous flap
Tranexamic acid, in blood transfusion, 116–117
Transanal endoscopic microsurgery
description of, 470–471, 472f, 1264
in rectal cancer, 1299
Transanal excision in rectal cancer, 1299
Transanal minimally invasive surgery, 1264, 1299
Transarterial chemoembolization, 1484–1485
Trans-Atlantic Inter-Society Consensus classification of aortoiliac occlusive disease, 942, 944f, 944r
of lower extremity arterial occlusive disease, 952, 955f, 955r
Transcatheter aortic valve replacement, 829, 2056
Transcription
after blunt trauma, 52
control of, 484, 485f
description of, 311, 479, 480f, 481, 483
epigenetic regulation of, 53
in eukaryotes, 483
in prokaryotes, 483
Transcription factors, 485
Transdifferentiation, 492
Transesophageal echocardiography, for anesthetic monitoring, 2033, 2033r
Transfection
description of, 499f, 499–500
small interfering RNA in, 504
transient, 500
“Transfer effectiveness ratio,” 2169
Transformation procedure in recombinant DNA technology, 493
Transformational leaders, 12
Transforming growth factor-β, 51
platelet expression of, 56
signaling pathway of
in cancer, 489–490
description of, 489f, 489–490
in wound healing
description of, 272, 275, 277f
hypertrophic scars and keloids, 293
Transfusions, 115–123
ABO blood groups in, 115
allergic reactions in, 121, 122r
in anemia, 117
banked whole blood, 115
in burns, 255
in children, 1707–1708
circulatory overload in, 122r
Traditional Chinese medicine
complications of, 121–123, 122t, 423t, 423–424
in damage control resuscitation, 117, 119
febrile reactions in, 121, 122r
fresh whole blood, 115, 207
hemoglobin in, human polymerized, 117
hemolytic reactions in, 122r, 122–123
in hemorhagic shock, 144–145
indications for, 117, 207, 423
infection transmission in, 123, 423, 423r
massive, 117, 120r
protocol in, 207, 208rf
thrombocytopenia and hemorrhage in, 124
plasma in, fresh-frozen, 116, 207
platelet, 109, 116, 118r
in massive transfusions, 120r
in thrombocytopenia, 109, 423
prehospital, 119, 121
purrpa in, 124
red blood cell
in children, 1707–1708
frozen, 115–116
in hemorhagic shock, 144–145
indications for, 117
packed, 119, 207
prehospital, 119
in resuscitation strategies, 117, 119, 120r
in thalassemia, 1526
washed/leukocyte-reduced, 116
in immunobiology in, 357
transfusion in, 1568
of spleen. See Spleen, trauma of
Brunicardi_Index_Vol II.indd   89 22/02/19   11:26 AM
Trauma (Cont.):
of testis, 1767–1768
of thoracic and lumbar spine, 1902
of trachea
acute management of, 664
description of, 185, 186f; 199, 222, 224
in intubation, 661, 663–664, 664f
transfusions and resuscitation strategies in, 117
of ureters, 234, 1765–1766
of urethra, 234, 1766–1767
of urinary tract, 234, 1764–1768
in laparoscopy, 1821
vascular injuries in. See Vascular injuries
wound healing in, 271–300
Traumatic brain injury, 217, 1836–1842
anticoagulation reversal and prophylaxis in, 1840
blood pressure management in, 1840
cerebral perfusion pressure in, 1840
in children, 244, 1752
contre-coup, 1837
contusion in, 1837, 1837f
decompressive craniectomy for, 1841
diffuse axonal, 1837
high-mobility group protein B1 levels after, 30
hypertension in, 1840
infratentorial pressure increase in, 447–448
intraparenchymal hemorrhage in, 1839
invasive intracranial monitoring in, 1840–1841
management of, 1839–1842
monitoring in, 447–449
pneumocephalus associated with, 1839
steroids for, 1839–1840
Treacher Collins syndrome, 1990, 1991f
Tremors
deep brain stimulation in, 1869
in Parkinson’s disease, 1870
Trendelenburg gait, 1905
Trendelenburg’s test, 983
Treponema pallidum infections, 1792f.
See also Syphilis
Trial Assessing Individualized Options for Treatment (TAILORx), 334, 580
Triangle of Calot, 1393
Triangle of dome, 1603, 1605f
Triangle of pain, 1603, 1605f
Triangular ligaments of liver, 1346, 1347f
Tricarboxylic acid cycle, 64
Trichilemmal cyst, 527
Trichobezoars, 1154, 1154f
Trichomoniasis, vulvovaginitis in, 1789f,
1793, 1794f
Tricuspid valve disorders, 829–831
annuloplasty in, 830
atresia, 770–773
classification of, 771
Fontan procedures in, 772–773, 773f
Glenn shunts in, 772f, 772–773
treatment of, 771–773
in Ebstein’s anomaly, 777–780
insufficiency, 829–831
prosthetic replacement in
in Ebstein’s anomaly, 779
stenosis, 829–831
Trigeminal neuralgia, 1871
Triglycerides
long-chain, 63, 1224–1225
medium-chain, in enteral nutrition
formulas, 64, 1225
metabolism of
in fasting, 60f–61f
in injury, 62f–63f, 63–64
in small intestine, 1224–1225, 1225f
several levels in chronic pancreatitis, 1452
short-chain, 1225
Trigonocephaly, 1993
Triiodothyronine, 1630
Trypsinogen
activation of, 1442
description of, 1224, 1224f, 1436, 1437f
Trypsinogen
activation of, 1442
inhibition of, 1436
description of, 1436, 1437f
TSC1, 323t
TSC2, 323t
TSHR gene, 323t
Tube thoracostomy, 417
Tubular carcinoma of breast, 566
Tuberculosis, 710–711, 712
Tube thoracostomy, 417
Tuberculosis, 710–711, 712
hypercalcemia in, 99, 681–682, 1668
in kidney transplant recipients, 369
osteonectin in, 1868
surgical treatment in, 711, 712
surgical treatment for, 711, 712f
Tuberous sclerosis, 323
Tuberculous sclerosis, 323f
Tubulocarcinoma of breast, 566
Tuftsin, 1521
Tularemia in biologic warfare, 179
Tumor(s), 305–350.
See also specific anatomic locations
adjuvant therapy in, 305
of adrenal gland, 1691–1692
anal, 1304–1305
of appendix vermiform, 1338–1339
of biliary tract, 1421–1425
biopsy of, 310–318, 313f
angiogenesis in, 316
apoptosis alterations in, 314–315
autophagy in, 315
cell cycle dysregulation in, 311–312, 313f
cell proliferation and transformation in, 310
invasion in, 315–316
oncogenes in, 310–311, 313–314
stem cells in, 318
steps in tumorigenesis in, 310–311
biomarkers of, 334–337. See also Biomarkers, tumor
of bladder, 1770–1771
of breast, 555–602
of carotid body, 918f, 918–919
cartilage-forming, 1912
of central nervous system, 1854–1861
cerebellar, 613f
chemotherapy in, 340–342, 491.
See also Chemotherapy
of chest wall, 720–726
in children, 1745–1746
circulating tumor cells in, 336–337
colorectal, 1288–1303
diagnosis of, 332–333
dormancy phenomenon in, 316
of endometrium, 1813–1815, 1814f
epidemiology of, 306–310
incidence rates, 306–310, 307f,
308f–309f, 311f
initiation of, 310–311
mortality rates, 306–310, 307f, 308f
survival rates, 306–307, 310f
of esophagus, 1068–1082.
See also Adenocarcinoma, of esophagus
etiology of, 318–328
of fallopian tubes, 1815–1816
fluid and electrolyte balance in, 99
of gallbladder, 1378, 1421–1422
genetic factors in, 320–324, 490
gene therapy and, 344f, 345, 491–492
Knudson’s two-hit hypothesis on, 320, 324f
surgical prevention and, 348–349
in targeted therapies, 342, 343t
global surgery and, 2082–2083
of hand and wrist, 1952–1958
of head and neck, 629–651
of heart, 841–846
heterogeneity of, 319f, 319–320
hormone therapy in, 342
hypercalcemia in, 99, 681–682, 1668
immunotherapy in, 342, 344–345, 491
individualized therapy in, 349–350
intracranial, 1854–1859
invasive, 315
of kidneys, 1772–1773
of liver, 1376–1379
in children, 1751
classification of, 1373t
incidental discovery of, 1372–1373
of lungs, 680–701
mediastinal, 726–736
mesenteric, 1560
mesothelioma, 743–744
metastasis of, 316–317, 317f. See also Metastasis
minimally invasive surgery in, 473–474
mortality to incidence ratio, 2083, 2083f
multidisciplinary approach to, 337
of omentum, 1559
of ovary, 1745–1746, 1815–1820
of pancreas, 1480–1505
perianal, 1304–1305
of peripheral nerves, 1866–1867
pleural, 743–744
pleural effusions in, 737t–738t, 739, 739t
prevention of, 348–349
prognostic and predictive markers in, 334–335
of prostate, 1774–1775
radiation therapy in, 346–348. See also Radiation therapy retroperitoneal, 1582f, 1582–1583
retrorectal/presacral, 1304
risk assessment for, 328–330
screening for, 330–332, 331–332t.
See also Screening serum markers, 335–336
of skin, 527–536
of small intestine, 1241–1245
small intestine obstruction in, 1228
of soft tissues, 528, 723–726, 1567–1590
of spine, 1859–1861, 1861f
staging of, 333–334, 337. See also Staging of tumors of stomach, 1139–1155
surgical approaches to, 305, 337–340
lymph node management in, 337–339
negative surgical margins in, 337
as primary therapy, 305
recent trends and evolving technologies in, 349
targeted therapy in, 342, 343t
of testis, 1771–1772
of thyroid, 1645–1663
of trachea, 665–667
transforming growth factor-beta–
signaling pathway in, 490
in transplant recipients, 364, 369
trends and evolving technology in, 349–350
of vulva, 1809–1810
Tumor lysis syndrome, 99
Tumor necrosis factor, 315
in inflammatory bowel disease, 1282
in injury response, 42
in shock response, 138
in tumor necrosis factor receptors of, 42, 51
TNF1, 40, 41f
Tumor necrosis factor-α-converting enzyme, 42
Tumor necrosis factor–related apoptosis–
inducing ligand (TRAIL), 315
Tumor promoters, 326
Tumor suppressor genes, 310–311, 490
in colorectal cancer, 1289–1290
in hereditary cancer syndromes, 320–324
in soft tissue sarcomas, 1570
Tumorigenesis, 310–311
chemical carcinogens in, 326
as multistep process, 490
physical carcinogens in, 326
proliferation and transformation of
cancer cells in, 490, 490f
in stomach cancer, 1141f
in thyroid cancer, 1645, 1645f, 1646f, 1647–1648
viral carcinogens in, 326, 328, 328t–329t
Tunica albuginea, 1760
traumatic rupture of, 1767
Turcot’s syndrome, 1292
Turner’s syndrome, 544
Tyler’s syndrome, 544
Tylectomy in breast cancer, 591
Tympanic membrane perforation, 615
Tymanoplasty, 615
Type I error, 2147
Type II error, 2147
Typhlitis, 1323
Tyrer-Cuzick model on breast cancer risk, 1647
U
Ultrasonography
in abdominal trauma, 203f
in aortic aneurysm, 858–859
in appendicitis, 1334–1335
in arterial disease, 900
in biliary disorders
description of, 1398, 1398f
endoscopic technique, 1401
in brain death, 1842
in breast cancer diagnosis, 571–572, 572f–574f
in Budd-Chiari syndrome, 1369
cardiac output estimated in, 443
in carotid artery occlusive disease, 909, 909f
in cholelithiasis, 1407
contrast-enhanced, 1357–1358
description of, 984
Doppler technique. See Doppler ultrasonography
endoanal, 1265–1266, 1269, 1269f
endobronchial, in lung cancer, 685
endorectal, 1265–1266, 1266f
in rectal cancer, 1296, 1296f
in esophageal cancer, endoscopic technique, 1070, 1073
in fecal incontinence, 1269, 1269f
in gallbladder cancer, 1422
in gallstones, 1398, 1398f, 1404, 1406f
of hand and wrist, 1932–1933
in inguinal hernia, 1607–1608
laparoscopic, 1413
of liver, 1357–1358, 1358f
elastography, 1358
intraoperative, 1358, 1358f
in lung cancer, 687–688
in lymphedema, 1002
in mesenteric ischemia, 931
in pancreatitis
chronic, 1457–1458, 1458f
endoscopic technique, 1457–1458,
1458f, 1464
in portal hypertension, 1366
prenatal
in diaphragmatic hernia, 1712, 1712f
in gastrointestinal, 1742, 1742f
in ovarian cysts, 1746
in renal artery stenosis, 936, 936f
in soft tissue sarcomas, 1570
of spleen, 1540
of stomach, 1116
endoscopic technique in cancer, 1116, 1145
in thoracic aortic aneurysm, 858–859
of thyroid, 1634, 1635f
in solitary nodule, 1644
in venous insufficiency, 997–998, 998f
in venous thromboembolism, 986, 986f
Ultraviolet radiation
lip cancer associated with, 630
skin damage from, 518–519, 528
cancer associated with, 528
UVA, 518–519, 528
UVB, 518–519, 528
UVC, 518
umbilical hernia, 1554, 1740, 1740f
umbilical ligament, medial, 1602f
umbilical vein, in portal hypertension, 1367
Undersea and Hyperbaric Medical Society,
299
Undescended testis
description of, 1744–1745, 1777
embryology in, 1744
two-stage repair in, 1745
Unfolded protein response, 39–40
United Nations and global surgery, 2088
Universal precautions in blood-borne infections, 177
Unnecessary medical care, 398–399

Unerve ligation, 816

Autografts in, 817

Brunicardi_Index_Vol II.indd   92 22/02/19   11:26 AM
Zion, Libby, 400, 410
Zoledronic acid in bone pain, 2071
Zollinger-Ellison syndrome, 1136–1137, 2142
diagnosis of, 1481
gastric carcinoid tumors in, 1151
in gastrinoma, 1136–1137, 1482
treatment of, 1482
Zona pellucida, 1984
Z-plasty
in burns of hand, 1960, 1960
description of, 1977, 1977
Zuska's disease, 550
Zygoma fractures, 628
Zygomatic fractures, 2004–2005
Zygomaticomaxillary buttress, 2002, 2005
Zygomaticomaxillary complex fractures, 2004–2005
Zymogenic cells, 1105
Zymogens, 1442